# An introduction to the symmetric group algebra

**[Math 701, Spring 2024 lecture notes]**

Darij Grinberg

July 16, 2025 (unfinished!)

**Abstract.** This is an introduction to the group algebras of the symmetric groups, written for a quarter-long graduate course. After recalling the definition of group algebras (and monoid algebras) in general, as well as basic properties of permutations, we introduce several families of elements in the symmetric group algebras $\mathbf{k}\left[S_n\right]$ such as the Young–Jucys–Murphy elements, the (sign-)integrals and the conjugacy class sums. Then comes a chapter on group actions and representations in general, followed by the core of this text: a study of the representations of symmetric groups (i.e., of left $\mathbf{k}\left[S_n\right]$-modules), including the classical theory of Young tableaux and Young symmetrizers. We prove in detail the main facts including the characterization of irreducible representations (in characteristic 0), the Garnir relations, the standard basis theorem, the description of duals of Specht modules, and the hook length formula, as well as a number of less known results. Finally, we describe several bases of $\mathbf{k}\left[S_n\right]$ that arise from the study of Specht modules.

The methods used are elementary and computational. We aim to assume as little as possible of the base ring $\mathbf{k}$, and to use as little as possible from representation theory (nothing more advanced than Maschke and Jordan–Hölder).

Over 100 exercises (without solutions) are scattered through the text.

## Contents

**1. What is this?** **8**

**2. Before we start... 8**
2.1. What is this? . . . 8
2.2. Prerequisites . . . 9
2.3. A note on the exercises . . . 10
2.4. Notations and elementary facts . . . 10
2.4.1. Basics . . . 10
2.4.2. Permutations . . . 11
2.4.3. Algebraic preliminaries . . . 15

**3. Introduction to the symmetric group algebras 16**
3.1. Monoid algebras and group algebras . . . 16
3.1.1. Monoids . . . 16
3.1.2. Monoid algebras . . . 17
3.1.3. Group algebras . . . 19
3.1.4. Some properties . . . 20
3.1.5. Some conventions . . . 21
3.1.6. Examples of monoid algebras . . . 22
3.2. The symmetric group algebra . . . 23
3.3. The integral and the sign-integral . . . 25
3.4. The Young–Jucys–Murphy elements (aka YJM elements, or jucys–murphies) . . . 28
3.4.1. Definition and commutativity . . . 28
3.4.2. A product . . . 30
3.4.3. More about jucys–murphies . . . 36
3.5. Interlude: Subalgebras generated by a set . . . 37
3.5.1. The general case . . . 37
3.5.2. The Gelfand–Tsetlin algebra . . . 39
3.6. The somewhere-to-below shuffles . . . 39
3.7. Partial integrals . . . 40
3.7.1. Definition and properties . . . 40
3.7.2. A recursion . . . 46
3.7.3. The transposition cancellation theorem . . . 50
3.7.4. Disjoint commutativity . . . 51
3.7.5. Products of $X$-integrals . . . 53
3.8. Orbits and cycles . . . 63
3.8.1. Orbits, cycles and reflection lengths . . . 63
3.8.2. Nonstarters and starters . . . 70
3.8.3. Products of distinct jucys–murphies . . . 72
3.9. Conjugacy classes and the center . . . 83
3.9.1. The general case . . . 83
3.9.2. Symmetric groups . . . 92
3.9.3. Centers and jucys–murphies . . . 96
3.10. A potpourri of other elements . . . 98
3.10.1. Descent numbers and the Eulerian subalgebra . . . 98

3.10.2. V-permutations . . . 100
3.10.3. Random-to-random shuffles . . . 103
3.10.4. Excedance and antiexcedance elements . . . 105
3.11. The antipode and the sign-twist . . . 106
3.11.1. Morphisms . . . 107
3.11.2. The sign-twist . . . 107
3.11.3. Anti-morphisms . . . 112
3.11.4. The antipode . . . 113
3.12. Other maps around group algebras . . . 115
3.12.1. The counit . . . 115
3.12.2. Morphisms of monoids yield morphisms of monoid algebras . . . 118
3.12.3. Base change . . . 128
3.13. Antipodal conjugacy . . . 131
3.14. Computer algebra . . . 132
3.14.1. SageMath . . . 132
3.14.2. Permutations in SageMath . . . 133
3.14.3. Symmetric group algebras in SageMath . . . 136
3.14.4. Exercises . . . 139
3.15. A look at the structure of symmetric group algebras . . . 140
3.15.1. $\mathbf{k}[S_2]$ . . . 140
3.15.2. $\mathbf{k}[S_3]$ . . . 142
3.15.3. About the general case . . . 146
3.16. More exercises . . . 147

**4. Actions and representations** **149**
4.1. Group actions and $G$-sets . . . 150
4.1.1. Motivation . . . 150
4.1.2. Definition . . . 151
4.1.3. Examples . . . 152
4.1.4. $G$-sets vs. morphisms $G \to S_X$ . . . 157
4.1.5. $G$-equivariant maps aka $G$-set morphisms . . . 160
4.1.6. Combining $G$-sets . . . 164
4.1.7. $G$-subsets and quotients . . . 166
4.1.8. Restriction . . . 168
4.1.9. Understanding $G$-actions for cyclic groups $G$ . . . 169
4.1.10. Faithfulness . . . 174
4.1.11. Further reading . . . 175
4.2. Linear group actions and representations of $G$ . . . 176
4.2.1. Linear actions, aka representations . . . 176
4.2.2. The curried form . . . 179
4.2.3. Outlook on representation theory . . . 181
4.2.4. Representations of $S_n$: first examples . . . 182
4.2.5. Morphisms of representations . . . 184

4.2.6. Subrepresentations . . . 185
4.2.7. Quotients . . . 188
4.3. Modules over rings, and a different way of thinking about representations . . . 196
4.3.1. Definitions . . . 196
4.3.2. Examples . . . 198
4.3.3. Properties and features . . . 201
4.3.4. Representations of $G$ as left $\mathbf{k}[G]$-modules . . . 202
4.3.5. Irreducible representations . . . 208
4.3.6. Some further properties of representations/modules . . . 210
4.4. Direct addends, projections, filtrations, Maschke, Jordan–Hölder 213
4.4.1. Direct addends and projections . . . 213
4.4.2. Maschke's theorem . . . 221
4.4.3. Filtrations . . . 225
4.4.4. An example: $\mathbf{k}[S_3]$ . . . 230
4.4.5. Artin–Wedderburn . . . 236
4.5. Noncommutative polynomials . . . 238
4.5.1. Definition and symmetric group actions . . . 238
4.5.2. Application: the Dynkin element $\mathbf{V}_n$ . . . 243

**5. Young and Specht modules 250**
5.1. Partitions, compositions, diagrams . . . 251
5.1.1. Definitions and examples . . . 251
5.1.2. Basic properties . . . 255
5.1.3. Some notations regarding partitions . . . 259
5.1.4. Conjugate partitions . . . 260
5.2. Tableaux . . . 263
5.2.1. Definitions and examples . . . 263
5.2.2. First properties . . . 266
5.2.3. The hook length formula . . . 268
5.3. Young modules . . . 271
5.3.1. Horizontal and vertical permutations . . . 271
5.3.2. Row-equivalence and column-equivalence . . . 271
5.3.3. Tabloids . . . 273
5.3.4. The symmetric group acts on tableaux and tabloids . . . 275
5.3.5. Young modules . . . 282
5.4. Specht modules . . . 284
5.4.1. Definition and examples . . . 284
5.4.2. Properties . . . 288
5.4.3. More examples . . . 290
5.5. Symmetrizers, antisymmetrizers and the left ideal avatar . . . 293
5.5.1. Row symmetrizers and column antisymmetrizers . . . 293
5.5.2. The left ideal avatars of Young and Specht modules . . . 301
5.5.3. The row-to-row sums . . . 302

5.5.4. Proving the left ideal avatars . . . . . 305
5.5.5. Explicit formulas for $\alpha^{-1}$ . . . . . 310
5.5.6. A restatement of the left ideal avatar of the Specht module 311
5.6. Determinantal avatars of Young and Specht modules . . . . . 312
5.6.1. The Specht–Vandermonde avatar . . . . . 313
5.6.2. The letterplace avatar . . . . . 330
5.7. The standard polytabloids are linearly independent . . . . . 332
5.7.1. Introduction and main result . . . . . 332
5.7.2. The Young last letter order . . . . . 333
5.7.3. Proof of the linear independence . . . . . 339
5.8. The Garnir relations . . . . . 343
5.8.1. Notations and some harmless lemmas . . . . . 343
5.8.2. Transversals (systems of coset representatives) in groups . 350
5.8.3. The Garnir relations . . . . . 354
5.9. The standard basis theorem . . . . . 359
5.9.1. The theorem . . . . . 359
5.9.2. Column-tabloids . . . . . 360
5.9.3. The column last letter order . . . . . 361
5.9.4. Proof of the standard basis theorem . . . . . 367
5.9.5. Remarks on bad shapes . . . . . 372
5.9.6. The rank of $\mathcal{S}^\lambda$ . . . . . 374
5.10. The Young alternative . . . . . 375
5.10.1. A bit of combinatorics: the Young alternative . . . . . 375
5.10.2. Algebraic consequences: the sandwich lemmas . . . . . 383
5.11. Young symmetrizers . . . . . 389
5.11.1. Definitions . . . . . 389
5.11.2. Quasi-idempotency . . . . . 391
5.11.3. Orthogonality . . . . . 400
5.11.4. The opposite Young symmetrizer . . . . . 405
5.12. Centralized Young symmetrizers . . . . . 409
5.12.1. Definition and centrality . . . . . 409
5.12.2. Proof of the basic claims . . . . . 411
5.12.3. A stronger quasi-idempotence . . . . . 416
5.12.4. A partition of unity into centralized Young symmetrizers 421
5.12.5. A combinatorial consequence: the sum of all $\left(f^\lambda\right)^2$ . . . . 423
5.12.6. Explicit expansions in centralized Young symmetrizers . . 424
5.12.7. The structure of $Z\left(\mathbf{k}\left[S_n\right]\right)$ . . . . . 432
5.13. Irreducibility, non-isomorphism and completeness . . . . . 437
5.13.1. Irreducibility . . . . . 438
5.13.2. Non-isomorphism . . . . . 442
5.13.3. Completeness . . . . . 445
5.14. The Artin–Wedderburn theorem . . . . . 449
5.14.1. The theorem . . . . . 449
5.14.2. Linear algebra lemmas . . . . . 449

5.14.3. The proof . . . 451
5.14.4. Proof of Theorem 3.15.1 **(a)** . . . 455
5.14.5. The kernel and the surjectivity of $\rho_\lambda : \mathbf{k}[S_n] \to \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ 456
5.15. The Young natural basis of $\mathbf{k}[S_n]$ . . . 459
5.15.1. The $w_{P,Q}$ and the $\mathbf{E}_{P,Q}$ . . . 459
5.15.2. Transposes of Young tableaux . . . 464
5.15.3. Straightening left and right . . . 470
5.15.4. Decomposing $\mathcal{A}$ and $\mathcal{A}\mathbf{E}_\lambda$ as left $\mathcal{A}$-modules . . . 475
5.15.5. The Young natural bases of $\mathcal{A}\mathbf{E}_\lambda$ and $\mathcal{A}$ . . . 482
5.15.6. Proof of Theorem 5.14.9 . . . 485
5.16. The action of conjugacy class sums on Specht modules . . . 489
5.16.1. A general formula . . . 489
5.16.2. The sum of all transpositions . . . 493
5.17. Characters . . . 500
5.17.1. Definition of a character . . . 500
5.17.2. Examples . . . 503
5.17.3. Characters and isomorphism . . . 508
5.17.4. The central element corresponding to a character . . . 509
5.17.5. Quotients and direct sums . . . 511
5.17.6. The quasiidempotent character formula . . . 520
5.17.7. Characters and centralized Young symmetrizers . . . 526
5.17.8. Characters of small symmetric groups . . . 533
5.17.9. An explicit formula for the character of $\mathcal{S}^\lambda$ . . . 535
5.17.10. An outlook on the Murnaghan–Nakayama rule . . . 539
5.17.11. Characters uniquely determine $S_n$-representations . . . 542
5.18. The sign-twist of an $S_n$-representation . . . 544
5.18.1. The sign-twist in general . . . 544
5.18.2. The sign-twist of a Specht module . . . 553
5.19. The dual of an $S_n$-representation . . . 557
5.19.1. Reminders on dual $\mathbf{k}$-modules . . . 557
5.19.2. Bilinear forms . . . 559
5.19.3. The dual of a group representation . . . 564
5.19.4. $G$-invariant bilinear forms and duals . . . 575
5.19.5. The dual of a Specht module: The case of a field . . . 581
5.19.6. The dual of a Specht module: The case of a skew Young diagram . . . 587
5.19.7. Consequences in characteristic 0 . . . 595
5.19.8. Characters of duals . . . 597
5.19.9. Proving Theorem 5.18.16 . . . 597
5.20. Application: Proof of antipodal conjugacy . . . 598
5.20.1. Conjugacy and the Taussky–Zassenhaus theorem . . . 598
5.20.2. Using Artin–Wedderburn . . . 600
5.20.3. The proof . . . 602

5.21. The $\mathbf{F}_U w_{U,V} \mathbf{E}_V$ basis . . . . . 603
5.21.1. The basis . . . . . 604
5.21.2. Proof of the basis property . . . . . 605
5.21.3. A combinatorial consequence: the sum of all $f^\lambda$ . . . . . 612
5.22. The quotient presentation of a Specht module . . . . . 619
5.22.1. Definitions and theorems . . . . . 619
5.22.2. Proofs . . . . . 629
5.23. Linear equations defining the Specht module . . . . . 638
5.23.1. The coGarnir equations . . . . . 638
5.23.2. Transposing the Garnir relations . . . . . 641
5.23.3. Proof of the coGarnir equations: Correctness . . . . . 645
5.23.4. Proof of the coGarnir equations: Sufficiency . . . . . 647
5.23.5. Triangulating row-Garnir sets . . . . . 656
5.24. Appendix: Young's proof of the hook length formula . . . . . 661
5.24.1. A consequence of the Garnir relations . . . . . 661
5.24.2. Arms, legs and hooks . . . . . 669
5.24.3. A few last lemmas . . . . . 672
5.24.4. Proof of the hook length formula . . . . . 675

**6. Bases of $\mathbf{k}[S_n]$** **694**
6.1. Notations . . . . . 695
6.2. What to do with a basis . . . . . 696
6.2.1. The size of a basis . . . . . 696
6.2.2. The matrix representing a linear map . . . . . 697
6.2.3. The endomorphisms $L(\mathbf{a})$ and $R(\mathbf{a})$ and the matrices $L^{\rightarrow}_{\mathbf{b}}(\mathbf{a})$ and $R^{\rightarrow}_{\mathbf{b}}(\mathbf{a})$ . . . . . 697
6.2.4. Sub-spans . . . . . 700
6.3. The standard basis of $\mathbf{k}[S_n]$, revisited . . . . . 700
6.4. The Young natural basis of $\mathbf{k}[S_n]$, revisited . . . . . 702
6.5. The $\mathbf{F}_U w_{U,V} \mathbf{E}_V$ basis of $\mathbf{k}[S_n]$, revisited . . . . . 707
6.6. The Artin–Wedderburn–Specht bases . . . . . 709
6.7. Combinatorial interlude: Lexicographic orders . . . . . 711
6.7.1. Partitions . . . . . 711
6.7.2. Standard tableaux of partition shapes . . . . . 713
6.7.3. Standard $n$-bitableaux . . . . . 717
6.8. The Murphy bases . . . . . 719
6.8.1. Introductory example: $n = 3$ . . . . . 719
6.8.2. Definitions . . . . . 721
6.8.3. Basic identities . . . . . 724
6.8.4. The basis theorem . . . . . 731
6.8.5. The dot product on $\mathcal{A}$ . . . . . 732
6.8.6. Some lemmas from linear algebra . . . . . 735
6.8.7. Proof of the basis property . . . . . 739
6.8.8. The Murphy spans . . . . . 743

6.8.9. The Murphy span filtrations . . . . . 757

**A. Omitted details and proofs** **768**
A.1. Tableaux . . . . . 768
A.2. The Young natural basis of $\mathbf{k}[S_n]$ . . . . . 769
A.3. The action of conjugacy class sums on Specht modules . . . . . 771
A.4. What to do with a basis . . . . . 775
A.5. Combinatorial interlude: Lexicographic orders . . . . . 776
A.6. The Murphy bases . . . . . 785

This work is licensed under a Creative Commons "CC0 1.0 Universal" license.

# 1. What is this?

These are the notes for a graduate-level course on the group algebras of the symmetric groups. In the present state, they tell perhaps half of what I wanted to tell, but this half is in a finished state. More will be added as I find the time.

## Acknowledgments

I thank Daniel Harrer, Victor Reiner, Franco Saliola and Jessica Tomasko for interesting comments and corrections. Questions, corrections and suggestions are welcome!

# 2. Before we start...

## 2.1. What is this?

This is a course on the *group algebras of the symmetric groups*: one of the most famous and feature-rich objects in algebraic combinatorics.

These group algebras have been studied since the early 20th Century, under the name "substitutional analysis" (see [Ruther48] for a textbook treatment on those early results). Here, "substitution" is an antiquated word for "permutation", whereas "analysis" is used in the literal sense (analyzing things, not limits and derivatives).

Originally, this was used as an abstraction for what happens if you permute the arguments of a multivariate function. Given a function $f : \mathbb{R}^n \to \mathbb{R}$, we can obtain $n!$ other functions $f_\sigma : \mathbb{R}^n \to \mathbb{R}$, one for each permutation $\sigma$ of $\{1, 2, \ldots, n\}$, by permuting the arguments of $f$, as follows:

$$f_\sigma(x_1, x_2, \ldots, x_n) = f\left(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}\right).$$

Then, we can take linear combinations of these $n!$ many functions, for example the symmetrization $\dfrac{1}{n!}\sum\limits_{\sigma\in S_n} f_\sigma$. The input set $\mathbb{R}^n$ could be replaced by $I^n$ for any set $I$; this includes "tensors" in the vulgar meaning of the word ("higher-dimensional matrices"), as well as the multilinear maps known from linear algebra. In particular, the study of "symmetries of tensors" in the 1950s (by Hermann Weyl and others) rests on the group algebras of the symmetric groups.

We will study these group algebras abstractly – meaning that we won't mainly be applying them to specific functions $f : \mathbb{R}^n \to \mathbb{R}$ but we will be computing formally with the permutations. This is similar to studying polynomials without evaluating them at specific points, or studying permutations as "abstract" maps rather than using them to permute a given tuple.

Permutations are indeed the backbone of this theory. In a way, symmetric group algebras are to permutations as polynomial rings are to monomials. An element of the group algebra of a symmetric group is a "formal" linear combination of the permutations in this group, where the coefficients come from a given commutative ring $\mathbf{k}$. These kinds of combinations can be taken for any group, not just a group of permutations, but we will say only few things about this general case.

This is **not** a course on representation theory. But we will cover, among other things, the main results of the representation theory of symmetric groups over fields of characteristic 0.

## 2.2. Prerequisites

We will use

- some abstract algebra: some basic group theory (Lagrange's theorem, cosets; see [McClea23], [DumFoo04, Chapters 1–4] or [Conrad]) and some theory of rings and modules (as covered in [23wa]);
- some combinatorics of permutations (as covered in [21s, Chapter 5] and to some extent in [22fco]);
- some theory of determinants ([21s, Chapter 6]) and characteristic polynomials (the Cayley–Hamilton theorem; see, e.g., [StoLui19, Chapter 3]).

No representation theory is required to read these notes; what little of it will be used will be introduced in Chapter 4.

We will be covering both classical and fairly recent research, overlapping with the purview of several textbooks. I can recommend [Wildon18], [Sagan01], [Meliot17], [Fulton97], [JamKer81], [Ruther48], [Kerber99], [Lorenz18], [EGHetc11] in particular (further sources will be cited in the respective chapters and sections). However, we will not follow any source for too long, and we will try to approach the material in somewhat novel ways.

## 2.3. A note on the exercises

All exercises in these notes are equipped with a difficulty rating. This is the number in the square (such as $\boxed{3}$) placed at the beginning of the exercise. The harder the exercise, the larger this number is (although the rating has been artificially lowered for some exercises that are too tangential to the subject at hand). For example, a $\boxed{5}$ represents a good graduate-level homework problem that requires thinking and work. A $\boxed{3}$ usually requires some thinking **or** work. A $\boxed{1}$ is a warm-up question that might be obvious after 15 seconds of thought. A $\boxed{7}$ should be somewhat too hard for homework. Exercises with ratings above $\boxed{10}$ are hard research results on which I would enjoy hearing your ideas but don't really expect a solution. The rating $\boxed{\text{open}}$ represents an open problem (conjecture or question). Multi-part exercises sometimes have points split between the parts.

When I teach from these notes, I assign the exercises as homework. By solving exercises, students gain points towards their total score. The points gained from a given exercise equal its difficulty rating (or less when the solution is incomplete or flawed). In Spring 2024, the requirement for a perfect grade (A+) was to gain at least 50 points in total. In hindsight, it would have probably been better to break the text into homework sets and require (say) at least 10 points from each homework set.

In solving an exercise, you can freely use (without proof) the claims of all exercises above it.

## 2.4. Notations and elementary facts

Before we properly begin, let me list some notations and basic facts that we will be using.

### 2.4.1. Basics

First, we introduce some notations and terminology for basic combinatorial concepts:

- The symbol $\mathbb{N}$ will denote the set $\{0, 1, 2, 3, \ldots\}$ of nonnegative integers.
- The size (i.e., cardinality) of a set $A$ will be denoted by $|A|$.
- The symbol "#" means "number". For example, the size $|A|$ of a set $A$ is the # of elements of $A$.
- For any integer $k$, we let $[k]$ denote the set $\{1, 2, \ldots, k\}$. This is a $k$-element set if $k \geq 0$, and is the empty set if $k \leq 0$.

### 2.4.2. Permutations

Permutations are among the main characters of this course, and unsurprisingly we will use some of their properties. Here we collect a few basic ones; we refer to [21s, Chapter 5], [Grinbe15, Chapter 5], [22fco, Chapter 4] for proofs and details.

- A *permutation* of a set $X$ means a bijection (i.e., bijective map) from $X$ to $X$. Permutations are composed like any functions: If $\alpha$ and $\beta$ are two permutations of $X$, then

$$(\alpha \circ \beta)(x) = \alpha(\beta(x)) \qquad \text{for any } x \in X.$$

  (Beware that some authors – particularly British ones – disagree.)

  The composition $\alpha \circ \beta$ is again a permutation of $X$.

- The *inverse* of a permutation $\alpha \in X$ is the inverse $\alpha^{-1}$ of the map $\alpha$. It is again a permutation of $X$.

- The identity map $\mathrm{id}_X : X \to X$ (often just called id) is the map sending every $x$ to $x$. It is also a permutation of $X$.

- For a set $X$, the *symmetric group* $S_X$ is the group of all permutations of $X$, where the group operation is composition (so that $\alpha\beta = \alpha \circ \beta$ for any $\alpha, \beta \in S_X$). Its size is $|S_X| = |X|!$ (when $X$ is finite).

  (Alternative notations for $S_X$ are $\mathfrak{S}_X$ and $\Sigma_X$ and $\mathcal{S}_X$ and $\operatorname{Aut} X$. The last one is justified by the fact that the permutations are the automorphisms in the category of sets. The other three are justified by the lack of letters that authors often experience in combinatorics.)

  Since $S_X$ is a group, we shall use the notation $\alpha\beta$ as a synonym for $\alpha \circ \beta$, and we shall use the notation 1 as a synonym for $\mathrm{id}_X$ (the identity map of $X$, aka the neutral element of $S_X$).

- For a given $n \in \mathbb{N}$, the symmetric group $S_{[n]} = S_{\{1,2,\ldots,n\}}$ is denoted by $S_n$. This is called the *n-th symmetric group*, and has size $|S_n| = n!$. (Again, some authors call it $\mathfrak{S}_n$ or $\Sigma_n$ or $\mathcal{S}_n$.)

- The easiest way to write down a permutation in $S_n$ is using its *one-line notation*.

  The *one-line notation* (short: *OLN*) of a permutation $\sigma \in S_n$ is the $n$-tuple $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. We denote it by $\operatorname{OLN}\sigma$. Thus,

$$\operatorname{OLN}\sigma = (\sigma(1), \sigma(2), \ldots, \sigma(n)).$$

  Sometimes we omit the parentheses and the commas in the OLN. For example, if $\sigma \in S_3$ is the permutation that sends $1, 2, 3$ to $3, 1, 2$, then

$\operatorname{OLN}\sigma = (3,1,2)$, and we abbreviate it as 312. (This kind of abbreviation becomes ambiguous when $\sigma \in S_n$ for $n > 10$, but fortunately such large numbers are rarely used in examples.)

Conversely, if $i_1, i_2, \ldots, i_n$ are $n$ distinct elements of $[n]$, then there exists a unique permutation $\sigma \in S_n$ with $\operatorname{OLN}\sigma = (i_1, i_2, \ldots, i_n)$. This permutation $\sigma$ is the permutation of $[n]$ that sends $1, 2, \ldots, n$ to $i_1, i_2, \ldots, i_n$, respectively. It will be denoted by $\operatorname{oln}(i_1, i_2, \ldots, i_n)$. Sometimes, for the sake of brevity, we omit the commas in this notation, so that we just write $\operatorname{oln}(i_1 i_2 \ldots i_n)$.

- Some of the most widely used permutations are the *cycles*. If $i_1, i_2, \ldots, i_k$ are $k$ distinct elements of a set $X$, then the *cycle* $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ denotes the permutation of $X$ that sends $i_1$ to $i_2$, sends $i_2$ to $i_3$, sends $i_3$ to $i_4$, and so on, sends $i_k$ to $i_1$, and leaves all the remaining elements of $X$ unchanged. In other words, it is defined by

$$\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(x) = \begin{cases} i_{h+1}, & \text{if } x = i_h \text{ for some } h \in [k]; \\ x, & \text{if not} \end{cases} \qquad \text{for each } x \in X,$$

where we set $i_{k+1} := i_1$.

(Many authors write $(i_1, i_2, \ldots, i_k)$ for $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$. Be careful if you do this, because the notation $(i_1, i_2, \ldots, i_k)$ already means a $k$-tuple, and thus you are creating an ambiguity. And it can bite: $(1,2,3)$ and $(2,3,1)$ are not the same 3-tuple, but $\operatorname{cyc}_{1,2,3}$ and $\operatorname{cyc}_{2,3,1}$ are the same cycle.)

Note that $\operatorname{cyc}_i = \operatorname{id}$ for any $i \in X$. See [21s, §5.2.2 and §5.5] and [22fco, §4.2 and §4.3] for more about cycles.

Strictly speaking, the cycle $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ depends not only on the elements $i_1, i_2, \ldots, i_k$, but also on the set $X$, even though the notation $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ neglects to mention this. In practice, this is rarely problematic, since the set $X$ is usually clear from the context or irrelevant. (It helps that $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(x) = x$ for all $x \notin \{i_1, i_2, \ldots, i_k\}$.)

- If $i$ and $j$ are two distinct elements of $X$, then we set

$$t_{i,j} := \operatorname{cyc}_{i,j} \in S_X. \tag{1}$$

This is called the *transposition* that swaps $i$ and $j$. It is the permutation of $X$ that swaps $i$ with $j$ and leaves the other elements unchanged.

Note that $t_{i,j} = t_{j,i}$ for any distinct $i$ and $j$. More generally,

$$\operatorname{cyc}_{i_1,i_2,\ldots,i_k} = \operatorname{cyc}_{i_2,i_3,\ldots,i_k,i_1} = \operatorname{cyc}_{i_3,i_4,\ldots,i_k,i_1,i_2} = \cdots = \operatorname{cyc}_{i_k,i_1,i_2,\ldots,i_{k-1}}$$

whenever $i_1, i_2, \ldots, i_k$ are $k$ distinct elements of $X$.

- Specifically for $X = [n]$, there is a specific sort of transpositions known as the *simple transpositions*:

  For any $n \in \mathbb{N}$ and any $i \in [n-1]$, we define

  $$s_i := t_{i,i+1} \in S_n. \tag{2}$$

  This is the transposition that swaps $i$ with $i+1$. It is called a *simple transposition*.

  An important fact (see [21s, Theorem 5.3.17 **(a)**] or most good courses on group theory) is that every permutation in $S_n$ can be written as a product of simple transpositions.

- If $\phi : X \to Y$ is a bijection between two sets $X$ and $Y$, then there is a group isomorphism

  $$\begin{aligned} S_\phi : S_X &\to S_Y, \\ \alpha &\mapsto \phi \circ \alpha \circ \phi^{-1}. \end{aligned} \tag{3}$$

  Essentially, it just transforms every permutation of $X$ into a permutation of $Y$ by "relabelling" each $x \in X$ as $\phi(x) \in Y$.

  In particular, if $X$ is a finite set of size $n \in \mathbb{N}$, then we can pick any bijection $\phi : X \to [n]$ and thus obtain a group isomorphism $S_\phi : S_X \to S_n$. This is why it is often sufficient to study $S_n$ when you want to understand $S_X$ for arbitrary finite sets $X$.

- For any set $X$, any cycle $\operatorname{cyc}_{i_1,i_2,\ldots,i_k} \in S_X$ and any permutation $\sigma \in S_X$, we have

  $$\sigma \circ \operatorname{cyc}_{i_1,i_2,\ldots,i_k} \circ \sigma^{-1} = \operatorname{cyc}_{\sigma(i_1),\sigma(i_2),\ldots,\sigma(i_k)} \tag{4}$$

  (see, e.g., [Grinbe15, Exercise 5.17 **(a)**] for the easy proof[1]). In particular,

  $$\sigma \circ t_{i,j} \circ \sigma^{-1} = t_{\sigma(i),\sigma(j)} \tag{5}$$

  for any transposition $t_{i,j} \in S_X$ and any permutation $\sigma \in S_X$.

- Each cycle is a product (i.e., composition) of transpositions: Indeed, for any set $X$ and any cycle $\operatorname{cyc}_{i_1,i_2,\ldots,i_k} \in S_X$, we have

  $$\operatorname{cyc}_{i_1,i_2,\ldots,i_k} = t_{i_1,i_2} \circ t_{i_2,i_3} \circ \cdots \circ t_{i_{k-1},i_k}. \tag{6}$$

  (The right hand side is understood to be $\mathrm{id}_X$ if $k = 1$. Generally, an empty product in a group is understood to be the neutral element of the group.)

  For a proof of this of (6), see [Grinbe15, Exercise 5.16][2].

[1] Note that [Grinbe15, Exercise 5.17 **(a)**] is only stated for the case $X = [n]$, but the proof works equally for any $X$.

[2] Again, [Grinbe15, Exercise 5.16] is only stated for $X = [n]$, but the proof applies equally well to all $X$.

- For any set $X$ and any $m$ distinct elements $i_1, i_2, \ldots, i_m$ of $X$, we have

$$\operatorname{cyc}_{i_1,i_2,\ldots,i_m} = \operatorname{cyc}_{i_1,i_2,\ldots,i_k} \circ \operatorname{cyc}_{i_k,i_{k+1},\ldots,i_m} \tag{7}$$

  for all $1 \leq k \leq m$. This is not hard to check, either by expressing all three cycles as products of transpositions using (6), or by directly comparing how the two sides act on each $x \in X$.

- Disjoint cycles commute: If $i_1, i_2, \ldots, i_k$ and $j_1, j_2, \ldots, j_\ell$ are $k + \ell$ distinct elements of $X$, then

$$\operatorname{cyc}_{i_1,i_2,\ldots,i_k} \circ \operatorname{cyc}_{j_1,j_2,\ldots,j_\ell} = \operatorname{cyc}_{j_1,j_2,\ldots,j_\ell} \circ \operatorname{cyc}_{i_1,i_2,\ldots,i_k}. \tag{8}$$

  This is easy to check directly by comparing the action on each $x \in X$.

  Note that the converse is not true: Non-disjoint cycles can sometimes also commute. In other words, (8) occasionally holds even if the $k + \ell$ elements $i_1, i_2, \ldots, i_k$ and $j_1, j_2, \ldots, j_\ell$ are not all distinct.

- To invert a cycle, we just read its indices right to left: In other words, for any cycle $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$, we have

$$\left(\operatorname{cyc}_{i_1,i_2,\ldots,i_k}\right)^{-1} = \operatorname{cyc}_{i_k,i_{k-1},\ldots,i_1}. \tag{9}$$

- Let $X$ be a finite set. The *sign* $(-1)^\sigma$ of a permutation $\sigma \in S_X$ is a number that is either 1 or $-1$, and can be defined in the following two equivalent ways:

  1. If $\sigma$ is a product of an even number of transpositions, then $(-1)^\sigma = 1$. If $\sigma$ is a product of an odd number of transpositions, then $(-1)^\sigma = -1$.
  2. Alternatively, we can define signs as follows: The sign of a permutation $\sigma \in S_n$ (for some $n \in \mathbb{N}$) is defined to be $(-1)^{\ell(\sigma)}$, where $\ell(\sigma)$ is the # of inversions of $\sigma$ (that is, of all pairs $(i, j) \in [n]^2$ such that $i < j$ and $\sigma(i) > \sigma(j)$). The sign of a permutation $\sigma \in S_X$ for an arbitrary finite set $X$ is defined by mapping $S_X$ to $S_n$ for $n = |X|$ using the above group isomorphism $S_\phi$ defined in (3), and computing the sign there.

  For details and proofs, see [Grinbe15, §5.3 and §5.6] (and [21s, §5.4] for a sketch). In particular, the second definition is the one given in [Grinbe15, Definition 5.14 and Exercise 5.12], whereas the first definition is equivalent to it by [Grinbe15, Exercise 5.15 **(c)**].

For us, the most important properties of signs are

$$(-1)^{\sigma\tau} = (-1)^{\sigma} \cdot (-1)^{\tau} \qquad \text{for any } \sigma, \tau \in S_X;$$
$$(-1)^{t_{i,j}} = -1 \qquad \text{for any distinct } i, j \in X;$$
$$(-1)^{\operatorname{cyc}_{i_1,i_2,\ldots,i_k}} = (-1)^{k-1} \qquad \text{for any cycle } \operatorname{cyc}_{i_1,i_2,\ldots,i_k} \in S_X;$$
$$(-1)^{\sigma^{-1}} = (-1)^{\sigma} \qquad \text{for any } \sigma \in S_X.$$

For the proofs, see [Grinbe15, Exercise 5.12 **(c)** and Exercise 5.15 **(a)**] as well as (6). The identity $(-1)^{\sigma\tau} = (-1)^{\sigma} \cdot (-1)^{\tau}$ is known as the *multiplicativity of the sign*, and shows that the *sign homomorphism* (i.e., the map $S_X \to \{1, -1\}$ that sends each $\sigma \in S_X$ to $(-1)^{\sigma}$) is a group morphism. An easy consequence of these properties is that conjugate permutations have equal signs – i.e., that

$$(-1)^{\sigma\tau\sigma^{-1}} = (-1)^{\tau} \qquad \text{for all } \sigma, \tau \in S_X. \tag{10}$$

- A permutation $\sigma \in S_X$ of a finite set $X$ is said to be *even* if its sign $(-1)^{\sigma}$ equals 1, and it is said to be *odd* if its sign $(-1)^{\sigma}$ equals $-1$.

### 2.4.3. Algebraic preliminaries

Next, we recall some concepts from abstract algebra. We will freely use the notion of a ring, and the notions of modules and algebras over a commutative ring. See [23wa, Chapters 2 and 3] for a detailed discussion of these notions, and [21s, §3.4.2] for a quick reminder. The word "ring", for us, always means an associative unital ring (but not necessarily commutative).

We fix a commutative ring $\mathbf{k}$ from now on. It will be called our *base ring*, and its elements will be called *scalars*. We recall the basic ideas behind modules and algebras:

- A $\mathbf{k}$-*module* means an additive abelian group whose elements can be scaled by elements of $\mathbf{k}$ and satisfy a few simple axioms. These axioms are precisely the vector space axioms known from linear algebra, and indeed the $\mathbf{k}$-modules are also known as "$\mathbf{k}$-vector spaces" when $\mathbf{k}$ is a field. But we will allow $\mathbf{k}$ to be an arbitrary commutative ring, in which case it is frowned upon to speak of "$\mathbf{k}$-vector spaces".

  The elements of a $\mathbf{k}$-module will nevertheless be called *vectors*.

  Just like in linear algebra, the simplest examples of $\mathbf{k}$-modules are $\mathbf{k}$ itself, its Cartesian powers $\mathbf{k}^n$ (consisting of $n$-tuples of elements of $\mathbf{k}$, added and scaled entrywise), its "matrix spaces" $\mathbf{k}^{n \times m}$ (consisting of $n \times m$-matrices with entries in $\mathbf{k}$), as well as various submodules thereof (such as the submodule of $\mathbf{k}^n$ that consists of all $n$-tuples $(a_1, a_2, \ldots, a_n)$ satisfying $a_1 + a_2 + \cdots + a_n = 0$).

- A $\mathbf{k}$-*algebra* means a $\mathbf{k}$-module $A$ that is also a ring (with the same addition and zero), with an extra compatibility requirement saying that

$$\lambda (ab) = (\lambda a) b = a (\lambda b) \qquad \text{for any } \lambda \in \mathbf{k} \text{ and } a, b \in A.$$

  For example, $\mathbf{k}$ itself is a $\mathbf{k}$-algebra, and so are the polynomial ring $\mathbf{k}[x]$ and the matrix ring $\mathbf{k}^{n \times n}$ for any $n \in \mathbb{N}$.

- The word "*morphism*" is short for "homomorphism".

- The word "*ideal*" means "two-sided ideal", unless it is part of the phrasing "left ideal" or "right ideal".

# 3. Introduction to the symmetric group algebras

## 3.1. Monoid algebras and group algebras

### 3.1.1. Monoids

Next, we recall the notion of a monoid.

Informally, a *monoid* is a "group without inverses".

Rigorously, a *monoid* is defined to be a triple $(M, \cdot, 1)$, where $M$ is a set, $\cdot$ is an associative binary operation on $M$ (that is, a map from $M \times M$ to $M$), and $1$ is an element of $M$ that is neutral for this operation $\cdot$. When $M$ is a monoid, we will write $mn$ for $m \cdot n$ whenever $m, n \in M$. The element $mn$ will be called the *product* of $m$ and $n$ in the monoid $M$. We will write $M$ for $(M, \cdot, 1)$ if the operation $\cdot$ and the element $1$ are clear from the context. The monoid $M$ is said to be *abelian* if all $m, n \in M$ satisfy $mn = nm$.

The notations $\cdot$ and $1$ are not set in stone; some authors use $*$ and $e$ for them. Depending on the monoid, they might have different names. For example, the most fundamental example of a monoid is the abelian monoid $(\mathbb{N}, +, 0)$. Here, the operation is actually the addition $+$ of the nonnegative integers, so it would be a bad idea to rename it as $\cdot$. Likewise, its neutral element is the number $0$, not the number $1$.

When we use the notations $\cdot$ and $1$ for the operation and the neutral element of a monoid, we say that our monoid is *written multiplicatively*.

Here are some examples of monoids:

- Any group is a monoid. For instance, $(\mathbb{Z}, +, 0)$ is a monoid, as is $(\mathbb{Z}, \cdot, 1)$. Actually, $(\mathbb{Z}, +, 0)$ is a group, but $(\mathbb{Z}, \cdot, 1)$ is not.

- More generally, if $R$ is any ring, then $(R, +, 0)$ and $(R, \cdot, 1)$ are monoids (where $+$ and $\cdot$ are the operations of $R$, and where $0$ and $1$ are the zero and the unity of $R$). The monoid $(R, +, 0)$ is always an abelian group, but the monoid $(R, \cdot, 1)$ is (in general) neither abelian nor a group.

- As we already mentioned, $(\mathbb{N}, +, 0)$ is a monoid as well (but not a group). So is $(\mathbb{N}, \cdot, 1)$.
- The Boolean monoid is defined to be the monoid $(\mathbb{B}, \vee, \mathrm{False})$, where $\mathbb{B} = \{\mathrm{True}, \mathrm{False}\}$.
- For any positive integer $n$, the triple $([n], \max, 1)$ is a monoid, where max is the binary operation sending each pair $(u, v) \in [n] \times [n]$ to $\max\{u, v\}$. The associativity follows from the fact that all $a, b, c \in [n]$ satisfy $\max\{a, \max\{b, c\}\} = \max\{\max\{a, b\}, c\}$, whereas the neutrality of 1 follows from the fact that each $u \in [n]$ satisfies $\max\{1, u\} = \max\{u, 1\} = u$.
- We can extend the set $\mathbb{Z}$ to a slightly larger set $\mathbb{Z}^* := \mathbb{Z} \cup \{-\infty\}$ by inserting a new element $-\infty$. We understand this new element $-\infty$ to be smaller than every integer, so that $\max\{k, -\infty\} = \max\{-\infty, k\} = k$ for all $k \in \mathbb{Z}^*$. Now, the triple $(\mathbb{Z}^*, \max, -\infty)$ is a monoid (where max is the binary operation sending each pair $(u, v) \in \mathbb{Z}^* \times \mathbb{Z}^*$ to $\max\{u, v\}$). It is called the *tropical monoid* on $\mathbb{Z}$. Similarly we can define a tropical monoid on every totally ordered set.

The neutral element of any monoid $M$ will be called $1_M$.

A *monoid morphism* is a map $f : M \to N$ between two monoids $M$ and $N$ that respects the neutral elements (i.e., satisfies $f(1_M) = 1_N$) and respects multiplication (i.e., satisfies $f(mm') = f(m) \cdot f(m')$ for any $m, m' \in M$).

### 3.1.2. Monoid algebras

Next, we shall define the "monoid algebra" of a monoid. For details, we refer to [23wa, §4.1].

Informally, the *monoid algebra* of a monoid $M$ over a commutative ring $\mathbf{k}$ is the $\mathbf{k}$-algebra obtained by "adjoining" the monoid $M$ to the ring $\mathbf{k}$, which means "inserting" the elements of $M$ into $\mathbf{k}$. That is, this algebra consists of "formal products" $\lambda m$ with $\lambda \in \mathbf{k}$ and $m \in M$, as well as their "formal sums" $\lambda_1 m_1 + \lambda_2 m_2 + \cdots + \lambda_i m_i$, which are called "formal $\mathbf{k}$-linear combinations of elements of $M$". These $\mathbf{k}$-linear combinations are multiplied using the multiplications of $\mathbf{k}$ and $M$: For example, if $\lambda_1 m_1 + \lambda_2 m_2$ and $\lambda_3 m_3 + \lambda_4 m_4$ are two such $\mathbf{k}$-linear combinations, then their product is

$$\begin{aligned}&(\lambda_1 m_1 + \lambda_2 m_2)(\lambda_3 m_3 + \lambda_4 m_4) \\ &= \underbrace{\lambda_1 \lambda_3}_{\substack{\text{a product} \\ \text{in } \mathbf{k}}} \underbrace{m_1 m_3}_{\substack{\text{a product} \\ \text{in } M}} + \lambda_1 \lambda_4 m_1 m_4 + \lambda_2 \lambda_3 m_2 m_3 + \lambda_2 \lambda_4 m_2 m_4.\end{aligned}$$

Let us now define this monoid algebra formally.

First, we define the notion of the *free* $\mathbf{k}$*-module* $\mathbf{k}^{(S)}$ for a given set $S$:

**Definition 3.1.1.** Let $S$ be an arbitrary set, and (as before) let $\mathbf{k}$ be a commutative ring.

As usual, $\mathbf{k}^S$ denotes the set of all families $(\lambda_s)_{s\in S}$ of elements of $\mathbf{k}$ indexed by elements of $S$. For example, if $S = [n]$ for some $n \in \mathbb{N}$, then such families are just the $n$-tuples $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ of elements of $\mathbf{k}$, so that $\mathbf{k}^{[n]} = \mathbf{k}^n$. For another example, if $S = \mathbb{N}$, then such families are the sequences $(\lambda_0, \lambda_1, \lambda_2, \ldots)$.

A family $(\lambda_s)_{s\in S} \in \mathbf{k}^S$ is said to be *essentially finite* if all but finitely many $s \in S$ satisfy $\lambda_s = 0$ (that is, if it has only finitely many nonzero entries).

We let $\mathbf{k}^{(S)}$ denote the set of all essentially finite families in $\mathbf{k}^S$. Thus,

$$\mathbf{k}^{(S)} = \left\{ (\lambda_s)_{s\in S} \in \mathbf{k}^S \ \mid \ \text{all but finitely many } s \in S \text{ satisfy } \lambda_s = 0 \right\}.$$

The set $\mathbf{k}^S$ is a $\mathbf{k}$-module, where addition and scaling are defined entry-wise:

$$(a_s)_{s\in S} + (b_s)_{s\in S} := (a_s + b_s)_{s\in S} \qquad \text{for any } (a_s)_{s\in S}, (b_s)_{s\in S} \in \mathbf{k}^S;$$
$$\lambda\,(a_s)_{s\in S} := (\lambda a_s)_{s\in S} \qquad \text{for any } \lambda \in \mathbf{k} \text{ and } (a_s)_{s\in S} \in \mathbf{k}^S.$$

The subset $\mathbf{k}^{(S)}$ of $\mathbf{k}^S$ is a $\mathbf{k}$-submodule of $\mathbf{k}^S$, because the union of two finite sets is still finite. This submodule $\mathbf{k}^{(S)}$ is called the *free* $\mathbf{k}$*-module on the set* $S$. (We often omit the $\mathbf{k}$, and just say "free module".)

This $\mathbf{k}$-module $\mathbf{k}^{(S)}$ is indeed free. Indeed, we can construct a basis (called the *standard basis*) as follows: For each $t \in S$, we define the *standard basis vector* $e_t$ to be the family $(\delta_{s,t})_{s\in S} \in \mathbf{k}^S$, where $\delta_{s,t}$ is the Kronecker delta (that is, 1 if $s = t$ and 0 otherwise). This standard basis vector $e_t$ has only one nonzero entry (namely, its $t$-th entry, which is 1), and thus belongs to $\mathbf{k}^{(S)}$. The family $(e_t)_{t\in S}$ is a basis of $\mathbf{k}^{(S)}$, called the *standard basis*.

For $S = [n]$, this all is standard linear algebra: Here, $e_t = (0, 0, \ldots, 0, 1, 0, 0, \ldots, 0)$ with the 1 in the $t$-th position, and thus the standard basis is really the standard basis you know from linear algebra.

Here are a few more remarks about Definition 3.1.1:

- We have $\mathbf{k}^{(S)} = \mathbf{k}^S$ when the set $S$ is finite. When $S$ is infinite, $\mathbf{k}^{(S)}$ is a free $\mathbf{k}$-module, while $\mathbf{k}^S$ is not always (for example, $\mathbb{Z}^{\mathbb{N}}$ is not a free $\mathbb{Z}$-module).
- We always have $\mathbf{k}^S = \prod_{s\in S} \mathbf{k}$ and $\mathbf{k}^{(S)} = \bigoplus_{s\in S} \mathbf{k}$ as $\mathbf{k}$-modules.
- If $\mathbf{k} = \mathbb{R}$, then an element $(\alpha_s)_{s\in S} \in \mathbb{R}^S$ with all its entries $\alpha_s$ nonnegative and with $\sum_{s\in S} \alpha_s = 1$ is called a *probability distribution* on $S$. This distribution is said to be *finitely supported* if it belongs to $\mathbb{R}^{(S)}$. Thus, one may think

of $(\alpha_s)_{s\in S}$ in the general case as generalized probability distributions (a useful mental picture if you find probabilities intuitive).

When $S$ is just a set, $\mathbf{k}^{(S)}$ is just a free module. But any further structures on $S$ yield corresponding extra structures on $\mathbf{k}^{(S)}$. In particular, when $S$ is a monoid, $\mathbf{k}^{(S)}$ inherits a multiplication and becomes a $\mathbf{k}$-algebra:

**Definition 3.1.2.** Let $M$ be a monoid, written multiplicatively (so that $M = (M, \cdot, 1)$).

The *monoid algebra of $M$ over $\mathbf{k}$* (also known as the *monoid ring of $M$ over $\mathbf{k}$*) is the $\mathbf{k}$-algebra $\mathbf{k}[M]$ defined as follows: As a $\mathbf{k}$-module, it is just the free $\mathbf{k}$-module $\mathbf{k}^{(M)}$. Its multiplication is defined to be the unique $\mathbf{k}$-bilinear map $\mu : \mathbf{k}^{(M)} \times \mathbf{k}^{(M)} \to \mathbf{k}^{(M)}$ that satisfies

$$\mu\left(e_m, e_n\right) = e_{mn} \qquad \text{for all } m, n \in M. \tag{11}$$

(Here, $(e_m)_{m\in M}$ is the standard basis of $\mathbf{k}^{(M)}$, as defined above. Moreover, a map is said to be $\mathbf{k}$-*bilinear* if it is $\mathbf{k}$-linear in each of its two arguments.)

The unity of this $\mathbf{k}$-algebra $\mathbf{k}[M]$ is the basis element $e_1$ corresponding to the neutral element 1 of $M$.

**Theorem 3.1.3.** This is indeed a well-defined $\mathbf{k}$-algebra.

*Proof.* See [23wa, Theorem 4.1.2]. □

Since the bilinear map $\mu$ in Definition 3.1.2 is used as the multiplication of $\mathbf{k}[M]$, we can rewrite (11) as follows:

$$e_m \cdot e_n = e_{mn} \qquad \text{for all } m, n \in M.$$

Since the multiplication $\mu$ is $\mathbf{k}$-bilinear, we can multiply not just two basis elements, but any two of their $\mathbf{k}$-linear combinations. For example, for any $\alpha, \beta, \gamma, \delta \in \mathbf{k}$ and $m, n, p, q \in M$, we have

$$\begin{aligned}
&\left(\alpha e_m + \beta e_n\right) \cdot \left(\gamma e_p + \delta e_q\right) \\
&= \alpha e_m \left(\gamma e_p + \delta e_q\right) + \beta e_n \left(\gamma e_p + \delta e_q\right) && \left(\begin{array}{c}\text{since } \mu \text{ is linear in} \\ \text{the first argument}\end{array}\right) \\
&= \alpha\gamma \underbrace{e_m e_p}_{=e_{mp}} + \alpha\delta \underbrace{e_m e_q}_{=e_{mq}} + \beta\gamma \underbrace{e_n e_p}_{=e_{np}} + \beta\delta \underbrace{e_n e_q}_{=e_{nq}} && \left(\begin{array}{c}\text{since } \mu \text{ is linear in} \\ \text{the second argument}\end{array}\right) \\
&= \alpha\gamma e_{mp} + \alpha\delta e_{mq} + \beta\gamma e_{np} + \beta\delta e_{nq}.
\end{aligned}$$

### 3.1.3. Group algebras

**Definition 3.1.4.** When a monoid $M$ is a group, its monoid algebra $\mathbf{k}[M]$ is called its *group algebra* (or *group ring*).

### 3.1.4. Some properties

Some conventions and some simple facts make it easier to work in monoid algebras. First, we note that the commutativity of a monoid is inherited by its monoid algebra:

**Proposition 3.1.5.** Let $M$ be an **abelian** monoid. Then, the algebra $\mathbf{k}[M]$ is commutative.

*Proof.* See [23wa, Proposition 4.1.9]. □

The next two propositions show that the monoid algebra $\mathbf{k}[M]$ contains a "copy" of $\mathbf{k}$ and (unless $\mathbf{k}$ is trivial) a "copy" of $M$.

**Proposition 3.1.6.** Let $M$ be a monoid with neutral element 1. Then, the map

$$\begin{aligned} \mathbf{k} &\to \mathbf{k}[M], \\ \lambda &\mapsto \lambda e_1 \end{aligned}$$

is an injective $\mathbf{k}$-algebra morphism. Thus, its image is a $\mathbf{k}$-subalgebra of $\mathbf{k}[M]$ that is isomorphic to $\mathbf{k}$.

*Proof.* See [23wa, Proposition 4.1.10]. □

**Proposition 3.1.7.** Let $M$ be a monoid. Then, the map

$$\begin{aligned} M &\to \mathbf{k}[M], \\ m &\mapsto e_m \end{aligned}$$

is a monoid morphism from $M$ to the monoid $(\mathbf{k}[M], \cdot, 1)$. If the ring $\mathbf{k}$ is nontrivial (i.e., if $0 \neq 1$ in $\mathbf{k}$), then this morphism is injective, so that its image is a submonoid of $(\mathbf{k}[M], \cdot, 1)$ isomorphic to $M$.

*Proof.* See [23wa, Proposition 4.1.12] for the first claim. The second claim (about injectivity) is clear, since the $e_m$ for different $m \in M$ are distinct elements of the standard basis $(e_m)_{m \in M}$ and therefore distinct (when $\mathbf{k}$ is nontrivial[3]). □

The above two propositions show that $\mathbf{k}[M]$ contains both a copy of $\mathbf{k}$ and (when $\mathbf{k}$ is nontrivial) a copy of $M$. Moreover, $\mathbf{k}[M]$ is "generated" by these two pieces: Every element of $\mathbf{k}[M]$ has the form

$$\lambda_1 e_{m_1} + \lambda_2 e_{m_2} + \cdots + \lambda_k e_{m_k}$$

[3] Indeed, if $m$ and $n$ are two distinct elements of $M$, then the $m$-th entry of the family $e_m$ is 1, whereas the $m$-th entry of the family $e_n$ is 0. Thus, these two entries are distinct, unless $1 = 0$ in $\mathbf{k}$.

for some scalars $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbf{k}$ and some monoid elements $m_1, m_2, \ldots, m_k \in M$. This is just because the standard basis $(e_m)_{m \in M}$ is a basis of $\mathbf{k}[M]$.

### 3.1.5. Some conventions

It gets tiresome to always write $e_m$ for a standard basis vector in $\mathbf{k}[M]$. Wouldn't it be easier to just write $m$ instead? (After all, Proposition 3.1.7 tells us that the map sending each $m$ to $e_m$ is a monoid morphism, so that $e_m$ is a "copy" of $m$ that behaves just like $m$ with respect to multiplication.) This is indeed a good idea, and is unproblematic as long as it does not get ambiguous, i.e., as long as expressions like "$m + n$" (for $m, n \in M$) and "$\lambda m$" (for $\lambda \in \mathbf{k}$ and $m \in M$) don't already have a different meaning. So let us agree to do this:

**Convention 3.1.8.** Let $M$ be a monoid. The standard basis vectors $e_m$ in $\mathbf{k}[M]$ will just be denoted by $m$ (by abuse of notation). Thus, for example, a linear combination of the form $\alpha e_m + \beta e_n + \gamma e_p$ will be written $\alpha m + \beta n + \gamma p$. With this notation, an element of $\mathbf{k}[M]$ really looks like a $\mathbf{k}$-linear combination of elements of $M$ (although in truth, it is not the elements of $M$ but the respective standard basis vectors that are being combined).

Avoid this notation when the elements of $M$ can be scaled or added in another way (i.e., when there already is a scaling or an addition defined on the monoid $M$). For example, if $M$ is the monoid $(\mathbb{N}, \cdot, 1)$, then this notation must not be used, since there is already an addition defined on $\mathbb{N}$ which has nothing to do with the addition in $\mathbf{k}[M]$ (for example, $2 + 3 = 5$ but $e_2 + e_3 \neq e_5$). For another example in which this notation would be confusing, consider the quaternion group: a group of size 8 with elements $1, i, j, k, -1, -i, -j, -k$. If we wrote $m$ for $e_m$ here, then the nonzero sum $e_1 + e_{-1}$ would look like the zero sum $1 + (-1)$, which is not a good thing.

If $M$ is a monoid, then every element $(\alpha_m)_{m \in M}$ of the monoid algebra $\mathbf{k}[M]$ can be rewritten as $\sum_{m \in M} \alpha_m e_m$ (just like the vector $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$ in linear algebra is rewritten as $\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n$). If we use Convention 3.1.8, we can rewrite this element further as $\sum_{m \in M} \alpha_m m$ (since we write $m$ for $e_m$). Thus, using Convention 3.1.8, we have

$$(\alpha_m)_{m \in M} = \sum_{m \in M} \alpha_m e_m = \sum_{m \in M} \alpha_m m \tag{12}$$

for any family $(\alpha_m)_{m \in M} \in \mathbf{k}[M]$.

We will often use the notation $\sum_{m \in M} \alpha_m m$ instead of the "native" notation $(\alpha_m)_{m \in M}$, because it allows us to multiply two elements of $\mathbf{k}[M]$ by just expanding the product:

$$\left( \sum_{m \in M} \alpha_m m \right) \left( \sum_{n \in M} \beta_n n \right) = \sum_{m \in M} \sum_{n \in M} \underbrace{\alpha_m \beta_n}_{\text{a product in } \mathbf{k}} \underbrace{mn}_{\text{a product in } M}.$$

**Remark 3.1.9.** Let $M$ be a monoid. Let $(\alpha_m)_{m\in M}$ and $(\beta_m)_{m\in M}$ be two elements of the monoid algebra $\mathbf{k}[M]$. Then, the equation

$$\sum_{m\in M} \alpha_m m = \sum_{m\in M} \beta_m m \qquad \text{in } \mathbf{k}[M] \tag{13}$$

is equivalent to the statement that

$$(\alpha_m = \beta_m \qquad \text{for each } m \in M). \tag{14}$$

This fact is completely trivial (since we can use (12) to rewrite the equation (13) as $(\alpha_m)_{m\in M} = (\beta_m)_{m\in M}$, but this is clearly equivalent to the statement (14)). Nevertheless, it is worth pointing out, since it provides a dictionary between equalities in the monoid algebra and equalities in the base ring $\mathbf{k}$. We will often use it to deduce (14) from (13) (since the other direction is obvious). This deduction is called *comparing coefficients*.

Here is another useful convention:

**Convention 3.1.10.** Let $M$ be a monoid. Then, we shall identify each scalar $\lambda \in \mathbf{k}$ with the corresponding element $\lambda e_1 \in \mathbf{k}[M]$ of the monoid algebra. In other words, we identify each $\lambda \in \mathbf{k}$ with its image under the injective $\mathbf{k}$-algebra morphism

$$\begin{aligned} &\mathbf{k} \to \mathbf{k}[M], \\ &\lambda \mapsto \lambda e_1. \end{aligned}$$

Thus, for example, $\alpha e_1 + \beta e_m$ becomes $\alpha 1 + \beta m$ (by Convention 3.1.8) or just $\alpha + \beta m$ (by Convention 3.1.10).

Note that Convention 3.1.10 is an instance of the standard convention to embed the base ring $\mathbf{k}$ into each $\mathbf{k}$-algebra $A$, at least if the canonical morphism

$$\begin{aligned} &\mathbf{k} \to A, \\ &\lambda \mapsto \lambda 1_A \end{aligned}$$

is injective.

### 3.1.6. Examples of monoid algebras

Here are some of the most famous examples of monoid algebras:

1. Let $M$ be the monoid $\{x^0, x^1, x^2, \ldots\}$ with multiplication given by $x^i x^j = x^{i+j}$ and with neutral element $x^0$. This monoid is just the familiar monoid $(\mathbb{N}, +, 0)$, rewritten multiplicatively (by renaming each $n \in \mathbb{N}$ as $x^n$).

The monoid ring $\mathbf{k}[M]$ then has the standard basis $(e_{x^0}, e_{x^1}, e_{x^2}, \ldots)$, which (by Convention 3.1.8) we can just write as $\left(x^0, x^1, x^2, \ldots\right)$. Thus, each element of this monoid ring is a $\mathbf{k}$-linear combination of $x^0, x^1, x^2, \ldots$. Moreover, the multiplication is given by $e_{x^i} e_{x^j} = e_{x^i x^j} = e_{x^{i+j}}$, that is, by $x^i x^j = x^{i+j}$ (again using Convention 3.1.8). This should sound familiar: It shows that this monoid ring $\mathbf{k}[M]$ is just the univariate polynomial ring $\mathbf{k}[x]$.

Note that this is not the same as the ring $\mathbf{k}[[x]]$ of formal power series, because we only allow essentially finite $\mathbf{k}$-linear combinations (i.e., all but finitely many coefficients must be 0).

2. Let $M$ be the monoid $\{\ldots, x^{-1}, x^0, x^1, x^2, \ldots\}$ with multiplication given by $x^i x^j = x^{i+j}$ and with neutral element $x^0$. This monoid is just $(\mathbb{Z}, +, 0)$, rewritten multiplicatively (by renaming each $n \in \mathbb{Z}$ as $x^n$).

   Then, the monoid ring $\mathbf{k}[M]$ is the Laurent polynomial ring $\mathbf{k}\left[x^{\pm 1}\right]$.

3. You can similarly obtain multivariate polynomial rings (or Laurent polynomial rings) by starting with multivariate monomial monoids (i.e., monoids of the form $\left(\mathbb{N}^k, +, 0\right)$ or $\left(\mathbb{Z}^k, +, 0\right)$, rewritten multiplicatively).

## 3.2. The symmetric group algebra

**Convention 3.2.1.** From now on, we fix a nonnegative integer $n$.

**Definition 3.2.2.** The *$n$-th symmetric group algebra* (over $\mathbf{k}$) means the group algebra of the $n$-th symmetric group $S_n$ over $\mathbf{k}$. In other words, it means the $\mathbf{k}$-algebra $\mathbf{k}[S_n]$. We will be using Convention 3.1.8 when working in this algebra.

So the elements of $\mathbf{k}[S_n]$ are $\mathbf{k}$-linear combinations of permutations of $[n]$. Note that the monoid $S_n$ is finite, so that we don't need to distinguish between $\mathbf{k}^{S_n}$ and $\mathbf{k}^{(S_n)}$.

Let us play around with $\mathbf{k}[S_n]$ for $n = 2$ and $n = 3$:

**Example 3.2.3.** The symmetric group $S_2$ has just two elements: the permutation id and the simple transposition $s_1$ (recall that $s_1$ is the transposition $t_{1,2}$, swapping with 1 with 2). Thus, the standard basis of $\mathbf{k}[S_2]$ is $(e_{\mathrm{id}}, e_{s_1}) = (\mathrm{id}, s_1)$ (using Convention 3.1.8). Hence, each element of $\mathbf{k}[S_2]$ has the form

$$\alpha\,\mathrm{id} + \beta s_1 \qquad \text{for some } \alpha, \beta \in \mathbf{k}$$

(of course, $\alpha\,\mathrm{id} + \beta s_1$ really means $\alpha e_{\mathrm{id}} + \beta e_{s_1}$, since we are using Convention 3.1.8). Since id is the neutral element of $S_2$, we can also write 1 for it, and we

can rewrite $\alpha \operatorname{id}$ as $\alpha$ (by Convention 3.1.10). Hence, each element of $\mathbf{k}[S_2]$ has the form

$$\alpha + \beta s_1 \qquad \text{for some } \alpha, \beta \in \mathbf{k}.$$

Let us multiply two such elements:

$$\begin{aligned}
(\alpha + \beta s_1)(\gamma + \delta s_1) &= \alpha\gamma + \alpha\delta s_1 + \beta\gamma s_1 + \beta\delta \underbrace{s_1 s_1}_{\substack{=\mathrm{id} \\ (\text{because } t_{i,j} t_{i,j} = \mathrm{id} \\ \text{for any } i,j)}} \\
&= \alpha\gamma + \alpha\delta s_1 + \beta\gamma s_1 + \underbrace{\beta\delta \operatorname{id}}_{=\beta\delta} \\
&= \alpha\gamma + \alpha\delta s_1 + \beta\gamma s_1 + \beta\delta \\
&= (\alpha\gamma + \beta\delta) + (\alpha\delta + \beta\gamma) s_1.
\end{aligned}$$

The unity of the algebra $\mathbf{k}[S_2]$ is the permutation id, or more precisely the corresponding basis vector $e_{\mathrm{id}}$ (which we just call id).

The algebra $\mathbf{k}[S_2]$ is commutative, since the group $S_2$ is abelian. (Indeed, $S_2$ is a cyclic group of order 2.)

**Example 3.2.4.** The symmetric group $S_3$ has 6 elements: the permutations

$$\operatorname{oln}(123),\ \operatorname{oln}(132),\ \operatorname{oln}(213),\ \operatorname{oln}(231),\ \operatorname{oln}(312),\ \operatorname{oln}(321)$$

(see Subsection 2.4.2 for the meaning of this notation). Hence, each element of the symmetric group algebra $\mathbf{k}[S_3]$ has the form

$$\alpha \underbrace{\operatorname{oln}(123)}_{=\mathrm{id}=1} + \beta \underbrace{\operatorname{oln}(132)}_{=s_2} + \gamma \underbrace{\operatorname{oln}(213)}_{=s_1} + \delta \underbrace{\operatorname{oln}(231)}_{\substack{=\mathrm{cyc}_{1,2,3} \\ =s_1 s_2}} + \varepsilon \underbrace{\operatorname{oln}(312)}_{\substack{=\mathrm{cyc}_{1,3,2} \\ =s_2 s_1}} + \zeta \underbrace{\operatorname{oln}(321)}_{\substack{=t_{1,3} \\ =s_1 s_2 s_1 \\ =s_2 s_1 s_2}}$$

for some $\alpha, \beta, \gamma, \delta, \varepsilon, \zeta \in \mathbf{k}$. Just like in Example 3.2.3, we could compute the product of any two such elements. Instead of the general formula, let me show two examples of such products:

$$\begin{aligned}
(1 + s_1)(1 - s_1) &= 1 + s_1 - s_1 - s_1 s_1 = 1 - \underbrace{s_1 s_1}_{=\mathrm{id}=1} = 1 - 1 \\
&= 0
\end{aligned} \tag{15}$$

and

$$\begin{aligned}
(1 + s_2)(1 + s_1 + s_1 s_2) &= 1 + s_1 + s_1 s_2 + s_2 + s_2 s_1 + s_2 s_1 s_2 \\
&= \sum_{w \in S_3} w.
\end{aligned} \tag{16}$$

The product (15) can be easily generalized: We have

$$\left(1 + t_{i,j}\right)\left(1 - t_{i,j}\right) = 0 \qquad \text{for any transposition } t_{i,j} \in S_n.$$

Even more generally,

$$(1 + w)(1 - w) = 0 \qquad \text{for any involution } w \in S_n.$$

(An *involution* means a permutation $w$ such that $w^2 = \mathrm{id}$, that is, $w = w^{-1}$.)

Soon we will also generalize the product (16).

We note that the 0-th and the 1-st symmetric group algebras $\mathbf{k}\left[S_0\right]$ and $\mathbf{k}\left[S_1\right]$ are both isomorphic to $\mathbf{k}$, since the groups $S_0$ and $S_1$ are trivial (i.e., groups of size 1). These are not very interesting objects, although their existence can be helpful in induction proofs.

## 3.3. The integral and the sign-integral

We now define two peculiar elements of any symmetric group algebra.

**Definition 3.3.1.** The *integral* (aka *symmetrizer*) of the symmetric group algebra $\mathbf{k}\left[S_n\right]$ is defined to be the element

$$\nabla := \sum_{w \in S_n} w \in \mathbf{k}\left[S_n\right].$$

The *sign-integral* (aka *antisymmetrizer*) of the symmetric group algebra $\mathbf{k}\left[S_n\right]$ is defined to be the element

$$\nabla^- := \sum_{w \in S_n} \underbrace{(-1)^w}_{\text{the sign of } w} w \in \mathbf{k}\left[S_n\right].$$

**Example 3.3.2.** For $n = 2$, we have

$$\nabla = 1 + s_1 \qquad \text{and} \qquad \nabla^- = 1 - s_1.$$

For $n = 3$, we have

$$\begin{aligned} \nabla &= 1 + s_1 + s_2 + s_1 s_2 + s_2 s_1 + s_2 s_1 s_2 \qquad \text{and} \\ \nabla^- &= 1 - s_1 - s_2 + s_1 s_2 + s_2 s_1 - s_2 s_1 s_2. \end{aligned}$$

For $n = 0$ or $n = 1$, we have $\nabla = \nabla^- = 1$.

Multiplying the integral or the sign-integral by another permutation has a rather simple outcome:

**Proposition 3.3.3.** Let $w \in S_n$. Then,

$$w\nabla = \nabla w = \nabla \qquad \text{and} \tag{17}$$
$$w\nabla^- = \nabla^- w = (-1)^w \nabla^-. \tag{18}$$

*Proof.* The symmetric group $S_n$ is a group. Thus, the map

$$\begin{aligned} S_n &\to S_n, \\ u &\mapsto wu \end{aligned}$$

is a bijection (with inverse given by $u \mapsto w^{-1}u$). But the definition of $\nabla$ can be rewritten as $\nabla = \sum_{u \in S_n} u$. Thus,

$$\begin{aligned} w\nabla &= w \sum_{u \in S_n} u = \sum_{u \in S_n} wu = \sum_{u \in S_n} u \\ &\qquad \left( \begin{array}{c} \text{here, we substituted } u \text{ for } wu \text{ in the sum, since} \\ \text{the map } S_n \to S_n,\ u \mapsto wu \text{ is a bijection} \end{array} \right) \\ &= \nabla. \end{aligned}$$

Similarly, $\nabla w = \nabla$. Combining these two equalities, we obtain (17).

Next, let us prove (18). One of the basic properties of signs of permutations says that $(-1)^{\sigma\tau} = (-1)^\sigma \cdot (-1)^\tau$ for any $\sigma, \tau \in S_n$. Thus, for any $u \in S_n$, we have

$$(-1)^{wu} = (-1)^w \cdot (-1)^u,$$

so that

$$\begin{aligned} (-1)^u &= \underbrace{\frac{1}{(-1)^w}}_{=(-1)^w} \cdot (-1)^{wu} \\ &\qquad \begin{array}{c} \text{(since the number } (-1)^w \text{ is either 1 or } -1, \\ \text{and thus equals its own reciprocal)} \end{array} \\ &= (-1)^w \cdot (-1)^{wu}. \end{aligned} \tag{19}$$

Now, rewrite the definition of $\nabla^-$ as $\nabla^- = \sum\limits_{u \in S_n} (-1)^u u$. Thus,

$$\begin{aligned} w\nabla^- &= w \sum_{u \in S_n} (-1)^u u = \sum_{u \in S_n} \underbrace{(-1)^u}_{\substack{=(-1)^w \cdot (-1)^{wu} \\ \text{(by (19))}}} wu \\ &= \sum_{u \in S_n} (-1)^w \cdot (-1)^{wu} wu = \sum_{u \in S_n} (-1)^w \cdot (-1)^u u \\ &\qquad \left( \begin{array}{c} \text{here, we substituted } u \text{ for } wu \text{ in the sum, since} \\ \text{the map } S_n \to S_n,\ u \mapsto wu \text{ is a bijection} \end{array} \right) \\ &= (-1)^w \cdot \underbrace{\sum_{u \in S_n} (-1)^u u}_{=\nabla^-} = (-1)^w \nabla^-. \end{aligned}$$

Similarly, $\nabla^- w = (-1)^w \nabla^-$. Combining these two equalities, we obtain (18). This completes the proof of Proposition 3.3.3. □

Note that commutativity in $\mathbf{k}[S_n]$ is rare (for $n \geq 3$), so that the equalities $w\nabla = \nabla w$ and $w\nabla^- = \nabla^- w$ in Proposition 3.3.3 should be somewhat surprising (but not too much, since $\nabla$ and $\nabla^-$ are two of the simplest elements of $\mathbf{k}[S_n]$).

The relation (17), in a sense, characterizes $\nabla$ (up to scalar multiplication): Indeed, the only elements $\mathbf{a} \in \mathbf{k}[S_n]$ that satisfy $\mathbf{a}w = \mathbf{a}$ for all $w \in S_n$ are $\nabla$ and its scalar multiples. This is part of the following exercise:

**Exercise 3.3.1.** Let $\mathbf{a} \in \mathbf{k}[S_n]$. Prove the following:

**(a)** $\boxed{1}$ If $\mathbf{a}w = \mathbf{a}$ for all $w \in S_n$, then $\mathbf{a} = \lambda\nabla$ for some $\lambda \in \mathbf{k}$.

**(b)** $\boxed{1}$ Even better: If $\mathbf{a}s_i = \mathbf{a}$ for all $i \in [n-1]$, then $\mathbf{a} = \lambda\nabla$ for some $\lambda \in \mathbf{k}$.

**Corollary 3.3.4.** We have $\nabla^2 = n! \cdot \nabla$.

*Proof.* We have

$$\begin{aligned} \nabla^2 = \nabla\nabla &= \left( \sum_{w \in S_n} w \right) \nabla \qquad \left( \text{since } \nabla = \sum_{w \in S_n} w \right) \\ &= \sum_{w \in S_n} \underbrace{w\nabla}_{\substack{=\nabla \\ \text{(by (17))}}} = \sum_{w \in S_n} \nabla = n! \cdot \nabla \end{aligned}$$

(since there are $n!$ many permutations $w \in S_n$). This proves Corollary 3.3.4. □

Corollary 3.3.4 can be restated as follows: $\nabla(\nabla - n!) = 0$. In other words, $P(\nabla) = 0$, where $P \in \mathbb{Z}[x]$ is the polynomial $x(x - n!)$. This is the first sound

of a recurring theme: Often, you define an element $\mathbf{a}$ of $\mathbf{k}[S_n]$ and find that there is a "nice" polynomial $P \in \mathbf{k}[x]$ such that $P(\mathbf{a}) = 0$. Actually, we will soon see that there always is a monic polynomial $P \in \mathbf{k}[x]$ of degree $n!$ such that $P(\mathbf{a}) = 0$. But in many cases, there exist such polynomials $P$ of much smaller degree than $n!$, and sometimes these $P$'s factor into linear factors over $\mathbb{Z}$. Many interesting results have this form; Corollary 3.3.4 is the simplest of them.

**Remark 3.3.5.** We can generalize $\nabla$ to an arbitrary finite group $G$. Indeed, if $G$ is any finite group, then we can define an element $\nabla_G := \sum_{g \in G} g \in \mathbf{k}[G]$, and then every $w \in G$ satisfies

$$w\nabla_G = \nabla_G w = \nabla_G. \tag{20}$$

This can be proved in the same way as we proved (17) (and indeed, (17) is the particular case of (20) for $G = S_n$). As a consequence, $\nabla_G^2 = |G| \cdot \nabla_G$ (which generalizes Corollary 3.3.4).

For infinite groups $G$, we cannot define $\nabla_G$, since linear combinations are not allowed to have infinitely many nonzero coefficients.

The equality (18) can also be generalized with a bit more work: We let $\mathbf{k}^\times$ denote the group of units of $\mathbf{k}$ (that is, the group of all invertible elements of $\mathbf{k}$, with respect to multiplication). If $f : G \to \mathbf{k}^\times$ is a group morphism from $G$ to $\mathbf{k}^\times$, then we can define an element $\nabla_G^f := \sum_{g \in G} f(g)\, g \in \mathbf{k}[G]$, and then every $w \in G$ satisfies

$$w\nabla_G^f = \nabla_G^f w = f\left(w^{-1}\right) \cdot \nabla_G^f. \tag{21}$$

If we let $G = S_n$ and let $f$ be the sign homomorphism (so that $f(w) = \operatorname{sign} w = (-1)^w$ for each $w \in S_n$), then this recovers (18).

## 3.4. The Young–Jucys–Murphy elements (aka YJM elements, or jucys–murphies)

### 3.4.1. Definition and commutativity

It is time to define the first family of elements of $\mathbf{k}[S_n]$ that have nontrivial interesting properties:

**Definition 3.4.1.** For each $k \in [n]$, we define the *$k$-th Young–Jucys–Murphy element* (for short: *YJM element*, or just *jucys–murphy*) in $\mathbf{k}[S_n]$ to be the element

$$\mathbf{m}_k := t_{1,k} + t_{2,k} + \cdots + t_{k-1,k} = \sum_{i=1}^{k-1} t_{i,k}.$$

This is the sum of all transpositions $t_{i,k}$ with $i < k$.

In particular,

$$\begin{aligned}
\mathbf{m}_1 &= 0; \\
\mathbf{m}_2 &= t_{1,2}; \\
\mathbf{m}_3 &= t_{1,3} + t_{2,3}; \\
\mathbf{m}_4 &= t_{1,4} + t_{2,4} + t_{3,4}; \\
&\cdots; \\
\mathbf{m}_n &= t_{1,n} + t_{2,n} + \cdots + t_{n-1,n}.
\end{aligned}$$

Thus,

$$\mathbf{m}_1 + \mathbf{m}_2 + \cdots + \mathbf{m}_n = \sum_{k=1}^{n} \mathbf{m}_k = \sum_{k=1}^{n} \sum_{i=1}^{k-1} t_{i,k}$$

is the sum of **all** transpositions in $S_n$.

Here is our first surprise:

**Theorem 3.4.2.** The jucys–murphies $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n$ commute. In other words,

$$\mathbf{m}_i \mathbf{m}_j = \mathbf{m}_j \mathbf{m}_i \qquad \text{for all } i, j \in [n].$$

*Proof.* Let $i, j \in [n]$. We must prove that $\mathbf{m}_i \mathbf{m}_j = \mathbf{m}_j \mathbf{m}_i$. If $i = j$, then this is obvious. Thus, we WLOG assume that $i \neq j$. Furthermore, our claim ($\mathbf{m}_i \mathbf{m}_j = \mathbf{m}_j \mathbf{m}_i$) is symmetric in $i$ and $j$, so that we can swap $i$ with $j$ at will. Hence, we furthermore WLOG assume that $i \leq j$. Thus, $i < j$ (since $i \neq j$).

Recall that

$$\mathbf{m}_i = t_{1,i} + t_{2,i} + \cdots + t_{i-1,i} = \sum_{p=1}^{i-1} t_{p,i} \qquad \text{and}$$

$$\mathbf{m}_j = t_{1,j} + t_{2,j} + \cdots + t_{j-1,j} = \sum_{k=1}^{j-1} t_{k,j}.$$

Now we shall show that

$$t_{p,i} \mathbf{m}_j = \mathbf{m}_j t_{p,i} \text{ for each } p \in [i-1]. \tag{22}$$

[*Proof of (22):* Let $p \in [i-1]$. Then, $p \leq i-1 < i < j$. Hence, each $k \in [j-1] \setminus \{p, i\}$ satisfies $k \notin \{j, p, i\}$ and therefore

$$t_{p,i} t_{k,j} = t_{k,j} t_{p,i} \tag{23}$$

(this is a particular case of (8), since the four numbers $p, i, k, j$ are distinct). Moreover, we can easily check that

$$t_{p,i} t_{p,j} = t_{i,j} t_{p,i} \tag{24}$$

(indeed, both sides here equal $\operatorname{cyc}_{i,p,j}$) and

$$t_{p,i}t_{i,j} = t_{p,j}t_{p,i} \tag{25}$$

(indeed, both sides here equal $\operatorname{cyc}_{i,j,p}$).

Now, $p$ and $i$ are two distinct elements of $[j]$ (since $p < i < j$). Recall that

$$\mathbf{m}_j = \sum_{k=1}^{j-1} t_{k,j} = t_{p,j} + t_{i,j} + \sum_{k \in [j-1] \setminus \{p,i\}} t_{k,j}$$

(here, we have split off the addends for $k = p$ and for $k = i$ from the sum, since $p$ and $i$ are two distinct elements of $[j]$). Hence,

$$\begin{aligned}
t_{p,i}\mathbf{m}_j &= t_{p,i}\left(t_{p,j} + t_{i,j} + \sum_{k \in [j-1] \setminus \{p,i\}} t_{k,j}\right) \\
&= \underbrace{t_{p,i}t_{p,j}}_{\substack{=t_{i,j}t_{p,i} \\ \text{(by (24))}}} + \underbrace{t_{p,i}t_{i,j}}_{\substack{=t_{p,j}t_{p,i} \\ \text{(by (25))}}} + \sum_{k \in [j-1] \setminus \{p,i\}} \underbrace{t_{p,i}t_{k,j}}_{\substack{=t_{k,j}t_{p,i} \\ \text{(by (23))}}} \\
&= t_{i,j}t_{p,i} + t_{p,j}t_{p,i} + \sum_{k \in [j-1] \setminus \{p,i\}} t_{k,j}t_{p,i} \\
&= \underbrace{\left(t_{i,j} + t_{p,j} + \sum_{k \in [j-1] \setminus \{p,i\}} t_{k,j}\right)}_{\substack{=t_{p,j}+t_{i,j}+\sum\limits_{k \in [j-1] \setminus \{p,i\}} t_{k,j} \\ =\mathbf{m}_j}} t_{p,i} = \mathbf{m}_j t_{p,i}.
\end{aligned}$$

This proves (22).]

Now, from $\mathbf{m}_i = \sum_{p=1}^{i-1} t_{p,i}$, we obtain

$$\mathbf{m}_i\mathbf{m}_j = \left(\sum_{p=1}^{i-1} t_{p,i}\right)\mathbf{m}_j = \sum_{p=1}^{i-1} \underbrace{t_{p,i}\mathbf{m}_j}_{\substack{=\mathbf{m}_j t_{p,i} \\ \text{(by (22))}}} = \sum_{p=1}^{i-1} \mathbf{m}_j t_{p,i} = \mathbf{m}_j \underbrace{\left(\sum_{p=1}^{i-1} t_{p,i}\right)}_{=\mathbf{m}_i} = \mathbf{m}_j\mathbf{m}_i.$$

This completes the proof of Theorem 3.4.2. □

### 3.4.2. A product

Now let us compute the product $(1 + \mathbf{m}_1)(1 + \mathbf{m}_2)\cdots(1 + \mathbf{m}_n)$. For $n = 3$, it is

$$\begin{aligned}
(1 + \mathbf{m}_1)(1 + \mathbf{m}_2)(1 + \mathbf{m}_3) &= 1\,(1 + t_{1,2})\,(1 + t_{1,3} + t_{2,3}) \\
&= 1 + t_{1,2} + t_{1,3} + t_{2,3} + t_{1,2}t_{1,3} + t_{1,2}t_{2,3} \\
&= \sum_{w \in S_3} w = \nabla.
\end{aligned}$$

This is not a coincidence:

**Theorem 3.4.3.** We have

$$(1+\mathbf{m}_1)(1+\mathbf{m}_2)\cdots(1+\mathbf{m}_n) = \nabla.$$

*First proof of Theorem 3.4.3.* We have so far defined the transpositions $t_{i,j}$ only for $i \neq j$. Now, let us also set $t_{k,k} := \mathrm{id}$ for each $k \in [n]$. This is not really a transposition, but it still fits the pattern of the $t_{i,j}$'s, since it "swaps $k$ with $k$ and leaves everything else unchanged". In other words, we still have $t_{i,j}(i) = j$ and $t_{i,j}(j) = i$ for all $i, j \in [n]$, as well as $t_{i,j}(k) = k$ for all $i, j, k \in [n]$ that satisfy $k \notin \{i, j\}$.

Now, for each $k \in [n]$, we have

$$\underbrace{1}_{=\mathrm{id}=t_{k,k}} + \underbrace{\mathbf{m}_k}_{=\sum_{j=1}^{k-1} t_{j,k}} = t_{k,k} + \sum_{j=1}^{k-1} t_{j,k} = \sum_{j=1}^{k} t_{j,k}.$$

Multiplying these equalities for all $k \in [n]$, we obtain

$$\begin{aligned}
&(1+\mathbf{m}_1)(1+\mathbf{m}_2)\cdots(1+\mathbf{m}_n) \\
&= \left(\sum_{j=1}^{1} t_{j,1}\right)\left(\sum_{j=1}^{2} t_{j,2}\right)\cdots\left(\sum_{j=1}^{n} t_{j,n}\right) \\
&= \sum_{(j_1,j_2,\ldots,j_n)\in[1]\times[2]\times\cdots\times[n]} t_{j_1,1}t_{j_2,2}\cdots t_{j_n,n}
\end{aligned}$$

(by the product rule, i.e., by expanding the product).[4]

Our goal is to prove that this sum equals $\nabla$, i.e., that it contains each permutation $w \in S_n$ exactly once. In other words, we must prove the following theorem:

**Theorem 3.4.4.** Each $w \in S_n$ can be written as $t_{j_1,1}t_{j_2,2}\cdots t_{j_n,n}$ for exactly one $n$-tuple $(j_1, j_2, \ldots, j_n) \in [1] \times [2] \times \cdots \times [n]$. In other words, the map

$$\begin{aligned}
[1]\times[2]\times\cdots\times[n] &\to S_n, \\
(j_1, j_2, \ldots, j_n) &\mapsto t_{j_1,1}t_{j_2,2}\cdots t_{j_n,n}
\end{aligned}$$

is a bijection.

[4] The condition $(j_1, j_2, \ldots, j_n) \in [1] \times [2] \times \cdots \times [n]$ under the summation sign says precisely that $j_1, j_2, \ldots, j_n$ are positive integers satisfying $j_k \leq k$ for each $k$.

*Proof of Theorem 3.4.4.* Let $\Phi$ be the map

$$[1] \times [2] \times \cdots \times [n] \to S_n,$$
$$(j_1, j_2, \ldots, j_n) \mapsto t_{j_1,1} t_{j_2,2} \cdots t_{j_n,n}.$$

We must show that this map $\Phi$ is a bijection.

Let us first prove that it is injective. In other words, let us prove that any $n$-tuple $(j_1, j_2, \ldots, j_n) \in [1] \times [2] \times \cdots \times [n]$ can be reconstructed from its image $\Phi(j_1, j_2, \ldots, j_n) = t_{j_1,1} t_{j_2,2} \cdots t_{j_n,n}$.

To do so, we pick an $n$-tuple $(j_1, j_2, \ldots, j_n) \in [1] \times [2] \times \cdots \times [n]$, and we set

$$w := \Phi(j_1, j_2, \ldots, j_n) = t_{j_1,1} t_{j_2,2} \cdots t_{j_n,n}.$$

How can we reconstruct $(j_1, j_2, \ldots, j_n)$ from this $w$ ? Let us define the permutation

$$w_k := t_{j_1,1} t_{j_2,2} \cdots t_{j_k,k} \in S_n \qquad \text{for each } k \in \{0, 1, \ldots, n\}.$$

Thus, $w_0 = (\text{empty product}) = 1 = \mathrm{id}$ and $w_n = t_{j_1,1} t_{j_2,2} \cdots t_{j_n,n} = w$. The definition of $w_k$ shows that

$$w_k = w_{k-1} t_{j_k,k} \qquad \text{for each } k \in [n] \tag{26}$$

(since $\underbrace{w_{k-1}}_{= t_{j_1,1} t_{j_2,2} \cdots t_{j_{k-1},k-1}} t_{j_k,k} = \left( t_{j_1,1} t_{j_2,2} \cdots t_{j_{k-1},k-1} \right) t_{j_k,k} = t_{j_1,1} t_{j_2,2} \cdots t_{j_k,k} = w_k$).

We can now reconstruct (in this order) $w_n, j_n, w_{n-1}, j_{n-1}, w_{n-2}, j_{n-2}, \ldots$ (from $w$) by the following recursive procedure:

1. We know $w_n$ already, since $w_n = w$.
2. I claim that $j_n = w_n^{-1}(n)$. To see this, note that $w_n = t_{j_1,1} t_{j_2,2} \cdots t_{j_n,n}$ entails

$$w_n(j_n) = \left( t_{j_1,1} t_{j_2,2} \cdots t_{j_n,n} \right)(j_n) = \left( t_{j_1,1} t_{j_2,2} \cdots t_{j_{n-1},n-1} \right) \left( \underbrace{t_{j_n,n}(j_n)}_{=n} \right)$$
$$= \left( t_{j_1,1} t_{j_2,2} \cdots t_{j_{n-1},n-1} \right)(n) = n,$$

   since each of the transpositions $t_{j_1,1}, t_{j_2,2}, \ldots, t_{j_{n-1},n-1}$ leaves $n$ unchanged (because $j_1 \in [1]$, $j_2 \in [2]$, ..., $j_{n-1} \in [n-1]$ ensures that all the numbers $j_1, j_2, \ldots, j_{n-1}$ are $< n$). Hence, $j_n = w_n^{-1}(n)$, as I claimed.

   Thus, $j_n$ has been reconstructed from $w$.

3. Next, I claim that $w_{n-1} = w_n t_{j_n,n}^{-1}$. (Note that $t_{j_n,n}^{-1}$ is the same as $t_{j_n,n}$, but we write it as $t_{j_n,n}^{-1}$ for clarity.) Indeed, (26) (applied to $k = n$) yields $w_n = w_{n-1} t_{j_n,n}$. Solving this for $w_{n-1}$, we obtain $w_{n-1} = w_n t_{j_n,n}^{-1}$.

   Thus, $w_{n-1}$ has been reconstructed (since $j_n$ and $w_n$ have been reconstructed).

4. Just as we showed that $j_n = w_n^{-1}(n)$, we can show that $j_{n-1} = w_{n-1}^{-1}(n-1)$ (since $t_{j_{n-1},n-1}(j_{n-1}) = n-1$, and since each of the transpositions $t_{j_1,1}, t_{j_2,2}, \ldots, t_{j_{n-2},n-2}$ leaves $n-1$ unchanged). Thus, $j_{n-1}$ has been reconstructed from $w$.

5. Just as we showed that $w_{n-1} = w_n t_{j_n,n}^{-1}$, we can show that $w_{n-2} = w_{n-1} t_{j_{n-1},n-1}^{-1}$. Thus, $w_{n-2}$ has been reconstructed.

6. We continue along the same lines. That is, for each $k$ ranging from $n$ to 1, we obtain $j_k$ from $w_k$ by the formula $j_k = w_k^{-1}(k)$, and then obtain $w_{k-1}$ by the formula $w_{k-1} = w_k t_{j_k,k}^{-1}$.

At the end of this procedure, all of $j_1, j_2, \ldots, j_n$ have been reconstructed, so that the $n$-tuple $(j_1, j_2, \ldots, j_n)$ has been reconstructed. This shows that our map $\Phi$ is injective.

As we recall, our goal is to prove that $\Phi$ is a bijection, i.e., bijective. A very quick shortcut is available towards this goal: The sets $[1] \times [2] \times \cdots \times [n]$ and $S_n$ are two finite sets of the same size (namely, $n!$), and thus any injective map from one to the other must be bijective (by the pigeonhole principle). Thus, $\Phi$ is bijective (since $\Phi$ is injective), and Theorem 3.4.4 is proved.

This trick argument is perfectly legitimate, but let me also show a slower but more "natural" way to finish this proof of Theorem 3.4.4. We know that $\Phi$ is injective; we must prove that $\Phi$ is bijective. So it remains to prove that $\Phi$ is surjective. In other words, we must show that each $w \in S_n$ can be written as $\Phi(j_1, j_2, \ldots, j_n)$ for some $(j_1, j_2, \ldots, j_n) \in [1] \times [2] \times \cdots \times [n]$.

So let $w \in S_n$ be arbitrary. Recall the recursive procedure that we used to reconstruct $j_1, j_2, \ldots, j_n$ in the above proof of the injectivity of $\Phi$. We can apply this procedure to $w$ without knowing a-priori that $j_1, j_2, \ldots, j_n$ exist. Thus,

1. we define $w_n \in S_n$ by $w_n := w$;

2. we define $j_n \in [n]$ by $j_n := w_n^{-1}(n)$;

3. we define $w_{n-1} \in S_n$ by $w_{n-1} := w_n t_{j_n,n}^{-1}$;

4. we define $j_{n-1} \in [n]$ by $j_{n-1} := w_{n-1}^{-1}(n-1)$;

5. we define $w_{n-2} \in S_n$ by $w_{n-2} := w_{n-1} t_{j_{n-1},n-1}^{-1}$;

6. and so on.

Thus, we have recursively defined $n+1$ permutations $w_n, w_{n-1}, \ldots, w_0 \in S_n$ and $n$ elements $j_n, j_{n-1}, \ldots, j_1 \in [n]$ by the equalities $w_n := w$ and

$$w_{k-1} := w_k t_{j_k,k}^{-1} \qquad \text{and} \tag{27}$$

$$j_k := w_k^{-1}(k) \tag{28}$$

for all $k \in [n]$. It remains to show that $(j_1, j_2, \ldots, j_n) \in [1] \times [2] \times \cdots \times [n]$ and that $\Phi(j_1, j_2, \ldots, j_n) = w$. Let us now do this. We claim that[5]

$$w_k \text{ fixes } k+1, k+2, \ldots, n \tag{29}$$

(i.e., that $w_k(i) = i$ for each $i \in \{k+1, k+2, \ldots, n\}$) for each $k \in \{0, 1, \ldots, n\}$. We shall prove this claim by descending induction on $k$ (that is, by an induction with base case $k = n$ and with induction step $k \mapsto k-1$). The *base case* ($k = n$) is vacuously true, since in this case the list of numbers $k+1, k+2, \ldots, n$ is empty. For the *induction step*, we fix some $k \in [n]$, and assume (as our induction hypothesis) that $w_k$ fixes $k+1, k+2, \ldots, n$; our goal is to show that $w_{k-1}$ fixes $k, k+1, \ldots, n$. From (28), we see that $w_k(j_k) = k$, and thus $j_k \leq k$ (because if we had $j_k > k$, then $j_k$ would be one of the numbers $k+1, k+2, \ldots, n$, which (by our induction hypothesis) would imply that $w_k$ fixes $j_k$, so that $w_k(j_k) = j_k > k$, which would contradict $w_k(j_k) = k$). Hence, both $j_k$ and $k$ are $\leq k$. Therefore, the permutation $t_{j_k,k}$ fixes $k+1, k+2, \ldots, n$. Hence, its inverse $t_{j_k,k}^{-1}$ fixes $k+1, k+2, \ldots, n$ as well. So does the permutation $w_k$ (by our induction hypothesis). But (27) yields $w_{k-1} = w_k t_{j_k,k}^{-1}$. Hence, $w_{k-1}$ fixes $k+1, k+2, \ldots, n$ (since both factors $w_k$ and $t_{j_k,k}^{-1}$ in the product $w_k t_{j_k,k}^{-1}$ fix $k+1, k+2, \ldots, n$). Moreover, $w_{k-1}$ fixes $k$ as well, since

$$\underbrace{w_{k-1}}_{=w_k t_{j_k,k}^{-1}}(k) = \left(w_k t_{j_k,k}^{-1}\right)(k) = w_k\left(\underbrace{t_{j_k,k}^{-1}(k)}_{=j_k}\right) = w_k(j_k) = k.$$

Thus, $w_{k-1}$ fixes $k, k+1, \ldots, n$ (since $w_{k-1}$ fixes $k$ and fixes $k+1, k+2, \ldots, n$). This completes the induction step. Thus, the proof of (29) is complete.

In the above induction step, we have additionally shown that each $k \in [n]$ satisfies $j_k \leq k$ and thus $j_k \in [k]$. Hence, $(j_1, j_2, \ldots, j_n) \in [1] \times [2] \times \cdots \times [n]$. This proves one of our two goals. It remains to prove the other goal, which is $\Phi(j_1, j_2, \ldots, j_n) = w$.

This is not too hard any more. First, we observe that the permutation $w_0$ fixes $1, 2, \ldots, n$ (by (29), applied to $k = 0$), and thus equals id. In other words, $w_0 = \mathrm{id}$. Next, we note that each $k \in [n]$ satisfies

$$w_k = w_{k-1} t_{j_k,k} \qquad \text{(by (27))}. \tag{30}$$

[5] A map $f$ is said to *fix* an element $x$ if $f(x) = x$.

Now,

$$\begin{aligned} w = w_n &= \underbrace{w_{n-1}}_{\substack{=w_{n-2}t_{j_{n-1},n-1} \\ \text{(by (30))}}} t_{j_n,n} \qquad \text{(by (30))} \\ &= \underbrace{w_{n-2}}_{\substack{=w_{n-3}t_{j_{n-2},n-2} \\ \text{(by (30))}}} t_{j_{n-1},n-1}t_{j_n,n} \\ &= \underbrace{w_{n-3}}_{\substack{=w_{n-4}t_{j_{n-3},n-3} \\ \text{(by (30))}}} t_{j_{n-2},n-2}t_{j_{n-1},n-1}t_{j_n,n} \\ &= \cdots \\ &= \underbrace{w_0}_{=\mathrm{id}} t_{j_1,1}t_{j_2,2}\cdots t_{j_n,n} = t_{j_1,1}t_{j_2,2}\cdots t_{j_n,n} = \Phi\left(j_1, j_2, \ldots, j_n\right). \end{aligned}$$

Thus, we have written our (arbitrary) permutation $w \in S_n$ as $\Phi\left(j_1, j_2, \ldots, j_n\right)$ for some $\left(j_1, j_2, \ldots, j_n\right) \in [1] \times [2] \times \cdots \times [n]$. This proves that $\Phi$ is surjective. Thus, our second proof of Theorem 3.4.4 is complete. □

(Note that Theorem 3.4.4 is equivalent to [Grinbe15, Exercise 5.9], since $t_{j_k,k} = t_{k,j_k}$ for all $k$ and $j_k$.)

Theorem 3.4.3 now easily follows: Indeed, we have

$$\left(1+\mathbf{m}_1\right)\left(1+\mathbf{m}_2\right)\cdots\left(1+\mathbf{m}_n\right) = \sum_{\left(j_1,j_2,\ldots,j_n\right)\in[1]\times[2]\times\cdots\times[n]} t_{j_1,1}t_{j_2,2}\cdots t_{j_n,n} = \sum_{w\in S_n} w$$

(here, we have substituted $w$ for $t_{j_1,1}t_{j_2,2}\cdots t_{j_n,n}$ in the sum, since Theorem 3.4.4 shows that the map

$$\begin{aligned} [1]\times[2]\times\cdots\times[n] &\to S_n, \\ \left(j_1, j_2, \ldots, j_n\right) &\mapsto t_{j_1,1}t_{j_2,2}\cdots t_{j_n,n} \end{aligned}$$

is a bijection), so that

$$\left(1+\mathbf{m}_1\right)\left(1+\mathbf{m}_2\right)\cdots\left(1+\mathbf{m}_n\right) = \sum_{w\in S_n} w = \nabla$$

(by the definition of $\nabla$). Thus, Theorem 3.4.3 is proved. □

We will soon give a different (not in substance, but in presentation) proof of Theorem 3.4.3, which has the advantage of being slicker. But the above proof has its own claim to importance: It shows that Theorem 3.4.3 is, at heart, a combinatorial result (Theorem 3.4.4). Indeed, while we derived Theorem 3.4.3 from Theorem 3.4.4, we could just as well revert this argument and obtain Theorem 3.4.4 from Theorem 3.4.3 (provided that the latter theorem has been

proved in some other way), since a sum of elements of $\mathbf{k}[S_n]$ equals $\nabla$ if and only if this sum contains each permutation $w \in S_n$ exactly once. (To be precise, this reasoning requires $\mathbf{k}$ to be $\mathbb{Z}$ or a ring that contains $\mathbb{Z}$.) The takeaway from this is that (almost) every identity that we prove in $\mathbf{k}[S_n]$ can (by expanding its two sides) be turned into a combinatorial statement about permutations. The former usually looks slicker, but the latter gives it intuition and grounding.

### 3.4.3. More about jucys–murphies

**Exercise 3.4.1.** $\boxed{2}$ Let $i \in [n]$. Show the following:

**(a)** We have $s_i \mathbf{m}_j = \mathbf{m}_j s_i$ for all $j \in [n] \setminus \{i, i+1\}$.

**(b)** We have $s_i \mathbf{m}_i + 1 = \mathbf{m}_{i+1} s_i$.

**(c)** We have $s_i \mathbf{m}_{i+1} - 1 = \mathbf{m}_i s_i$.

Soon, we will compute $\mathbf{m}_2 \mathbf{m}_3 \cdots \mathbf{m}_n$ as well (see Corollary 3.8.21). But let us try it out for $n = 3$:

$$\mathbf{m}_2 \mathbf{m}_3 = t_{1,2} (t_{1,3} + t_{2,3}) = t_{1,2} t_{1,3} + t_{1,2} t_{2,3} = \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2}.$$

This is the sum of the two 3-cycles in $S_3$ (that is, the cycles of the form $\mathrm{cyc}_{i,j,k}$). Is this the beginning of a pattern?

Here is another natural question (inspired by Corollary 3.3.4, which shows that $P(\nabla) = 0$ for the polynomial $P = x(x - n!)$): What polynomials $P \in \mathbb{Z}[x]$ satisfy $P(\mathbf{m}_k) = 0$ for a given $k$ ? Let us see:

- For $k = 1$, this is easy: We have $\mathbf{m}_1 = 0$, so we can pick $P = x$.
- For $k = 2$, this is still easy: We have $\mathbf{m}_2 = t_{1,2} = s_1$, so $\mathbf{m}_2^2 = s_1^2 = \mathrm{id} = 1$ and therefore $\mathbf{m}_2^2 - 1 = 0$. Thus, we can pick $P = x^2 - 1 = (x-1)(x+1)$.
- For $k = 3$, we have $\mathbf{m}_3 = t_{1,3} + t_{2,3}$, and a bit of computation shows that
$$(\mathbf{m}_3 - 2)(\mathbf{m}_3 - 1)(\mathbf{m}_3 + 1)(\mathbf{m}_3 + 2) = 0,$$
so that we can pick $P = (x-2)(x-1)(x+1)(x+2)$.

In each of these three cases, we observe that $P$ is a polynomial that splits over $\mathbb{Z}$ (meaning that it is a product of linear factors $x - r$ for integers $r$). This generalizes to all $k$ and $n$:

**Theorem 3.4.5.** Let $k \in [n]$. Then,

$$\prod_{i=-(k-1)}^{k-1} (\mathbf{m}_k - i) = 0.$$

Alas, this theorem is too hard to be proved right now. The most elementary proof I know for this theorem uses some tricky linear algebra (see Igor Makhlin's MathOverflow post [Makhli11]). Another proof uses the so-called *seminormal basis* of $\mathbf{k}[S_n]$, one of the main results of this course. I will do this later in this course.

Note that the factor $\mathbf{m}_k - 0$ in Theorem 3.4.5 is unnecessary when $k = 2$ or $k = 3$, as we have seen above. However, all $2k-1$ factors in the product in Theorem 3.4.5 are necessary for $k \geq 4$ (as long as $\mathbf{k} = \mathbb{Z}$).

## 3.5. Interlude: Subalgebras generated by a set

### 3.5.1. The general case

Next, we recall/introduce a general concept in algebra:

**Definition 3.5.1.** Let $A$ be a $\mathbf{k}$-algebra. Let $U$ be a subset of $A$. Then, the *subalgebra of $A$ generated by $U$* is defined to be the subset of $A$ consisting of all (finite) $\mathbf{k}$-linear combinations of products of elements of $U$. In other words, it is the $\mathbf{k}$-linear span

$$\operatorname{span}_{\mathbf{k}} \{u_1 u_2 \cdots u_i \mid i \in \mathbb{N} \text{ and } u_1, u_2, \ldots, u_i \in U\}.$$

(Note that $i$ is allowed to be 0, in which case $u_1 u_2 \cdots u_i = (\text{empty product in } A) = 1_A$.)

This $\mathbf{k}$-linear span is really a $\mathbf{k}$-subalgebra of $A$ that contains $U$ as a subset. Moreover, it is the smallest such $\mathbf{k}$-subalgebra (i.e., it is a subset of any such $\mathbf{k}$-subalgebra). We denote it by $\operatorname{AlgGen} U$ or by $\mathbf{k}[U]$.

**Exercise 3.5.1.** [1] Prove the above claims.

**Example 3.5.2.** **(a)** The subalgebra of the $\mathbb{Z}$-algebra $\mathbb{R}$ generated by the set $\left\{\sqrt{2}\right\}$ is the $\mathbb{Z}$-subalgebra

$$\left\{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\right\}$$

of $\mathbb{R}$. Indeed, any product of the form $\sqrt{2}\sqrt{2}\cdots\sqrt{2}$ is either an integer or $\sqrt{2}$ times an integer, and thus belongs to the set $\left\{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\right\}$; therefore, any $\mathbb{Z}$-linear combination of such products belongs to this set as well.

**(b)** The subalgebra of the $\mathbb{Z}$-algebra $\mathbb{R}$ generated by the set $\left\{\sqrt[3]{2}\right\}$ is the $\mathbb{Z}$-subalgebra

$$\left\{a + b\sqrt[3]{2} + c\sqrt[3]{4} \mid a, b, c \in \mathbb{Z}\right\}$$

of $\mathbb{R}$.

**(c)** More generally, if $A$ is any $\mathbf{k}$-algebra, and if $u$ is any element of $A$, then the subalgebra of $A$ generated by the one-element subset $\{u\}$ is the $\mathbf{k}$-linear span of all powers $u^0, u^1, u^2, \ldots$ of $u$. In other words, it is the set of all polynomials in $u$ (that is, of all values $p(u)$ for $p \in \mathbf{k}[x]$).

**(d)** Now consider the polynomial ring $\mathbf{k}[x]$ as a $\mathbf{k}$-algebra. The subalgebra of $\mathbf{k}[x]$ generated by the set $\left\{x^2, x^3\right\}$ is the set of all polynomials $p \in \mathbf{k}[x]$ whose $x^1$-coefficient is 0 (that is, whose derivative at 0 is 0). Proving this is a quick way to check your understanding of polynomials and subalgebras.

**(e)** The Laurent polynomial ring $\mathbf{k}\left[x^{\pm 1}\right]$ is its own subalgebra generated by the two-element set $\left\{x, x^{-1}\right\}$. This is just saying that every Laurent polynomial in $x$ can be written as a (usual, not Laurent) polynomial in $x$ and $x^{-1}$.

The following exercise gives another example for subalgebras generated by a subset:

**Exercise 3.5.2.** Consider the Laurent polynomial ring $\mathbf{k}\left[x^{\pm 1}\right]$.

**(a)** [2] Prove that

$$\left\{p \in \mathbf{k}\left[x^{\pm 1}\right] \mid p\left(x^{-1}\right) = p(x)\right\} = \operatorname{AlgGen}\left\{x + x^{-1}\right\}.$$

In other words, the Laurent polynomials $p \in \mathbf{k}\left[x^{\pm 1}\right]$ that satisfy $p\left(x^{-1}\right) = p(x)$ are precisely the polynomials in $x + x^{-1}$.

**(b)** [1] Assume that 2 is a regular element of $\mathbf{k}$ (meaning that $2u = 0$ implies $u = 0$ whenever $u \in \mathbf{k}$). Prove that

$$\left\{p \in \mathbf{k}\left[x^{\pm 1}\right] \mid p(-x) = p(x)\right\} = \operatorname{AlgGen}\left\{x^2, x^{-2}\right\}$$

and

$$\{p \in \mathbf{k}[x] \mid p(-x) = p(x)\} = \operatorname{AlgGen}\left\{x^2\right\}.$$

The following proposition is easy but quite important, as it provides an easy way to prove the commutativity of an algebra:

**Proposition 3.5.3.** Let $A$ and $U$ be as in Definition 3.5.1. Assume that every two elements of $U$ commute. Then, the $\mathbf{k}$-algebra $\operatorname{AlgGen} U$ is commutative.

**Exercise 3.5.3.** [2] Prove this.

### 3.5.2. The Gelfand–Tsetlin algebra

Thanks to Proposition 3.5.3, the commutativity of an algebra needs only to be checked on its generators. This lets us easily show facts like the following:

**Corollary 3.5.4.** Let $\mathrm{GZ}_n$ be the subalgebra of $\mathbf{k}[S_n]$ generated by the set $\{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n\}$ of its jucys–murphies. (In other words, let $\mathrm{GZ}_n := \mathrm{AlgGen}\{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n\}$.) Then, $\mathrm{GZ}_n$ is a commutative $\mathbf{k}$-algebra.

*Proof.* Theorem 3.4.2 shows that every two elements of $\{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n\}$ commute. Thus, Proposition 3.5.3 (applied to $A = \mathbf{k}[S_n]$ and $U = \{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n\}$) shows that $\mathrm{AlgGen}\{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n\}$ is commutative. This proves Corollary 3.5.4. □

The subalgebra $\mathrm{GZ}_n$ defined in Corollary 3.5.4 is called the *Gelfand–Tsetlin subalgebra* (or *Gelfand–Cetlin subalgebra*, or *Gelfand–Zetlin subalgebra*) of $\mathbf{k}[S_n]$, and we will eventually learn some more of its properties.

## 3.6. The somewhere-to-below shuffles

Here is a newer family of elements of $\mathbf{k}[S_n]$, introduced by Lafrenière and Grinberg in [GriLaf22]:

**Definition 3.6.1.** The *somewhere-to-below shuffles* $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_n$ are the elements of $\mathbf{k}[S_n]$ defined by

$$\mathbf{t}_k := \mathrm{cyc}_k + \mathrm{cyc}_{k,k+1} + \mathrm{cyc}_{k,k+1,k+2} + \cdots + \mathrm{cyc}_{k,k+1,\ldots,n} = \sum_{i=k}^{n} \mathrm{cyc}_{k,k+1,\ldots,i} \in \mathbf{k}[S_n].$$

(Recall that $\mathrm{cyc}_k = \mathrm{id}$.)

In particular, $\mathbf{t}_n = \mathrm{cyc}_n = \mathrm{id} = 1$. The motivation for calling these elements "shuffles" will be eventually explained.

The following exercise generalizes (16) (why?):

**Exercise 3.6.1.** [2] Prove that $\mathbf{t}_n \mathbf{t}_{n-1} \cdots \mathbf{t}_1 = \nabla$.

The somewhere-to-below shuffles $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_n$ do not commute for $n \geq 3$. For instance, for $n = 3$, we have

$$\mathbf{t}_1 \mathbf{t}_2 - \mathbf{t}_2 \mathbf{t}_1 = t_{1,2} + \mathrm{cyc}_{1,2,3} - \mathrm{cyc}_{1,3,2} - t_{1,3} \neq 0.$$

However, they come fairly close to commuting. To make sense of this, let us define the *commutator* in a ring:

**Definition 3.6.2.** Let $a$ and $b$ be two elements of a ring. Then, their *commutator* $[a, b]$ is defined to be $ab - ba$.

Thus, two elements $a$ and $b$ satisfy $[a, b] = 0$ if and only if $a$ and $b$ commute.

Our above computation shows that $[\mathbf{t}_1, \mathbf{t}_2] \neq 0$ (for $n \geq 3$). However, one can also show that $[\mathbf{t}_1, \mathbf{t}_2]^2 = 0$. Thus, $[\mathbf{t}_1, \mathbf{t}_2]$ is a nilpotent element of $\mathbf{k}[S_n]$, with nilpotency order[6] 2. For comparison, a strictly upper-triangular $k \times k$-matrix is nilpotent with nilpotency order $\leq k$. Is there a connection? Yes, and we will explore this later.

This nilpotency furthermore generalizes: We have

$$\left[\mathbf{t}_i, \mathbf{t}_j\right]^{j-i+1} = 0 \qquad \text{for all } 1 \leq i < j \leq n.$$

This is [Grinbe23, Corollary 9.11]; I might outline a proof later.

**Exercise 3.6.2.** $\boxed{3}$ Prove that $\mathbf{t}_{i+1}\mathbf{t}_i = (\mathbf{t}_i - 1)\,\mathbf{t}_i$ for all $i \in [n-1]$.

We can again ask for polynomials $P \in \mathbf{k}[x]$ that annihilate these $\mathbf{t}_i$'s. For example, it turns out ([Grinbe18, Theorem 3]) that

$$P(\mathbf{t}_1) = 0 \qquad \text{for } P = x(x-1)(x-2)\cdots(x-(n-2))(x-n).$$

More generally,

$$P(\mathbf{t}_i) = 0 \qquad \text{for } P = x(x-1)(x-2)\cdots(x-(n-i-1))(x-(n-i+1)).$$

We might prove this eventually.

## 3.7. Partial integrals

### 3.7.1. Definition and properties

Let us now generalize the integral $\nabla$ and the sign-integral $\nabla^-$:

**Definition 3.7.1.** Let $X$ be a subset of $[n]$. Then, we define two elements $\nabla_X$ and $\nabla_X^-$ of $\mathbf{k}[S_n]$ by

$$\nabla_X := \sum_{\substack{w \in S_n; \\ w(i)=i \text{ for all } i \notin X}} w \qquad \text{and} \qquad \nabla_X^- := \sum_{\substack{w \in S_n; \\ w(i)=i \text{ for all } i \notin X}} (-1)^w w.$$

(Here, "for all $i \notin X$" means "for all $i \in [n] \setminus X$". Thus, both sums range over all permutations $w \in S_n$ that fix all elements $i \in [n]$ that don't belong to $X$.)

We shall call $\nabla_X$ the *$X$-integral*, and we shall call $\nabla_X^-$ the *$X$-sign-integral*.

[6] The *nilpotency order* (or *order of nilpotency*) of a nilpotent element $a$ of a ring $A$ is defined to be the smallest $k \in \mathbb{N}$ such that $a^k = 0$.

**Example 3.7.2.** Let $X$ be a subset of $[n]$.

**(a)** If $|X| = 0$, then $X = \varnothing$ and therefore $\nabla_X = \nabla_X^- = 1$ (because if a permutation $w \in S_n$ satisfies $w(i) = i$ for all $i \notin \varnothing$, then it simply satisfies $w(i) = i$ for all $i \in [n]$, which means that $w = \mathrm{id} = 1$).

**(b)** If $|X| = 1$, then we again have $\nabla_X = \nabla_X^- = 1$ (because if a permutation $w \in S_n$ fixes all but one element of $[n]$, then it must also fix the remaining element, and thus it equals $\mathrm{id} = 1$).

**(c)** If $|X| = 2$, then $X = \{i, j\}$ for some $i \neq j$, and thus

$$\nabla_X = 1 + t_{i,j} \qquad \text{and} \qquad \nabla_X^- = 1 - t_{i,j}$$

(since the permutations $w \in S_n$ that fix all elements of $[n]$ except for $i$ and $j$ are precisely $\mathrm{id} = 1$ and $t_{i,j}$).

**(d)** If $|X| = 3$, then $X = \{i, j, k\}$ for some distinct $i, j, k$, and thus

$$\begin{aligned} \nabla_X &= 1 + t_{i,j} + t_{i,k} + t_{j,k} + \mathrm{cyc}_{i,j,k} + \mathrm{cyc}_{i,k,j} \qquad \text{and} \\ \nabla_X^- &= 1 - t_{i,j} - t_{i,k} - t_{j,k} + \mathrm{cyc}_{i,j,k} + \mathrm{cyc}_{i,k,j}\,. \end{aligned}$$

(Here, we have used the fact that any transposition $t_{p,q}$ has sign $-1$ whereas any cycle $\mathrm{cyc}_{p,q,r}$ has sign 1.)

**Example 3.7.3.** The $[n]$-integral $\nabla_{[n]}$ just equals the usual integral $\nabla$. Indeed, the definition of $\nabla_{[n]}$ yields

$$\nabla_{[n]} = \sum_{\substack{w \in S_n; \\ w(i) = i \text{ for all } i \notin [n]}} w = \sum_{w \in S_n} w$$

(since the condition "$w(i) = i$ for all $i \notin [n]$" is vacuously true for any $w \in S_n$), so that

$$\nabla_{[n]} = \sum_{w \in S_n} w = \nabla.$$

Similarly, $\nabla_{[n]}^- = \nabla^-$.

The following proposition collects a number of basic properties of $\nabla_X$:

**Proposition 3.7.4.** Let $X$ be a subset of $[n]$. Let

$$S_{n,X} := \{w \in S_n \mid w(i) = i \text{ for all } i \notin X\}.$$

Then:

**(a)** The set $S_{n,X}$ is a subgroup of $S_n$, and is isomorphic to $S_X$ via the group

isomorphism

$$\begin{aligned} S_{n,X} &\to S_X, \\ w &\mapsto w\,|_X^X\,. \end{aligned}$$

(Here, $w\,|_X^X$ means the restriction of $w$ to $X$, regarded as a map from $X$ to $X$.)

**(b)** We have

$$\nabla_X = \sum_{w \in S_{n,X}} w \qquad \text{and} \qquad \nabla_X^- = \sum_{w \in S_{n,X}} (-1)^w\, w.$$

**(c)** For each $w \in S_{n,X}$, we have

$$\begin{aligned} w\nabla_X &= \nabla_X w = \nabla_X \qquad \text{and} \\ w\nabla_X^- &= \nabla_X^- w = (-1)^w\, \nabla_X^-. \end{aligned}$$

**(d)** Let $i$ and $j$ be two distinct elements of $X$. Let us set

$$\nabla_X^{\text{even}} := \sum_{\substack{w \in S_{n,X}; \\ (-1)^w = 1}} w.$$

Then,

$$\nabla_X = \nabla_X^{\text{even}} \cdot \left(1 + t_{i,j}\right) = \left(1 + t_{i,j}\right) \cdot \nabla_X^{\text{even}} \tag{31}$$

and

$$\nabla_X^- = \nabla_X^{\text{even}} \cdot \left(1 - t_{i,j}\right) = \left(1 - t_{i,j}\right) \cdot \nabla_X^{\text{even}}. \tag{32}$$

**Example 3.7.5.** Let $X$ be a three-element subset $\{i, j, k\}$ of $[n]$. Then, the element $\nabla_X^{\text{even}}$ in Proposition 3.7.4 **(d)** is given by

$$\nabla_X^{\text{even}} = 1 + \text{cyc}_{i,j,k} + \text{cyc}_{i,k,j}$$

(since the even permutations $w \in S_{n,X}$ are $\text{id} = 1$ and $\text{cyc}_{i,j,k}$ and $\text{cyc}_{i,k,j}$). Hence, for instance, the first equality in (31) says that

$$\underbrace{1 + t_{i,j} + t_{i,k} + t_{j,k} + \text{cyc}_{i,j,k} + \text{cyc}_{i,k,j}}_{=\nabla_X} = \underbrace{\left(1 + \text{cyc}_{i,j,k} + \text{cyc}_{i,k,j}\right)}_{=\nabla_X^{\text{even}}} \cdot \left(1 + t_{i,j}\right).$$

*Proof of Proposition 3.7.4.* **(a)** We must prove the following claims:

*Claim 1:* For any $w \in S_{n,X}$, we have $w\,(X) \subseteq X$ (so that the restriction $w\,|_X^X$ is really a well-defined map from $X$ to $X$) and $w\,|_X^X \in S_X$.

*Claim 2:* The set $S_{n,X}$ is a subgroup of $S_n$.

*Claim 3:* The map

$$\begin{aligned} S_{n,X} &\to S_X, \\ w &\mapsto w\left|_X^X\right. \end{aligned}$$

is a group morphism.

*Claim 4:* This group morphism is a group isomorphism.

All these claims are easy, so we just outline their proofs:

*Proof of Claim 1 (sketched).* Let $w \in S_{n,X}$. Let $x \in X$. From $w \in S_{n,X}$, we obtain $w(i) = i$ for all $i \notin X$. Thus, if we had $w(x) \notin X$, then we would have $w(w(x)) = w(x)$ (by applying the equality $w(i) = i$ to $i = w(x)$), which would entail $w(x) = x$ (since $w$ is a permutation and thus injective), and therefore $x = w(x) \notin X$, which would contradict $x \in X$. Hence, we cannot have $w(x) \notin X$. Thus, $w(x) \in X$.

Forget that we fixed $x$. We thus have shown that $w(x) \in X$ for each $x \in X$. That is, $w(X) \subseteq X$.

This yields that the restriction $w\left|_X^X\right.$ is well-defined as a map from $X$ to $X$. It remains to show that it is a permutation of $X$ (that is, belong to $S_X$). It clearly suffices to show that it is injective and surjective. The injectivity follows from the injectivity of $w$ (which is because $w$ is a permutation), since a restriction of an injective map is again injective. Now, let us prove the surjectivity: Let $y \in X$. Let $x = w^{-1}(y)$. Thus, $w(x) = y$. From $w \in S_{n,X}$, we obtain $w(i) = i$ for all $i \notin X$. Thus, if we had $x \notin X$, then we would have $w(x) = x$ and thus $x = w(x) = y \in X$, which would contradict $x \notin X$. Hence, we cannot have $x \notin X$. So we must have $x \in X$. Thus, $\left(w\left|_X^X\right.\right)(x) = w(x) = y$. This shows that $y$ is an image under the map $w\left|_X^X\right.$. Since we have proved this for each $y \in X$, we have thus shown that each element of $X$ is an image under $w\left|_X^X\right.$. In other words, $w\left|_X^X\right.$ is surjective. This completes our proof of Claim 1. □

*Proof of Claim 2 (sketched).* This is straightforward. For example, if $u, v \in S_{n,X}$, then $uv \in S_{n,X}$, since each $i \notin X$ satisfies $(uv)(i) = u\left(\underbrace{v(i)}_{=i}\right) = u(i) = i$. □

*Proof of Claim 3 (sketched).* This is again straightforward. (The well-definedness of the map follows from Claim 1.) □

*Proof of Claim 4 (sketched).* Let us denote this group morphism by $\Phi$.

Each permutation $w \in S_{n,X}$ is uniquely determined by its values on the elements of $X$ (since all its other values are uniquely determined by the requirement "$w(i) = i$ for all $i \notin X$"), that is, by its restriction $\Phi(w) = w\left|_X^X\right.$. Hence, the map $\Phi$ is injective.

Any permutation $u \in S_X$ can be extended to a permutation $w \in S_{n,X}$ satisfying $\Phi(w) = u$ (we just need to set $w(x) = u(x)$ for all $x \in X$, and $w(i) = i$ for all $i \notin X$). Thus, the map $\Phi$ is surjective.

So the map $\Phi$ is both injective and surjective, hence bijective. Since $\Phi$ is a group morphism, it thus follows that $\Phi$ is a group isomorphism. This proves Claim 4. □

With these four claims proved, Proposition 3.7.4 **(a)** follows.

**(b)** This is just a restatement of the definitions of $\nabla_X$ and $\nabla_X^-$, since the permutations $w \in S_n$ that satisfy $w(i) = i$ for all $i \notin X$ are precisely the permutations $w \in S_{n,X}$.

**(c)** This is a generalization of Proposition 3.3.3, and is proved in the exact same way (using the formulas for $\nabla_X$ and $\nabla_X^-$ given in Proposition 3.7.4 **(b)**), since $S_{n,X}$ is a group.

**(d)** Let $i$ and $j$ be two distinct elements of $X$. Then, Proposition 3.7.4 **(b)** yields

$$\nabla_X = \sum_{w \in S_{n,X}} w = \sum_{\substack{w \in S_{n,X}; \\ (-1)^w = 1}} w + \sum_{\substack{w \in S_{n,X}; \\ (-1)^w = -1}} w \tag{33}$$

(since each $w \in S_{n,X}$ satisfies either $(-1)^w = 1$ or $(-1)^w = -1$, but not both at the same time). Let us now simplify the second sum on the right hand side. Indeed, $S_{n,X}$ is a group, and $t_{i,j}$ is one of its elements (since both $i$ and $j$ belong to $X$, and thus the transposition $t_{i,j}$ fixes all elements $k \notin X$). Hence, the map

$$\begin{aligned} S_{n,X} &\to S_{n,X}, \\ w &\mapsto w t_{i,j} \end{aligned}$$

is well-defined. This map is furthermore a bijection (with inverse given by $w \mapsto w t_{i,j}^{-1}$). Thus, we can substitute $w t_{i,j}$ for $w$ in the sum $\sum_{\substack{w \in S_{n,X}; \\ (-1)^w = -1}} w$. We thus obtain

$$\sum_{\substack{w \in S_{n,X}; \\ (-1)^w = -1}} w = \sum_{\substack{w \in S_{n,X}; \\ (-1)^{w t_{i,j}} = -1}} w t_{i,j}. \tag{34}$$

However, the transposition $t_{i,j}$ has sign $(-1)^{t_{i,j}} = -1$ (like any transposition), and thus each $w \in S_{n,X}$ satisfies

$$\begin{aligned} (-1)^{w t_{i,j}} &= (-1)^w \underbrace{(-1)^{t_{i,j}}}_{=-1} \qquad \left(\text{by the rule } (-1)^{\sigma\tau} = (-1)^\sigma \cdot (-1)^\tau\right) \\ &= -(-1)^w. \end{aligned}$$

Hence, the condition "$(-1)^{w t_{i,j}} = -1$" under the summation sign on the right hand side of (34) is equivalent to the condition "$-(-1)^w = -1$", which is in turn equivalent to "$(-1)^w = 1$". Therefore, we can replace the former condition by the latter. Thus, (34) rewrites as

$$\sum_{\substack{w \in S_{n,X}; \\ (-1)^w = -1}} w = \sum_{\substack{w \in S_{n,X}; \\ (-1)^w = 1}} w t_{i,j} = \underbrace{\left( \sum_{\substack{w \in S_{n,X}; \\ (-1)^w = 1}} w \right)}_{= \nabla_X^{\text{even}}} t_{i,j} = \nabla_X^{\text{even}} t_{i,j}.$$

Hence, (33) becomes

$$\nabla_X = \underbrace{\sum_{\substack{w \in S_{n,X}; \\ (-1)^w = 1}} w}_{= \nabla_X^{\text{even}}} + \underbrace{\sum_{\substack{w \in S_{n,X}; \\ (-1)^w = -1}} w}_{= \nabla_X^{\text{even}} t_{i,j}} = \nabla_X^{\text{even}} + \nabla_X^{\text{even}} t_{i,j} = \nabla_X^{\text{even}} \cdot \left(1 + t_{i,j}\right).$$

A similar argument (but using the bijection $w \mapsto t_{i,j} w$ instead of $w \mapsto w t_{i,j}$) shows that $\nabla_X = \left(1 + t_{i,j}\right) \cdot \nabla_X^{\text{even}}$. Combining these two equalities, we obtain (31). The proof of (32) is similar, except that the signs need to be taken into account. Proposition 3.7.4 **(d)** is thus completely proved. □

We note that the $t_{i,j}$ in Proposition 3.7.4 **(d)** could be replaced by any odd permutation $\tau \in S_{n,X}$:

**Proposition 3.7.6.** Let $X$ be a subset of $[n]$. Let $\tau \in S_{n,X}$ be an odd permutation (where $S_{n,X}$ is as in Proposition 3.7.4). Let us set

$$\nabla_X^{\text{even}} := \sum_{\substack{w \in S_{n,X}; \\ (-1)^w = 1}} w.$$

Then,

$$\nabla_X = \nabla_X^{\text{even}} \cdot (1 + \tau) = (1 + \tau) \cdot \nabla_X^{\text{even}} \tag{35}$$

and

$$\nabla_X^- = \nabla_X^{\text{even}} \cdot (1 - \tau) = (1 - \tau) \cdot \nabla_X^{\text{even}}. \tag{36}$$

*Proof.* Exactly the same as for Proposition 3.7.4 **(d)**. □

**Remark 3.7.7.** The equality (35) can be generalized even further to arbitrary finite groups, albeit at the cost of concreteness. Indeed, let $G$ be a finite group, and let $H$ be a subgroup of $G$. Then, we can find a subset $P$ of $G$ such that each element of $G$ can be uniquely expressed as $hp$ with $h \in H$ and $p \in P$. (Such a subset $P$ is called a *right transversal of $H$ in $G$*, or a *system of representatives for the right cosets of $H$ in $G$*.) For such a subset $P$, we then get

$$\sum_{g \in G} g = \left(\sum_{h \in H} h\right) \left(\sum_{p \in P} p\right).$$

For $G = S_{n,X}$ and $H = S_{n,X}^{\text{even}} = \left\{w \in S_{n,X} \mid (-1)^w = 1\right\}$ and $P = \{1, \tau\}$, this recovers the first equality in (35).

We can also find a subset $Q$ of $G$ such that each element of $G$ can be uniquely expressed as $qh$ with $h \in H$ and $q \in Q$. Then we have

$$\sum_{g \in G} g = \left(\sum_{q \in Q} q\right) \left(\sum_{h \in H} h\right).$$

More interestingly, we can find a subset $R$ of $G$ that simultaneously satisfies the requirements on $P$ and on $Q$ (by a result of Miller from 1910 – see `https://math.stackexchange.com/questions/178186` for some references), so that we get

$$\sum_{g\in G} g = \left(\sum_{h\in H} h\right)\left(\sum_{r\in R} r\right) = \left(\sum_{r\in R} r\right)\left(\sum_{h\in H} h\right).$$

In the case of $G = S_{n,X}$ and $H = S_{n,X}^{\text{even}}$, we can take $R = \{1, \tau\}$ for example.

### 3.7.2. A recursion

We can also use $X$-integrals to give a new proof of Theorem 3.4.3. For this, we will show the following lemma:[7]

**Lemma 3.7.8.** For any $k \in [n]$, we have $\nabla_{[k]} = \nabla_{[k-1]}\left(1 + \mathbf{m}_k\right)$.

If this lemma is proved, then we get

$$\begin{aligned}
\nabla &= \nabla_{[n]} \qquad \text{(by Example 3.7.3)} \\
&= \underbrace{\nabla_{[n-1]}}_{\substack{=\nabla_{[n-2]}\left(1+\mathbf{m}_{n-1}\right) \\ \text{(by Lemma 3.7.8)}}} \left(1 + \mathbf{m}_n\right) \qquad \text{(by Lemma 3.7.8)} \\
&= \underbrace{\nabla_{[n-2]}}_{\substack{=\nabla_{[n-3]}\left(1+\mathbf{m}_{n-2}\right) \\ \text{(by Lemma 3.7.8)}}} \left(1 + \mathbf{m}_{n-1}\right)\left(1 + \mathbf{m}_n\right) \\
&= \cdots \\
&= \underbrace{\nabla_{[0]}}_{\substack{=1 \\ \text{(by Example 3.7.2 \textbf{(a)})}}} \left(1 + \mathbf{m}_1\right)\left(1 + \mathbf{m}_2\right)\cdots\left(1 + \mathbf{m}_n\right) \\
&= \left(1 + \mathbf{m}_1\right)\left(1 + \mathbf{m}_2\right)\cdots\left(1 + \mathbf{m}_n\right),
\end{aligned}$$

which proves Theorem 3.4.3 again. So we only need to prove Lemma 3.7.8 in order to obtain a second proof of Theorem 3.4.3.

We shall not prove Lemma 3.7.8 immediately, but rather extend it first, as this allows us to get a more general result at the same price. First, we extend it by replacing the subset $[k]$ and its element $k$ with an arbitrary subset $X$ and any element $x \in X$:

[7] Recall that $[i] = \{1, 2, \ldots, i\}$ for each $i \in \mathbb{N}$. Also, the Young–Jucys–Murphy elements $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n$ were defined in Definition 3.4.1.

**Lemma 3.7.9.** Let $X$ be any subset of $[n]$, and let $x \in X$. Then,

$$\nabla_X = \nabla_{X\setminus\{x\}} \left(1 + \sum_{y \in X\setminus\{x\}} t_{y,x}\right).$$

Even more generally, we can replace each transposition $t_{y,x}$ on the right hand side by an arbitrary permutation $\sigma_y \in S_{n,X}$ that sends $x$ to $y$ (where $S_{n,X}$ is as in Proposition 3.7.4), and also replace the 1 by an arbitrary permutation $\sigma_x \in S_{n,X}$ that sends $x$ to $x$. Thus we arrive at the following generalization:

**Lemma 3.7.10.** Let $X$ be any subset of $[n]$, and let $x \in X$. Let $S_{n,X}$ be as in Proposition 3.7.4. For each $y \in X$, fix a permutation $\sigma_y \in S_{n,X}$ that sends $y$ to $x$. Then,

$$\nabla_X = \nabla_{X\setminus\{x\}} \sum_{y \in X} \sigma_y.$$

*Proof of Lemma 3.7.10.* We first observe that

$$\{w \in S_{n,X} \mid w(x) = x\} = S_{n,X\setminus\{x\}}. \tag{37}$$

*Proof of (37).* We have $x \in X$ and thus $\{x\} \subseteq X$.

Recall that $S_{n,X} = \{w \in S_n \mid w(i) = i \text{ for all } i \notin X\}$ (by the definition of $S_{n,X}$). Thus,

$$\begin{aligned}
&\{w \in S_{n,X} \mid w(x) = x\} \\
&= \{w \in S_n \mid w(i) = i \text{ for all } i \notin X, \text{ and moreover } w(x) = x\} \\
&= \{w \in S_n \mid w(i) = i \text{ for all } i \notin X \text{ and also for } i = x\} \\
&= \{w \in S_n \mid w(i) = i \text{ for all } i \in [n] \setminus X \text{ and also for } i = x\} \\
&= \{w \in S_n \mid w(i) = i \text{ for all } i \in ([n] \setminus X) \cup \{x\}\} \\
&= \{w \in S_n \mid w(i) = i \text{ for all } i \in [n] \setminus (X \setminus \{x\})\} \\
&\qquad \left(\begin{array}{c}\text{since } ([n] \setminus X) \cup \{x\} = [n] \setminus (X \setminus \{x\}) \\ \text{(because } \{x\} \subseteq X \subseteq [n]\text{)}\end{array}\right) \\
&= \{w \in S_n \mid w(i) = i \text{ for all } i \notin X \setminus \{x\}\} \\
&= S_{n,X\setminus\{x\}} \qquad \left(\text{by the definition of } S_{n,X\setminus\{x\}}\right).
\end{aligned}$$

This proves (37). □

Proposition 3.7.4 **(b)** yields

$$\nabla_X = \sum_{w \in S_{n,X}} w \qquad \text{and} \tag{38}$$

$$\nabla_{X\setminus\{x\}} = \sum_{w \in S_{n,X\setminus\{x\}}} w. \tag{39}$$

If $w \in S_{n,X}$, then there is a unique $y \in X$ such that $w(y) = x$ [8]. Hence, we can split the sum $\sum_{w \in S_{n,X}} w$ according to the value of this $y$. We thus obtain

$$\sum_{w \in S_{n,X}} w = \sum_{y \in X} \sum_{\substack{w \in S_{n,X}; \\ w(y) = x}} w. \tag{40}$$

Let us next show that each $y \in X$ satisfies

$$\sum_{\substack{w \in S_{n,X}; \\ w(y) = x}} w = \nabla_{X \setminus \{x\}} \sigma_y. \tag{41}$$

*Proof of (41).* Let $y \in X$. Then, $\sigma_y \in S_{n,X}$ is a permutation that sends $y$ to $x$ (by its definition); thus, $\sigma_y(y) = x$. Moreover, $S_{n,X}$ is a group (by Proposition 3.7.4 **(a)**). Thus, $w\sigma_y \in S_{n,X}$ for each $w \in S_{n,X}$ (since $\sigma_y \in S_{n,X}$). Hence, the map

$$\begin{aligned} S_{n,X} &\to S_{n,X}, \\ w &\mapsto w\sigma_y \end{aligned}$$

is well-defined. This map is furthermore bijective (with inverse given by $w \mapsto w\sigma_y^{-1}$). Thus, we can substitute $w\sigma_y$ for $w$ in the sum $\sum_{\substack{w \in S_{n,X}; \\ w(y) = x}} w$. We thus find

$$\sum_{\substack{w \in S_{n,X}; \\ w(y) = x}} w = \sum_{\substack{w \in S_{n,X}; \\ \left(w\sigma_y\right)(y) = x}} w\sigma_y = \sum_{\substack{w \in S_{n,X}; \\ w(x) = x}} w\sigma_y \tag{42}$$

(since each $w \in S_{n,X}$ satisfies $\left(w\sigma_y\right)(y) = w\left(\underbrace{\sigma_y(y)}_{=x}\right) = w(x)$). But we have

$$\{w \in S_{n,X} \mid w(x) = x\} = S_{n,X \setminus \{x\}} \qquad \text{(by (37))}.$$

Thus, the summation sign $\sum_{\substack{w \in S_{n,X}; \\ w(x) = x}}$ on the right hand side of (42) can be simplified

[8] *Proof.* Let $w \in S_{n,X}$. We must prove that there is a unique $y \in X$ such that $w(y) = x$.

From $w \in S_{n,X}$, we see that $w(i) = i$ for all $i \notin X$. Thus, if we had $w^{-1}(x) \notin X$, then we would have $w\left(w^{-1}(x)\right) = w^{-1}(x)$ (by the equality $w(i) = i$, applied to $i = w^{-1}(x)$), so that $w^{-1}(x) = w\left(w^{-1}(x)\right) = x \in X$, which would contradict $w^{-1}(x) \notin X$. Thus, we cannot have $w^{-1}(x) \notin X$. Hence, we must have $w^{-1}(x) \in X$. Thus, there exists a $y \in X$ such that $w(y) = x$ (namely, $y = w^{-1}(x)$). Moreover, this $y$ is clearly unique (since $w$ is a permutation and thus injective). Hence, we have shown that there is a unique $y \in X$ such that $w(y) = x$.

to $\sum\limits_{w \in S_{n,X\setminus\{x\}}}$. Hence, (42) rewrites as

$$\sum_{\substack{w \in S_{n,X}; \\ w(y)=x}} w = \sum_{w \in S_{n,X\setminus\{x\}}} w\sigma_y = \underbrace{\left( \sum_{w \in S_{n,X\setminus\{x\}}} w \right)}_{\substack{=\nabla_{X\setminus\{x\}} \\ \text{(by (39))}}} \sigma_y = \nabla_{X\setminus\{x\}} \sigma_y.$$

This proves (41). □

Now, (38) becomes

$$\begin{aligned} \nabla_X &= \sum_{w \in S_{n,X}} w = \sum_{y \in X} \underbrace{\sum_{\substack{w \in S_{n,X}; \\ w(y)=x}} w}_{\substack{=\nabla_{X\setminus\{x\}}\sigma_y \\ \text{(by (41))}}} \qquad \text{(by (40))} \\ &= \sum_{y \in X} \nabla_{X\setminus\{x\}} \sigma_y = \nabla_{X\setminus\{x\}} \sum_{y \in X} \sigma_y. \end{aligned}$$

This proves Lemma 3.7.10. □

*Proof of Lemma 3.7.9.* Let $S_{n,X}$ be as in Proposition 3.7.4. Set $t_{x,x} := \mathrm{id} \in S_n$. Thus, a permutation $t_{y,x}$ is defined for all $y \in X$ (including $x$). It is clear that this permutation $t_{y,x}$ belongs to $S_{n,X}$ (since both $x$ and $y$ belong to $X$) and sends $y$ to $x$. Thus, Lemma 3.7.10 (applied to $\sigma_y = t_{y,x}$) yields

$$\begin{aligned} \nabla_X &= \nabla_{X\setminus\{x\}} \underbrace{\sum_{y \in X} t_{y,x}}_{\substack{= t_{x,x} + \sum\limits_{y \in X\setminus\{x\}} t_{y,x} \\ \text{(here, we have split off the addend for } y=x\text{)}}} \\ &= \nabla_{X\setminus\{x\}} \left( \underbrace{t_{x,x}}_{=\mathrm{id}=1} + \sum_{y \in X\setminus\{x\}} t_{y,x} \right) = \nabla_{X\setminus\{x\}} \left( 1 + \sum_{y \in X\setminus\{x\}} t_{y,x} \right). \end{aligned}$$

This proves Lemma 3.7.9. □

*Proof of Lemma 3.7.8.* Lemma 3.7.9 (applied to $X = [k]$ and $x = k$) yields

$$\begin{aligned}
\nabla_{[k]} &= \nabla_{[k]\setminus\{k\}} \left(1 + \sum_{y\in[k]\setminus\{k\}} t_{y,k}\right) \\
&= \nabla_{[k-1]} \left(1 + \underbrace{\sum_{y\in[k-1]} t_{y,k}}_{=t_{1,k}+t_{2,k}+\cdots+t_{k-1,k}}\right) \qquad (\text{since } [k]\setminus\{k\} = [k-1]) \\
&= \nabla_{[k-1]} \left(1 + \underbrace{t_{1,k}+t_{2,k}+\cdots+t_{k-1,k}}_{\substack{=\mathbf{m}_k \\ \text{(by the definition of } \mathbf{m}_k\text{)}}}\right) = \nabla_{[k-1]} \left(1+\mathbf{m}_k\right).
\end{aligned}$$

This proves Lemma 3.7.8. □

### 3.7.3. The transposition cancellation theorem

The following theorem generalizes (15) as well as the slightly more general formula $\left(1+t_{i,j}\right)\left(1-t_{i,j}\right) = 0$ for all $i \neq j$:

**Theorem 3.7.11** (transposition cancellation theorem)**.** Let $X$ and $Y$ be two subsets of $[n]$ such that $|X \cap Y| > 1$. Then,

$$\begin{aligned}
\nabla_X \nabla_Y^- &= 0 \qquad \text{and} \\
\nabla_X^- \nabla_Y &= 0.
\end{aligned}$$

For example, $\nabla_{\{1,2,3\}} \nabla^-_{\{1,2,4\}} = 0$ and $\nabla^-_{\{1,2,3\}} \nabla_{\{1,2,4\}} = 0$.

*Proof of Theorem 3.7.11.* There exist two distinct elements $i$ and $j$ of $X \cap Y$ (since $|X \cap Y| > 1$). Pick such $i$ and $j$. Then, $i$ and $j$ are two distinct elements of $X$ (since $i \in X \cap Y \subseteq X$ and $j \in X \cap Y \subseteq X$). Hence, (31) yields

$$\nabla_X = \nabla_X^{\text{even}} \cdot \left(1+t_{i,j}\right) = \left(1+t_{i,j}\right) \cdot \nabla_X^{\text{even}},$$

where $\nabla_X^{\text{even}}$ is defined in Proposition 3.7.4 **(d)**. However, $i$ and $j$ are also two distinct elements of $Y$ (since $i \in X \cap Y \subseteq Y$ and $j \in X \cap Y \subseteq Y$). Hence, (32) yields

$$\nabla_Y^- = \nabla_Y^{\text{even}} \cdot \left(1-t_{i,j}\right) = \left(1-t_{i,j}\right) \cdot \nabla_Y^{\text{even}},$$

where $\nabla_Y^{\text{even}}$ is defined similarly to the $\nabla_X^{\text{even}}$ in Proposition 3.7.4 **(d)**. Thus,

$$\underbrace{\nabla_X}_{=\nabla_X^{\text{even}}\cdot\left(1+t_{i,j}\right)} \underbrace{\nabla_Y^-}_{=\left(1-t_{i,j}\right)\cdot\nabla_Y^{\text{even}}} = \nabla_X^{\text{even}} \cdot \underbrace{\left(1+t_{i,j}\right)\cdot\left(1-t_{i,j}\right)}_{\substack{=1-t_{i,j}^2=0 \\ \text{(since } t_{i,j}^2=\text{id}=1\text{)}}} \cdot \nabla_Y^{\text{even}} = 0.$$

Thus, $\nabla_X \nabla_Y^- = 0$ is proved. Similarly, we can show $\nabla_X^- \nabla_Y = 0$ (using $\nabla_X^- = \nabla_X^{\text{even}} \cdot \left(1 + t_{i,j}\right)$ and $\nabla_Y = \left(1 + t_{i,j}\right) \cdot \nabla_Y^{\text{even}}$). The proof of Theorem 3.7.11 is thus complete. □

### 3.7.4. Disjoint commutativity

The following easy fact says that if $X$ and $Y$ are two disjoint subsets of $[n]$, then the $X$-(sign)-integral commutes with the $Y$-(sign-)integral:

**Proposition 3.7.12.** Let $X$ and $Y$ be two disjoint subsets of $[n]$. Then:

**(a)** If $u \in S_n$ is a permutation that satisfies $(u(i) = i$ for all $i \notin X)$, and if $v \in S_n$ is a permutation that satisfies $(v(i) = i$ for all $i \notin Y)$, then $uv = vu$.

**(b)** We have

$$\nabla_X \nabla_Y = \nabla_Y \nabla_X; \qquad \nabla_X \nabla_Y^- = \nabla_Y^- \nabla_X;$$
$$\nabla_X^- \nabla_Y = \nabla_Y \nabla_X^-; \qquad \nabla_X^- \nabla_Y^- = \nabla_Y^- \nabla_X^-.$$

**(c)** If $v \in S_n$ is a permutation that satisfies $(v(i) = i$ for all $i \notin Y)$, then

$$\nabla_X v = v \nabla_X \qquad \text{and} \qquad \nabla_X^- v = v \nabla_X^-.$$

**(d)** If $p$ and $q$ are two distinct elements of $Y$, then the transposition $t_{p,q} \in S_n$ satisfies

$$\nabla_X t_{p,q} = t_{p,q} \nabla_X \qquad \text{and} \qquad \nabla_X^- t_{p,q} = t_{p,q} \nabla_X^-.$$

*Proof.* **(a)** Let $u \in S_n$ be a permutation that satisfies $(u(i) = i$ for all $i \notin X)$. Let $v \in S_n$ be a permutation that satisfies $(v(i) = i$ for all $i \notin Y)$. We must prove that $uv = vu$. In other words, we must prove that $(uv)(p) = (vu)(p)$ for each $p \in [n]$.

So let $p \in [n]$ be arbitrary. We are in one of the following three cases:

*Case 1:* We have $p \in X$.

*Case 2:* We have $p \in Y$.

*Case 3:* We have neither $p \in X$ nor $p \in Y$.

Let us first consider Case 1. In this case, we have $p \in X$. Hence, $p \notin Y$ (since the sets $X$ and $Y$ are disjoint). Thus, $v(p) = p$ (since $v(i) = i$ for all $i \notin Y$).

If we had $u(p) \notin X$, then we would have $u(u(p)) = u(p)$ (since $u(i) = i$ for all $i \notin X$), which would yield $u(p) = p$ (since the map $u$ is a permutation and thus injective), which would in turn yield $p = u(p) \notin X$, which would contradict $p \in X$. Hence, we cannot have $u(p) \notin X$. Thus, we have $u(p) \in X$. Therefore, $u(p) \notin Y$ (since the sets $X$ and $Y$ are disjoint). Thus, $v(u(p)) = u(p)$ (since $v(i) = i$ for all $i \notin Y$).

Now, $(uv)(p) = u(v(p)) = u(p)$ (since $v(p) = p$). Comparing this with $(vu)(p) = v(u(p)) = u(p)$, we obtain $(uv)(p) = (vu)(p)$. Hence, we have proved $(uv)(p) = (vu)(p)$ in Case 1.

Let us now consider Case 2. In this case, we have $p \in Y$. Hence, the same argument as in Case 1 (but now with the roles of $X$ and $Y$ swapped, and with the roles of $u$ and $v$

swapped) shows that $(vu)(p) = (uv)(p)$. In other words, $(uv)(p) = (vu)(p)$. Hence, we have proved $(uv)(p) = (vu)(p)$ in Case 2.

Let us finally consider Case 3. In this case, we have neither $p \in X$ nor $p \in Y$. Hence, $p \notin X$ and $p \notin Y$. From $p \notin X$, we obtain $u(p) = p$ (since $u(i) = i$ for all $i \notin X$). Thus, $v(u(p)) = v(p)$. From $p \notin Y$, we obtain $v(p) = p$ (since $v(i) = i$ for all $i \notin Y$). Hence, $u(v(p)) = u(p)$. Now, comparing $(uv)(p) = u(v(p)) = u(p) = p$ with $(vu)(p) = v(u(p)) = v(p) = p$, we obtain $(uv)(p) = (vu)(p)$. Hence, we have proved $(uv)(p) = (vu)(p)$ in Case 3.

Now we have proved $(uv)(p) = (vu)(p)$ in all three Cases 1, 2 and 3. Hence, $(uv)(p) = (vu)(p)$ always holds.

Forget that we fixed $p$. We thus have shown that $(uv)(p) = (vu)(p)$ for all $p \in [n]$. In other words, $uv = vu$. This proves Proposition 3.7.12 **(a)**.

**(b)** The definition of $\nabla_X$ yields

$$\nabla_X = \sum_{\substack{w \in S_n; \\ w(i)=i \text{ for all } i \notin X}} w = \sum_{\substack{u \in S_n; \\ u(i)=i \text{ for all } i \notin X}} u. \tag{43}$$

Similarly,

$$\nabla_Y = \sum_{\substack{v \in S_n; \\ v(i)=i \text{ for all } i \notin Y}} v. \tag{44}$$

Multiplying these two equalities, we obtain

$$\begin{aligned}
\nabla_X \nabla_Y &= \left( \sum_{\substack{u \in S_n; \\ u(i)=i \text{ for all } i \notin X}} u \right) \left( \sum_{\substack{v \in S_n; \\ v(i)=i \text{ for all } i \notin Y}} v \right) \\
&= \sum_{\substack{u \in S_n; \\ u(i)=i \text{ for all } i \notin X}} \sum_{\substack{v \in S_n; \\ v(i)=i \text{ for all } i \notin Y}} \underbrace{uv}_{=vu} \quad \text{(by Proposition 3.7.12 \textbf{(a)})} \\
&= \sum_{\substack{u \in S_n; \\ u(i)=i \text{ for all } i \notin X}} \sum_{\substack{v \in S_n; \\ v(i)=i \text{ for all } i \notin Y}} vu \\
&= \underbrace{\left( \sum_{\substack{v \in S_n; \\ v(i)=i \text{ for all } i \notin Y}} v \right)}_{=\nabla_Y \text{ (by (44))}} \underbrace{\left( \sum_{\substack{u \in S_n; \\ u(i)=i \text{ for all } i \notin X}} u \right)}_{=\nabla_X \text{ (by (43))}} = \nabla_Y \nabla_X.
\end{aligned}$$

Analogous computations (with some signs injected in) prove the equalities $\nabla_X \nabla_Y^- = \nabla_Y^- \nabla_X$ and $\nabla_X^- \nabla_Y = \nabla_Y \nabla_X^-$ and $\nabla_X^- \nabla_Y^- = \nabla_Y^- \nabla_X^-$. This completes the proof of Proposition 3.7.12 **(b)**.

**(c)** Let $v \in S_n$ be a permutation that satisfies $(v(i) = i \text{ for all } i \notin Y)$. Then, from (43),

we obtain

$$\nabla_X v = \left( \sum_{\substack{u \in S_n; \\ u(i)=i \text{ for all } i \notin X}} u \right) v = \sum_{\substack{u \in S_n; \\ u(i)=i \text{ for all } i \notin X}} \underbrace{uv}_{=vu} \quad \text{(by Proposition 3.7.12 \textbf{(a)})}$$

$$= \sum_{\substack{u \in S_n; \\ u(i)=i \text{ for all } i \notin X}} vu = v \underbrace{\left( \sum_{\substack{u \in S_n; \\ u(i)=i \text{ for all } i \notin X}} u \right)}_{=\nabla_X \text{ (by (43))}} = v \nabla_X.$$

An analogous computation (with some signs injected in) proves the equality $\nabla_X^- v = v \nabla_X^-$. Thus, Proposition 3.7.12 **(c)** is proved.

**(d)** Let $p$ and $q$ be two distinct elements of $Y$. Then, $\{p, q\} \subseteq Y$. Hence, each $i \notin Y$ satisfies $i \notin \{p, q\}$.

But the transposition $t_{p,q}$ satisfies $t_{p,q}(i) = i$ for all $i \notin \{p, q\}$ (by its definition). Therefore, it satisfies $t_{p,q}(i) = i$ for all $i \notin Y$ (since each $i \notin Y$ satisfies $i \notin \{p, q\}$). Hence, Proposition 3.7.12 **(c)** (applied to $v = t_{p,q}$) yields

$$\nabla_X t_{p,q} = t_{p,q} \nabla_X \qquad \text{and} \qquad \nabla_X^- t_{p,q} = t_{p,q} \nabla_X^-.$$

This proves Proposition 3.7.12 **(d)**. □

### 3.7.5. Products of $X$-integrals

The next proposition will prove itself useful in the study of Specht modules.

**Proposition 3.7.13.** Let $X_1, X_2, \ldots, X_k$ be $k$ disjoint subsets of $[n]$ such that $X_1 \cup X_2 \cup \cdots \cup X_k = [n]$. Then:

**(a)** If $w \in S_n$ is any permutation, then the four statements

$$(w(X_i) \subseteq X_i \text{ for all } i \in [k]), \tag{45}$$

$$(w(X_i) = X_i \text{ for all } i \in [k]), \tag{46}$$

$$(w(X_i) \subseteq X_i \text{ for all } i \in [k-1]) \qquad \text{and} \tag{47}$$

$$(w(X_i) = X_i \text{ for all } i \in [k-1]) \tag{48}$$

are equivalent.

**(b)** We have

$$\sum_{\substack{w \in S_n; \\ w(X_i) \subseteq X_i \text{ for all } i \in [k]}} w = \nabla_{X_1} \nabla_{X_2} \cdots \nabla_{X_k} \tag{49}$$

and

$$\sum_{\substack{w \in S_n; \\ w(X_i) \subseteq X_i \text{ for all } i \in [k]}} (-1)^w w = \nabla_{X_1}^- \nabla_{X_2}^- \cdots \nabla_{X_k}^-. \tag{50}$$

**(c)** The factors $\nabla_{X_1}, \nabla_{X_2}, \ldots, \nabla_{X_k}$ on the right hand side of (49) pairwise commute. So do the factors $\nabla^-_{X_1}, \nabla^-_{X_2}, \ldots, \nabla^-_{X_k}$ on the right hand side of (50).

**Example 3.7.14.** Let $n = 8$ and $k = 4$ and $X_1 = \{2,5,6\}$ and $X_2 = \{1,3\}$ and $X_3 = \{4,8\}$ and $X_4 = \{7\}$ (so that the four sets $X_1, X_2, X_3, X_4$ are disjoint and their union is $[n]$). Then, (49) says that

$$\sum_{\substack{w \in S_8; \\ w(\{2,5,6\}) \subseteq \{2,5,6\}; \\ w(\{1,3\}) \subseteq \{1,3\}; \\ w(\{4,8\}) \subseteq \{4,8\}; \\ w(\{7\}) \subseteq \{7\}}} w = \nabla_{\{2,5,6\}} \nabla_{\{1,3\}} \nabla_{\{4,8\}} \nabla_{\{7\}}.$$

Note that the factor $\nabla_{\{7\}}$ equals 1 (by Example 3.7.2 **(b)**) and thus can be omitted.

*Proof of Proposition 3.7.13.* **(a)** Let $w \in S_n$ be any permutation. We must prove that the four statements (45), (46), (47) and (48) are equivalent. Clearly, (46) implies (45). Let us now show the converse implication:

*Proof that (45) implies (46):* Assume that (45) holds. We must prove that (46) holds.

Let $i \in [k]$. Then, $w(X_i) \subseteq X_i$ (since we assumed that (45) holds). But $w$ is a permutation, and thus is injective. Hence, $|w(P)| = |P|$ for any subset $P$ of $[n]$. In particular, this shows that $|w(X_i)| = |X_i|$.

But $X_i$ is a finite set. Hence, the only subset of $X_i$ that has the same size as $X_i$ is $X_i$ itself. In other words, if $P$ is a subset of $X_i$ such that $|P| = |X_i|$, then $P = X_i$. Applying this to $P = w(X_i)$, we obtain $w(X_i) = X_i$ (since $w(X_i) \subseteq X_i$ and $|w(X_i)| = |X_i|$).

Forget that we fixed $i$. We thus have shown that $w(X_i) = X_i$ for all $i \in [k]$. In other words, (46) holds. This shows that (45) implies (46). □

We have now shown that (45) implies (46), whereas we know that (46) implies (45). Hence, the statements (45) and (46) are equivalent. The same argument (with $k-1$ instead of $k$) shows that the statements (47) and (48) are equivalent.

It remains to prove that the statements (45) and (47) are equivalent. Clearly, (45) implies (47), so we only need to check that (47) implies (45).

*Proof that (47) implies (45):* Assume that (47) holds. We must prove that (45) holds. In other words, we must prove that $w(X_i) \subseteq X_i$ for all $i \in [k]$.

First let us prove that $w(X_k) \subseteq X_k$. Indeed, let $x \in X_k$ be arbitrary. We shall show that $w(x) \in X_k$. Indeed, assume the contrary. Then, $w(x) \notin X_k$. But $w(x) \in [n] = X_1 \cup X_2 \cup \cdots \cup X_k$, so that $w(x) \in X_i$ for some $i \in [k]$. Consider this $i$. If we had $i = k$, then we would obtain $w(x) \in X_k$ (since $w(x) \notin X_i$), which would contradict $w(x) \notin X_k$. Thus, we cannot have $i = k$. Hence, $i \neq k$, so that $i \in [k] \setminus \{k\} = [k-1]$.

But we have assumed that (47) holds. Thus, (48) also holds (since the statements (47) and (48) are equivalent). Therefore, $w(X_i) = X_i$ (since $i \in [k-1]$). Hence, $w(x) \in X_i = w(X_i)$, so that $w(x) = w(y)$ for some $y \in X_i$. Consider this $y$.

However, $w$ is a permutation and thus injective. Hence, from $w(x) = w(y)$, we obtain $x = y$, so that $x = y \in X_i$. Combining this with $x \in X_k$, we obtain $x \in X_i \cap X_k$.

But the sets $X_1, X_2, \ldots, X_k$ are disjoint. Hence, $X_i \cap X_k = \varnothing$ (since $i \neq k$). Hence, $x \in X_i \cap X_k = \varnothing$, which is absurd. This contradiction shows that our assumption was false. Thus, $w(x) \in X_k$ is proved.

Forget that we fixed $x$. We thus have shown that $w(x) \in X_k$ for each $x \in X_k$. In other words, $w(X_k) \subseteq X_k$.

Now, we know that the relation $w(X_i) \subseteq X_i$ holds for all $i \in [k-1]$ (by our assumption (47)), but we also know that the same relation holds for $i = k$ (since $w(X_k) \subseteq X_k$). Hence, this relation holds for all $i \in [k-1] \cup \{k\}$, that is, for all $i \in [k]$. In other words, (45) holds. This shows that (47) implies (45). □

We thus conclude that the statements (47) and (45) are equivalent (since (47) implies (45), whereas (45) implies (47)). Combined with the other equivalences we have shown above, this entails that all four statements (45), (46), (47) and (48) are equivalent. Thus, Proposition 3.7.13 **(a)** is proved.

**(b)** The following proof is long and yet free of any substance. The underlying idea is that a permutation $w \in S_n$ that satisfies $w(X_i) \subseteq X_i$ for all $i \in [k]$ is a permutation that permutes each of the $k$ subsets $X_1, X_2, \ldots, X_k$ individually, and thus can be split up into a permutation of $X_1$, a permutation of $X_2$, and so on. This splitting-up can be used to write $w$ as a product $w_1 w_2 \cdots w_k$ with $w_i \in S_{n,X_i}$ for each $i \in [k]$ (where $S_{n,X_i}$ is defined as in Proposition 3.7.4), and thus the sums on the left hand sides of (49) and (50) can be reindexed as sums over $k$-tuples $(w_1, w_2, \ldots, w_k)$, which then (by the product rule) become the right hand sides.

To make this proof rigorous, we need to formally define the "splitting-up" procedure and prove all its intuitively obvious properties. Thus we arrive at the following proof:

For any subset $X$ of $[n]$, define a subset $S_{n,X}$ of $S_n$ as in Proposition 3.7.4.

Define a subset $K$ of $S_n$ by

$$K := \{w \in S_n \mid w(X_i) \subseteq X_i \text{ for all } i \in [k]\}.$$

Clearly, this set $K$ is a subgroup of $S_n$ (because if two permutations $u, v \in S_n$ satisfy $u(X_i) \subseteq X_i$ and $v(X_i) \subseteq X_i$, then their product $uv$ satisfies $(uv)(X_i) = u\left(\underbrace{v(X_i)}_{\subseteq X_i}\right) \subseteq u(X_i) \subseteq X_i$).

We note the following simple fact first:

*Claim 1:* For any $j \in [k]$, the set $S_{n,X_j}$ is a subgroup of $K$.

*Proof of Claim 1.* Let $j \in [k]$. We must prove that the set $S_{n,X_j}$ is a subgroup of $K$. We already know (from Proposition 3.7.4 **(a)**, applied to $X = X_j$) that $S_{n,X_j}$ is a subgroup of $S_n$, and thus is closed under multiplication and taking inverses. Hence, it remains to prove that $S_{n,X_j}$ is a subset of $K$.

Let $u \in S_{n,X_j}$ be arbitrary. Thus,

$$u(i) = i \qquad \text{for all } i \notin X_j \tag{51}$$

(by the definition of $S_{n,X_j}$). Moreover, $u$ is a permutation of $[n]$, thus injective.

We shall now prove that $u\left(X_j\right) \subseteq X_j$.

Indeed, let $x \in X_j$. If we had $u(x) \notin X_j$, then we would have $u(u(x)) = u(x)$ (by (51), applied to $i = u(x)$), which would yield $u(x) = x$ (since $u$ is injective) and thus $x = u(x) \notin X_j$, which would contradict $x \in X_j$. Thus, we cannot have $u(x) \notin X_j$. Hence, $u(x) \in X_j$.

Forget that we fixed $x$. We thus have shown that $u(x) \in X_j$ for each $x \in X_j$. In other words, $u\left(X_j\right) \subseteq X_j$.

Next, we shall show that

$$u(X_i) \subseteq X_i \qquad \text{for each } i \in [k]. \tag{52}$$

Indeed, let $i \in [k]$ be arbitrary. We must prove that $u(X_i) \subseteq X_i$. If $i = j$, then this follows from $u\left(X_j\right) \subseteq X_j$. Thus, we WLOG assume that $i \neq j$. Hence, the sets $X_i$ and $X_j$ are disjoint (since the sets $X_1, X_2, \ldots, X_k$ are disjoint). Now, let $x \in X_i$. Therefore, $x \notin X_j$ (since the sets $X_i$ and $X_j$ are disjoint, so that an element of $X_i$ cannot belong to $X_j$). Hence, (51) (applied to $x$ instead of $i$) yields $u(x) = x \in X_i$.

Forget that we fixed $x$. We thus have shown that $u(x) \in X_i$ for each $x \in X_i$. In other words, $u(X_i) \subseteq X_i$.

Forget that we fixed $i$. Thus we have proved that $u(X_i) \subseteq X_i$ for each $i \in [k]$. In other words, $u \in K$ (by the definition of $K$).

Forget that we fixed $u$. We thus have shown that $u \in K$ for each $u \in S_{n,X_j}$. In other words, $S_{n,X_j} \subseteq K$. Hence, $S_{n,X_j}$ is a subgroup of $K$ (since $S_{n,X_j}$ is a subgroup of $S_n$). This proves Claim 1. $\square$

Now, each $k$-tuple $(w_1, w_2, \ldots, w_k) \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}$ satisfies $w_j \in S_{n,X_j} \subseteq K$ for each $j \in [k]$ (by Claim 1) and therefore $w_1 w_2 \cdots w_k \in K$ (since $K$ is a subgroup of $S_n$). Thus, we can define a map

$$\begin{aligned} \Psi : S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k} &\to K, \\ (w_1, w_2, \ldots, w_k) &\mapsto w_1 w_2 \cdots w_k. \end{aligned}$$

Consider this map $\Psi$.

For each $j \in [k]$ and each $w \in K$, we define a map $w^{(j)} : [n] \to [n]$ by setting

$$w^{(j)}(x) = \begin{cases} w(x), & \text{if } x \in X_j; \\ x, & \text{if } x \notin X_j \end{cases} \qquad \text{for each } x \in [n].$$

We claim the following:

*Claim 2:* Let $j \in [k]$ and $w \in K$. Then, $w^{(j)} \in S_{n,X_j}$.

*Proof of Claim 2.* The permutation $w$ belongs to $K$, and thus satisfies $(w(X_i) \subseteq X_i \text{ for all } i \in [k])$ (by the definition of $K$). Hence, it satisfies $(w(X_i) = X_i \text{ for all } i \in [k])$ (since Proposition 3.7.13 **(a)** shows that the two statements $(w(X_i) \subseteq X_i \text{ for all } i \in [k])$ and $(w(X_i) = X_i \text{ for all } i \in [k])$ are equivalent). Applying this to $i = j$, we obtain $w\left(X_j\right) = X_j$.

Now, we shall prove that the map $w^{(j)}$ is surjective. In other words, we shall prove that $y \in w^{(j)}([n])$ for each $y \in [n]$.

Indeed, fix $y \in [n]$. We must prove that $y \in w^{(j)}([n])$.

If $y \notin X_j$, then $w^{(j)}(y) = y$ (since the definition of $w^{(j)}$ yields that $w^{(j)}(x) = x$ whenever $x \notin X_j$) and thus $y = w^{(j)}(y) \in w^{(j)}([n])$. Hence, we can WLOG assume that we don't have $y \notin X_j$. Thus, $y \in X_j$. In other words, $y \in w\left(X_j\right)$ (since $w\left(X_j\right) = X_j$). In other words, $y = w(x)$ for some $x \in X_j$. Consider this $x$. The definition of $w^{(j)}$ yields $w^{(j)}(x) = w(x)$ (since $x \in X_j$). Hence, $y = w(x) = w^{(j)}(x) \in w^{(j)}([n])$.

Forget that we fixed $y$. We thus have proved that $y \in w^{(j)}([n])$ for each $y \in [n]$. In other words, $[n] \subseteq w^{(j)}([n])$. Thus, the map $w^{(j)}$ is surjective.

But the pigeonhole principle says that any surjective map from a finite set to itself must be bijective. Hence, the map $w^{(j)}$ is bijective (since $w^{(j)}$ is a surjective map from the finite set $[n]$ to itself). Thus, $w^{(j)}$ is a permutation of $[n]$. In other words, $w^{(j)} \in S_n$. Moreover, $w^{(j)}(x) = x$ for all $x \notin X_j$ (by the definition of $w^{(j)}$). In other words, $w^{(j)}(i) = i$ for all $i \notin X_j$. This entails $w^{(j)} \in S_{n,X_j}$ (by the definition of $S_{n,X_j}$, since $w^{(j)} \in S_n$). This proves Claim 2. □

Claim 2 shows that each $w \in K$ satisfies $w^{(j)} \in S_{n,X_j}$ for all $j \in [k]$, and therefore $\left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right) \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}$. Hence, we can define a map

$$\begin{aligned} \Phi : K &\to S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}, \\ w &\mapsto \left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right). \end{aligned}$$

Consider this map $\Phi$. We shall now show that the maps $\Phi$ and $\Psi$ are mutually inverse. First, we shall show the following claim:

*Claim 3:* Let $w \in K$. Then, $w = w^{(1)} w^{(2)} \cdots w^{(k)}$.

*Proof of Claim 3.* Let $x \in [n]$. Then, $x \in [n] = X_1 \cup X_2 \cup \cdots \cup X_k$. Hence, $x \in X_p$ for some $p \in [k]$. Consider this $p$. By the definition of $w^{(p)}$, we have $w^{(p)}(x) = \begin{cases} w(x), & \text{if } x \in X_p; \\ x, & \text{if } x \notin X_p \end{cases} = w(x)$ (since $x \in X_p$). But $w \in K$, so that $(w(X_i) \subseteq X_i$ for all $i \in [k])$ (by the definition of $K$). Applying this to $i = p$, we find $w(X_p) \subseteq X_p$.

Set $y := w(x)$. Then, $y = w(x) \in w(X_p)$ (since $x \in X_p$), so that $y \in w(X_p) \subseteq X_p$.

The $k$ sets $X_1, X_2, \ldots, X_k$ are disjoint. Thus, for each $i \in \{p+1, p+2, \ldots, k\}$, we have $w^{(i)}(x) = x$ [9]. In other words, all $k-p$ permutations $w^{(p+1)}, w^{(p+2)}, \ldots, w^{(k)}$ leave the element $x$ unchanged. Hence, so does their product $w^{(p+1)} w^{(p+2)} \cdots w^{(k)}$. In other words,

$$\left(w^{(p+1)} w^{(p+2)} \cdots w^{(k)}\right)(x) = x.$$

Hence,

$$\underbrace{\left(w^{(p)} w^{(p+1)} \cdots w^{(k)}\right)}_{= w^{(p)} \circ \left(w^{(p+1)} w^{(p+2)} \cdots w^{(k)}\right)}(x) = w^{(p)}\left(\underbrace{\left(w^{(p+1)} w^{(p+2)} \cdots w^{(k)}\right)(x)}_{=x}\right)$$
$$= w^{(p)}(x) = w(x) = y.$$

The $k$ sets $X_1, X_2, \ldots, X_k$ are disjoint. Thus, for each $i \in \{1, 2, \ldots, p-1\}$, we have $w^{(i)}(y) = y$ [10]. In other words, all $p-1$ permutations $w^{(1)}, w^{(2)}, \ldots, w^{(p-1)}$ leave the element $y$ unchanged. Hence, so does their product $w^{(1)} w^{(2)} \cdots w^{(p-1)}$. In other words,

$$\left(w^{(1)} w^{(2)} \cdots w^{(p-1)}\right)(y) = y.$$

[9] *Proof.* Let $i \in \{p+1, p+2, \ldots, k\}$. Thus, $i \neq p$. Hence, the sets $X_i$ and $X_p$ are disjoint (since $X_1, X_2, \ldots, X_k$ are disjoint). Thus, an element of $X_p$ cannot belong to $X_i$. From $x \in X_p$, we thus obtain $x \notin X_i$. Hence, $w^{(i)}(x) = x$ (since the definition of $w^{(i)}$ yields that $w^{(i)}(z) = z$ for each $z \notin X_i$).

[10] *Proof.* Let $i \in \{1, 2, \ldots, p-1\}$. Thus, $i \neq p$. Hence, the sets $X_i$ and $X_p$ are disjoint (since $X_1, X_2, \ldots, X_k$ are disjoint). Thus, an element of $X_p$ cannot belong to $X_i$. From $y \in X_p$, we thus obtain $y \notin X_i$. Hence, $w^{(i)}(y) = y$ (since the definition of $w^{(i)}$ yields that $w^{(i)}(z) = z$ for each $z \notin X_i$).

Now,

$$\begin{aligned}
&\underbrace{\left(w^{(1)}w^{(2)}\cdots w^{(k)}\right)}_{=\left(w^{(1)}w^{(2)}\cdots w^{(p-1)}\right)\circ\left(w^{(p)}w^{(p+1)}\cdots w^{(k)}\right)}(x) \\
&= \left(w^{(1)}w^{(2)}\cdots w^{(p-1)}\right)\left(\underbrace{\left(w^{(p)}w^{(p+1)}\cdots w^{(k)}\right)(x)}_{=y}\right) \\
&= \left(w^{(1)}w^{(2)}\cdots w^{(p-1)}\right)(y) = y = w(x).
\end{aligned}$$

Forget that we fixed $x$. We thus have shown that $\left(w^{(1)}w^{(2)}\cdots w^{(k)}\right)(x) = w(x)$ for each $x \in [n]$. In other words, $w^{(1)}w^{(2)}\cdots w^{(k)} = w$. This proves Claim 3. □

Claim 3 easily yields the following:

*Claim 4:* We have $\Psi \circ \Phi = \mathrm{id}$.

*Proof of Claim 4.* Let $w \in K$. Then, the definition of $\Phi$ yields

$$\Phi(w) = \left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right).$$

Applying the map $\Psi$ to this equality, we obtain

$$\begin{aligned}
\Psi(\Phi(w)) &= \Psi\left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right) \\
&= w^{(1)}w^{(2)}\cdots w^{(k)} \qquad \text{(by the definition of } \Psi\text{)} \\
&= w \qquad \text{(by Claim 3)}.
\end{aligned}$$

Hence, $(\Psi \circ \Phi)(w) = \Psi(\Phi(w)) = w = \mathrm{id}(w)$.

Forget that we fixed $w$. We thus have shown that $(\Psi \circ \Phi)(w) = \mathrm{id}(w)$ for each $w \in K$. In other words, $\Psi \circ \Phi = \mathrm{id}$. This proves Claim 4. □

On the other hand, the following holds:

*Claim 5:* Let $(w_1, w_2, \ldots, w_k) \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}$. Set $w = w_1 w_2 \cdots w_k$. Then, $w^{(j)} = w_j$ for each $j \in [k]$.

*Proof of Claim 5.* We have $(w_1, w_2, \ldots, w_k) \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}$. Thus, for each $j \in [k]$, we have

$$\begin{aligned}
w_j &\in S_{n,X_j} \\
&= \left\{w \in S_n \mid w(i) = i \text{ for all } i \notin X_j\right\}
\end{aligned} \tag{53}$$

(by the definition of $S_{n,X_j}$) and therefore

$$w_j(i) = i \qquad \text{for all } i \notin X_j. \tag{54}$$

Let $j \in [k]$. Let $x \in [n]$. We shall show that $w^{(j)}(x) = w_j(x)$.

We are in one of the following two cases:

*Case 1:* We have $x \in X_j$.

*Case 2:* We have $x \notin X_j$.

Let us first consider Case 1. In this case, we have $x \in X_j$. Hence, $w^{(j)}(x) = w(x)$ (by the definition of $w^{(j)}$).

Set $y := w_j(x)$. Note that $S_{n,X_j} \subseteq K$ (by Claim 1), so that (53) becomes $w_j \in S_{n,X_j} \subseteq K$. Hence, we have $w_j(X_i) \subseteq X_i$ for all $i \in [k]$ (by the definition of $K$). Applying this to $i = j$, we find $w_j(X_j) \subseteq X_j$. Hence, from $x \in X_j$, we obtain $w_j(x) \in w_j(X_j) \subseteq X_j$. In other words, $y \in X_j$ (since $y = w_j(x)$).

The $k$ sets $X_1, X_2, \ldots, X_k$ are disjoint. Thus, for each $i \in \{j+1, j+2, \ldots, k\}$, we have $w_i(x) = x$ [11]. In other words, all $k-j$ permutations $w_{j+1}, w_{j+2}, \ldots, w_k$ leave the element $x$ unchanged. Hence, so does their product $w_{j+1}w_{j+2}\cdots w_k$. In other words,

$$\left(w_{j+1}w_{j+2}\cdots w_k\right)(x) = x.$$

Hence,

$$\underbrace{\left(w_j w_{j+1}\cdots w_k\right)}_{=w_j \circ \left(w_{j+1}w_{j+2}\cdots w_k\right)}(x) = w_j\left(\underbrace{\left(w_{j+1}w_{j+2}\cdots w_k\right)(x)}_{=x}\right) = w_j(x) = w(x) = y.$$

The $k$ sets $X_1, X_2, \ldots, X_k$ are disjoint. Thus, for each $i \in \{1, 2, \ldots, j-1\}$, we have $w_i(y) = y$ [12]. In other words, all $j-1$ permutations $w_1, w_2, \ldots, w_{j-1}$ leave the element $y$ unchanged. Hence, so does their product $w_1 w_2 \cdots w_{j-1}$. In other words,

$$\left(w_1 w_2 \cdots w_{j-1}\right)(y) = y.$$

Now,

$$\begin{aligned} w^{(j)}(x) &= \underbrace{w}_{\substack{=w_1w_2\cdots w_k \\ =\left(w_1w_2\cdots w_{j-1}\right)\circ\left(w_jw_{j+1}\cdots w_k\right)}}(x) = \left(w_1w_2\cdots w_{j-1}\right)\left(\underbrace{\left(w_jw_{j+1}\cdots w_k\right)(x)}_{=y}\right) \\ &= \left(w_1w_2\cdots w_{j-1}\right)(y) = y = w_j(x). \end{aligned}$$

[11] *Proof.* Let $i \in \{j+1, j+2, \ldots, k\}$. Thus, $i \neq j$. Hence, the sets $X_i$ and $X_j$ are disjoint (since $X_1, X_2, \ldots, X_k$ are disjoint). Thus, an element of $X_j$ cannot belong to $X_i$. From $x \in X_j$, we thus obtain $x \notin X_i$. Hence, (54) (applied to $i$ and $x$ instead of $j$ and $i$) yields $w_i(x) = x$.

[12] *Proof.* Let $i \in \{1, 2, \ldots, j-1\}$. Thus, $i \neq j$. Hence, the sets $X_i$ and $X_j$ are disjoint (since $X_1, X_2, \ldots, X_k$ are disjoint). Thus, an element of $X_j$ cannot belong to $X_i$. From $y \in X_j$, we thus obtain $y \notin X_i$. Hence, (54) (applied to $i$ and $y$ instead of $j$ and $i$) yields $w_i(y) = y$.

Thus, we have proved $w^{(j)}(x) = w_j(x)$ in Case 1.

Let us now consider Case 2. In this case, we have $x \notin X_j$. Hence, $w^{(j)}(x) = x$ (by the definition of $w^{(j)}$) and $w_j(x) = x$ (by (54), applied to $i = x$). Comparing these two equalities, we obtain $w^{(j)}(x) = w_j(x)$. Thus, we have proved $w^{(j)}(x) = w_j(x)$ in Case 2.

We have now proved $w^{(j)}(x) = w_j(x)$ in both Cases 1 and 2. Hence, $w^{(j)}(x) = w_j(x)$ always holds.

Forget that we fixed $x$. We have now shown that $w^{(j)}(x) = w_j(x)$ for all $x \in [n]$. In other words, $w^{(j)} = w_j$. This proves Claim 5. □

Claim 5, in turn, easily yields the following:

*Claim 6:* We have $\Phi \circ \Psi = \mathrm{id}$.

*Proof of Claim 6.* Let $\mathbf{w} \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}$. Write this $k$-tuple $\mathbf{w}$ in the form $\mathbf{w} = (w_1, w_2, \ldots, w_k)$. Thus,

$$\Psi(\mathbf{w}) = \Psi(w_1, w_2, \ldots, w_k) = w_1 w_2 \cdots w_k \qquad \text{(by the definition of } \Psi\text{)}.$$

Set $w := w_1 w_2 \cdots w_k$. Thus, $\Psi(\mathbf{w}) = w_1 w_2 \cdots w_k = w$. Applying the map $\Phi$ to this equality, we obtain

$$\Phi(\Psi(\mathbf{w})) = \Phi(w) = \left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right) \qquad \text{(by the definition of } \Phi\text{)}.$$

However, Claim 5 yields that $w^{(j)} = w_j$ for each $j \in [k]$. In other words,

$$\left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right) = (w_1, w_2, \ldots, w_k).$$

Hence,

$$\begin{aligned}(\Phi \circ \Psi)(\mathbf{w}) &= \Phi(\Psi(\mathbf{w})) = \left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right) \\ &= (w_1, w_2, \ldots, w_k) = \mathbf{w} = \mathrm{id}(\mathbf{w}).\end{aligned}$$

Since we have proved this for each $\mathbf{w} \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}$, we thus conclude that $\Phi \circ \Psi = \mathrm{id}$. Thus, Claim 6 is proved. □

Combining Claim 4 with Claim 6, we see that the maps $\Phi$ and $\Psi$ are mutually inverse. Hence, the maps $\Phi$ and $\Psi$ are invertible, i.e., are bijections. Now, recall that $K$ was defined as the set of all permutations $w \in S_n$ that satisfy $(w(X_i) \subseteq X_i$ for all $i \in [k])$. Hence, the summation sign $\sum_{\substack{w \in S_n; \\ w(X_i) \subseteq X_i \text{ for all } i \in [k]}}$ can

be rewritten as $\sum_{w \in K}$. Thus,

$$\begin{aligned}
&\sum_{\substack{w \in S_n; \\ w(X_i) \subseteq X_i \text{ for all } i \in [k]}} (-1)^w w \\
&= \sum_{w \in K} (-1)^w w \\
&= \sum_{(w_1, w_2, \ldots, w_k) \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}} \underbrace{(-1)^{\Psi(w_1, w_2, \ldots, w_k)} \Psi(w_1, w_2, \ldots, w_k)}_{\substack{= (-1)^{w_1 w_2 \cdots w_k} w_1 w_2 \cdots w_k \\ \text{(since } \Psi(w_1, w_2, \ldots, w_k) = w_1 w_2 \cdots w_k \\ \text{(by the definition of } \Psi))}} \\
&\qquad \left( \begin{array}{c} \text{here, we have substituted } \Psi(w_1, w_2, \ldots, w_k) \\ \text{for } w \text{ in the sum, since the} \\ \text{map } \Psi : S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k} \to K \\ \text{is a bijection} \end{array} \right) \\
&= \sum_{(w_1, w_2, \ldots, w_k) \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}} \underbrace{(-1)^{w_1 w_2 \cdots w_k}}_{\substack{= (-1)^{w_1} (-1)^{w_2} \cdots (-1)^{w_k} \\ \text{(by the multiplicativity of the sign)}}} w_1 w_2 \cdots w_k \\
&= \sum_{(w_1, w_2, \ldots, w_k) \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}} \underbrace{(-1)^{w_1} (-1)^{w_2} \cdots (-1)^{w_k} w_1 w_2 \cdots w_k}_{= \left((-1)^{w_1} w_1\right) \left((-1)^{w_2} w_2\right) \cdots \left((-1)^{w_k} w_k\right)} \\
&= \sum_{(w_1, w_2, \ldots, w_k) \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}} \left((-1)^{w_1} w_1\right) \left((-1)^{w_2} w_2\right) \cdots \left((-1)^{w_k} w_k\right).
\end{aligned} \tag{55}$$

However, each $i \in [k]$ satisfies

$$\nabla^-_{X_i} = \sum_{w \in S_{n,X_i}} (-1)^w w \tag{56}$$

(by the second equality in Proposition 3.7.4 **(b)**, applied to $X = X_i$). Multiplying the equalities (55) over all $i \in \{1, 2, \ldots, k\}$, we obtain

$$\begin{aligned}
&\nabla^-_{X_1} \nabla^-_{X_2} \cdots \nabla^-_{X_k} \\
&= \left( \sum_{w \in S_{n,X_1}} (-1)^w w \right) \left( \sum_{w \in S_{n,X_2}} (-1)^w w \right) \cdots \left( \sum_{w \in S_{n,X_k}} (-1)^w w \right) \\
&= \sum_{(w_1, w_2, \ldots, w_k) \in S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}} \left((-1)^{w_1} w_1\right) \left((-1)^{w_2} w_2\right) \cdots \left((-1)^{w_k} w_k\right)
\end{aligned}$$

(by the product rule, i.e., by expanding the product). Comparing this with (55), we obtain

$$\sum_{\substack{w \in S_n; \\ w(X_i) \subseteq X_i \text{ for all } i \in [k]}} (-1)^w w = \nabla^-_{X_1} \nabla^-_{X_2} \cdots \nabla^-_{X_k}.$$

An analogous argument (but with all the signs removed) shows that

$$\sum_{\substack{w \in S_n; \\ w(X_i) \subseteq X_i \text{ for all } i \in [k]}} w = \nabla_{X_1} \nabla_{X_2} \cdots \nabla_{X_k}.$$

Thus, Proposition 3.7.13 **(b)** is proved.

**(c)** If $i$ and $j$ are two distinct elements of $[k]$, then the sets $X_i$ and $X_j$ are disjoint (since all $k$ sets $X_1, X_2, \ldots, X_k$ are disjoint) and thus satisfy $\nabla_{X_i} \nabla_{X_j} = \nabla_{X_j} \nabla_{X_i}$ (by the equality $\nabla_X \nabla_Y = \nabla_Y \nabla_X$ from Proposition 3.7.12 **(b)**, applied to $X = X_i$ and $Y = X_j$). In other words, the $k$ elements $\nabla_{X_1}, \nabla_{X_2}, \ldots, \nabla_{X_k}$ pairwise commute. Similarly, the $k$ elements $\nabla^-_{X_1}, \nabla^-_{X_2}, \ldots, \nabla^-_{X_k}$ pairwise commute (here we need to use the equality $\nabla^-_X \nabla^-_Y = \nabla^-_Y \nabla^-_X$ from Proposition 3.7.12 **(b)**). Thus, Proposition 3.7.13 **(c)** follows. □

## 3.8. Orbits and cycles

### 3.8.1. Orbits, cycles and reflection lengths

Let us now recall some elementary combinatorics of permutations (see, e.g., [22fco, Lectures 27–28, §4.3], [21s, §5.5]):

**Definition 3.8.1.** Let $X$ be a finite set. Let $\sigma$ be a permutation of $X$.

Consider the binary relation $\overset{\sigma}{\sim}$ on the set $X$ defined as follows: For any $i, j \in X$, we set

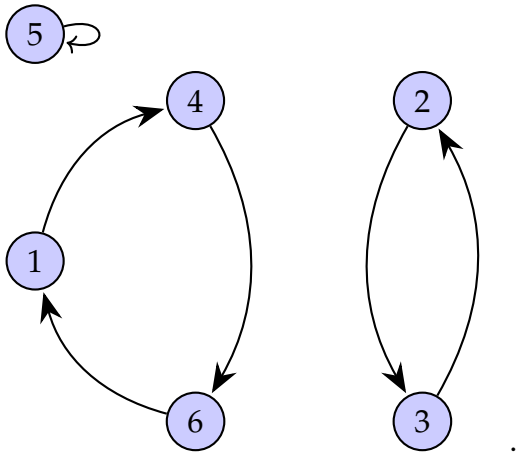
$$\left(i \overset{\sigma}{\sim} j\right) \iff \left(i = \sigma^k(j) \text{ for some } k \in \mathbb{N}\right).$$

It is not hard to see that this relation $\overset{\sigma}{\sim}$ is an equivalence relation (see [22fco, Lecture 27, Proposition 4.3.5]). Its equivalence classes are called the *orbits* of $\sigma$. Each element of $X$ belongs to exactly one orbit of $\sigma$.

**Example 3.8.2.** The orbits of $\operatorname{oln}(432651) \in S_6$ are $\{1,4,6\}$, $\{2,3\}$ and $\{5\}$. This is easiest to see using the *cycle digraph* of this permutation, which is a convenient way of visualizing it:

Generally, if $\sigma$ is a permutation of a finite set $X$, then the *cycle digraph* of $\sigma$ is the digraph (= directed graph) whose vertices are the elements of $X$, and whose arcs (= directed edges) are the pairs $(i, \sigma(i))$ for each vertex $i$. Here is

how this digraph looks like for $\sigma = \operatorname{oln}(432651) \in S_6$:




.

The digraph visibly consists of three vertex-disjoint cycles (i.e., three cycles that pairwise have no vertices in common). The vertex sets of these three cycles are exactly the three orbits of the permutation.

The orbits of a permutation have a rather simple structure (see [22fco, Lecture 27, Proposition 4.3.4] for a proof):

**Remark 3.8.3.** Let $X$ be a finite set. Let $\sigma$ be a permutation of $X$. Let $O$ be an orbit of $\sigma$. Let $a \in O$. Then,

$$\begin{aligned} O &= \left\{ \sigma^k(a) \mid k \in \mathbb{N} \right\} \\ &= \left\{ \sigma^0(a), \sigma^1(a), \sigma^2(a), \ldots \right\}. \end{aligned}$$

Moreover, if we set $m = |O|$, then $\sigma^m(a) = a$ and

$$O = \left\{ \sigma^0(a), \sigma^1(a), \ldots, \sigma^{m-1}(a) \right\}.$$

The orbits of a permutation do not uniquely determine it. For example, the permutations $\operatorname{cyc}_{1,2,3}$ and $\operatorname{cyc}_{1,3,2}$ in $S_3$ have the same orbits (namely, the single orbit $\{1,2,3\}$). However, there is a way to "refine" the structure of the orbits to get something that does determine the permutation. This refinement is the *disjoint cycle decomposition* of the permutation:

**Definition 3.8.4** (disjoint cycle decomposition)**.** Let $X$ be a finite set. Let $\sigma$ be

a permutation of $X$. Then, there is a list

$$\begin{aligned}&\Big( \left(a_{1,1}, a_{1,2}, \ldots, a_{1,n_1}\right), \\ &\quad \left(a_{2,1}, a_{2,2}, \ldots, a_{2,n_2}\right), \\ &\quad \ldots, \\ &\quad \left(a_{k,1}, a_{k,2}, \ldots, a_{k,n_k}\right) \Big)\end{aligned} \tag{57}$$

of nonempty lists of elements of $X$ such that:

- each element of $X$ appears exactly once in the composite list

$$\begin{aligned}&(a_{1,1}, a_{1,2}, \ldots, a_{1,n_1}, \\ &\quad a_{2,1}, a_{2,2}, \ldots, a_{2,n_2}, \\ &\quad \ldots, \\ &\quad a_{k,1}, a_{k,2}, \ldots, a_{k,n_k}),\end{aligned} \tag{58}$$

  and

- we have

$$\sigma = \operatorname{cyc}_{a_{1,1},a_{1,2},\ldots,a_{1,n_1}} \circ \operatorname{cyc}_{a_{2,1},a_{2,2},\ldots,a_{2,n_2}} \circ \cdots \circ \operatorname{cyc}_{a_{k,1},a_{k,2},\ldots,a_{k,n_k}}$$

(see [22fco, Lecture 27, Theorem 4.3.11] or [21s, Theorem 5.5.2] for the proof). Such a list is called a *disjoint cycle decomposition* (or, for short, *DCD*) of $\sigma$. Its entries (which themselves are lists of elements of $X$) are called the *cycles* of $\sigma$.

The DCD of $\sigma$ is unique up to permuting the cycles and rotating each cycle. See [22fco, Lecture 27, Theorem 4.3.11] or [21s, Theorem 5.5.2] for the details of this statement and for a proof outline.

**Example 3.8.5.** The permutation $\operatorname{oln}(432651) \in S_6$ has disjoint cycle decomposition

$$((1,4,6),\ (2,3),\ (5)),$$

since each element of $[6]$ appears exactly once in the composite list $(1,4,6,2,3,5)$, and because $\sigma = \operatorname{cyc}_{1,4,6} \circ \operatorname{cyc}_{2,3} \circ \operatorname{cyc}_5$. Thus, the cycles of $\sigma$ are $(1,4,6)$, $(2,3)$ and $(5)$. (Do not forget the length-1 cycle $(5)$. Even though the permutation $\operatorname{cyc}_5$ is just the identity, the cycle $(5)$ needs to be listed, since otherwise 5 would not appear in the composite list.)

As we said, the DCD of $\sigma$ is unique up to permuting the cycles and rotating each cycle. Thus, another DCD of $\sigma$ is

$$((2,3),\ (5),\ (6,1,4)).$$

We stress that the cycles of $\sigma$ are determined only up to cyclic rotation (e.g., we cannot distinguish between $(1,4,6)$ and $(6,1,4)$, although we can distinguish them both from $(1,6,4)$); but their lengths are uniquely determined by $\sigma$ (as a multiset, since the order in which the cycles appear in the DCD is also not fixed).

Obviously, a permutation $\sigma$ is uniquely determined by its cycles (unlike by its orbits). The orbits of $\sigma$ are just a shadow of the cycles: Indeed, they are just the cycles of $\sigma$ transformed into sets (i.e., the sets of entries of all cycles of $\sigma$). Let us state this rigorously:

**Proposition 3.8.6.** Let $X$ be a finite set. Let $\sigma \in S_X$ be any permutation. Let

$$\begin{aligned}\Big( &\left(a_{1,1}, a_{1,2}, \ldots, a_{1,n_1}\right), \\ &\left(a_{2,1}, a_{2,2}, \ldots, a_{2,n_2}\right), \\ &\ldots, \\ &\left(a_{k,1}, a_{k,2}, \ldots, a_{k,n_k}\right) \Big)\end{aligned} \tag{59}$$

be a DCD of $\sigma$. Then, the orbits of $\sigma$ are

$$\begin{aligned}&\left\{a_{1,1}, a_{1,2}, \ldots, a_{1,n_1}\right\}, \\ &\left\{a_{2,1}, a_{2,2}, \ldots, a_{2,n_2}\right\}, \\ &\ldots, \\ &\left\{a_{k,1}, a_{k,2}, \ldots, a_{k,n_k}\right\}.\end{aligned}$$

*Proof of Proposition 3.8.6.* We shall use the notation "$x \overset{f}{\mapsto} y$" for "$f(x) = y$", where $f$ is a map and where $x$ is some value that $f$ can be applied to.

All the elements $a_{u,v}$ (for $u \in [k]$ and $v \in [n_u]$) are distinct (because of the words "exactly once" in Definition 3.8.4). Moreover, each element of $X$ has the form $a_{u,v}$ for some $u \in [k]$ and $v \in [n_u]$ (for the same reason).

For each $i \in [k]$, let us set

$$c_i := \operatorname{cyc}_{a_{i,1}, a_{i,2}, \ldots, a_{i,n_i}} \qquad \text{and} \qquad a_{i,n_i+1} := a_{i,1}.$$

Thus, for each $i \in [k]$ and each $j \in [n_i]$, we have

$$c_i\left(a_{i,j}\right) = a_{i,j+1}. \tag{60}$$

Moreover, for each $i \in [k]$ and $r \in [k] \setminus \{i\}$ and $j \in [n_r + 1]$, we have

$$c_i\left(a_{r,j}\right) = a_{r,j} \tag{61}$$

(since $a_{r,j}$ is one of the elements $a_{r,1}, a_{r,2}, \ldots, a_{r,n_r}$ (because $a_{r,n_r+1} = a_{r,1}$), and thus is **not** one of the elements $a_{i,1}, a_{i,2}, \ldots, a_{i,n_i}$ (because all the elements $a_{u,v}$ are distinct), and therefore is fixed by the cycle $c_i = \operatorname{cyc}_{a_{i,1}, a_{i,2}, \ldots, a_{i,n_i}}$).

We assumed that the list-of-lists (59) is a DCD of $\sigma$. Thus,

$$\begin{aligned}\sigma &= \underbrace{\operatorname{cyc}_{a_{1,1},a_{1,2},\ldots,a_{1,n_1}}}_{=c_1} \circ \underbrace{\operatorname{cyc}_{a_{2,1},a_{2,2},\ldots,a_{2,n_2}}}_{=c_2} \circ \cdots \circ \underbrace{\operatorname{cyc}_{a_{k,1},a_{k,2},\ldots,a_{k,n_k}}}_{=c_k} \\ &\qquad \begin{array}{ccc} \text{(by the definition of } c_1\text{)} & \text{(by the definition of } c_2\text{)} & \text{(by the definition of } c_k\text{)} \end{array} \\ &\qquad \text{(by the definition of a DCD)} \\ &= c_1 \circ c_2 \circ \cdots \circ c_k.\end{aligned}$$

Now, let $i \in [k]$. Furthermore, let $j \in [n_i]$. Recall that $\sigma = c_1 \circ c_2 \circ \cdots \circ c_k$. Thus, in order to compute $\sigma\left(a_{i,j}\right)$, we just need to apply the permutations $c_k, c_{k-1}, \ldots, c_1$ to $a_{i,j}$ successively in this order (i.e., to first apply $c_k$ to $a_{i,j}$, then apply $c_{k-1}$ to the result, then apply $c_{k-2}$ to the result, and so on). Let us do this:

$$\begin{aligned} a_{i,j} &\overset{c_k}{\mapsto} a_{i,j} && \text{(by (61))} \\ &\overset{c_{k-1}}{\mapsto} a_{i,j} && \text{(by (61))} \\ &\overset{c_{k-2}}{\mapsto} \cdots \\ &\overset{c_{i+1}}{\mapsto} a_{i,j} && \text{(by (61))} \\ &\overset{c_i}{\mapsto} a_{i,j+1} && \text{(by (60))} \\ &\overset{c_{i-1}}{\mapsto} a_{i,j+1} && \text{(by (61))} \\ &\overset{c_{i-2}}{\mapsto} a_{i,j+1} && \text{(by (61))} \\ &\overset{c_{i-3}}{\mapsto} \cdots \\ &\overset{c_1}{\mapsto} a_{i,j+1} && \text{(by (61))}. \end{aligned}$$

The final result of this procedure is $a_{i,j+1}$. Thus, $\sigma\left(a_{i,j}\right) = a_{i,j+1}$. In other words, $a_{i,j} \overset{\sigma}{\mapsto} a_{i,j+1}$.

Forget that we fixed $j$. We thus have shown that $a_{i,j} \overset{\sigma}{\mapsto} a_{i,j+1}$ for each $j \in [n_i]$. Hence,

$$a_{i,1} \overset{\sigma}{\mapsto} a_{i,2} \overset{\sigma}{\mapsto} a_{i,3} \overset{\sigma}{\mapsto} \cdots \overset{\sigma}{\mapsto} a_{i,n_i} \overset{\sigma}{\mapsto} a_{i,n_i+1} = a_{i,1}.$$

This shows that the elements obtained from $a_{i,1}$ by repeated application of $\sigma$ (that is, the elements of the form $\sigma^k\left(a_{i,1}\right)$ for $k \in \mathbb{N}$) are precisely $a_{i,1}, a_{i,2}, \ldots, a_{i,n_i}$. Thus, the set $\left\{a_{i,1}, a_{i,2}, \ldots, a_{i,n_i}\right\}$ is an orbit of $\sigma$.

Forget that we fixed $i$. We thus have shown that $\left\{a_{i,1}, a_{i,2}, \ldots, a_{i,n_i}\right\}$ is an orbit of $\sigma$ for each $i \in [k]$. In other words, all the $k$ sets

$$\begin{aligned} &\left\{a_{1,1}, a_{1,2}, \ldots, a_{1,n_1}\right\}, \\ &\left\{a_{2,1}, a_{2,2}, \ldots, a_{2,n_2}\right\}, \\ &\ldots, \\ &\left\{a_{k,1}, a_{k,2}, \ldots, a_{k,n_k}\right\} \end{aligned}$$

are orbits of $\sigma$. Since these $k$ sets include all elements of $X$ (because each element of $X$ has the form $a_{u,v}$ for some $u \in [k]$ and $v \in [n_u]$), it thus follows that $\sigma$ cannot have any further orbits; thus, the orbits of $\sigma$ are precisely the $k$ sets that we found. Proposition 3.8.6 is now proved. □

**Proposition 3.8.7.** Let $X$ be a finite set. Let $\sigma \in S_X$ be any permutation. Then:

**(a)** The cycles of $\sigma^{-1}$ are the cycles of $\sigma$, read backwards. (In other words, if $(a_1, a_2, \ldots, a_k)$ is a cycle of $\sigma$, then $(a_k, a_{k-1}, \ldots, a_1)$ is a cycle of $\sigma^{-1}$, and vice versa.)

**(b)** The orbits of $\sigma^{-1}$ are the orbits of $\sigma$.

*Proof.* **(a)** Let (57) be a DCD of $\sigma$. Then, each element of $X$ appears exactly once in the composite list (58), and the equality

$$\sigma = \operatorname{cyc}_{a_{1,1},a_{1,2},\ldots,a_{1,n_1}} \circ \operatorname{cyc}_{a_{2,1},a_{2,2},\ldots,a_{2,n_2}} \circ \cdots \circ \operatorname{cyc}_{a_{k,1},a_{k,2},\ldots,a_{k,n_k}}$$

holds (by the definition of a DCD). From the latter equality, we obtain

$$\begin{aligned}
\sigma^{-1} &= \left(\operatorname{cyc}_{a_{1,1},a_{1,2},\ldots,a_{1,n_1}} \circ \operatorname{cyc}_{a_{2,1},a_{2,2},\ldots,a_{2,n_2}} \circ \cdots \circ \operatorname{cyc}_{a_{k,1},a_{k,2},\ldots,a_{k,n_k}}\right)^{-1} \\
&= \underbrace{\operatorname{cyc}^{-1}_{a_{k,1},a_{k,2},\ldots,a_{k,n_k}}}_{\substack{=\operatorname{cyc}_{a_{k,n_k},a_{k,n_k-1},\ldots,a_{k,1}} \\ \text{(by (9))}}} \circ \underbrace{\operatorname{cyc}^{-1}_{a_{k-1,1},a_{k-1,2},\ldots,a_{k-1,n_{k-1}}}}_{\substack{=\operatorname{cyc}_{a_{k-1,n_{k-1}},a_{k-1,n_{k-1}-1},\ldots,a_{k-1,1}} \\ \text{(by (9))}}} \circ \cdots \circ \underbrace{\operatorname{cyc}^{-1}_{a_{1,1},a_{1,2},\ldots,a_{1,n_1}}}_{\substack{=\operatorname{cyc}_{a_{1,n_1},a_{1,n_1-1},\ldots,a_{1,1}} \\ \text{(by (9))}}} \\
&= \operatorname{cyc}_{a_{k,n_k},a_{k,n_k-1},\ldots,a_{k,1}} \circ \operatorname{cyc}_{a_{k-1,n_{k-1}},a_{k-1,n_{k-1}-1},\ldots,a_{k-1,1}} \circ \cdots \circ \operatorname{cyc}_{a_{1,n_1},a_{1,n_1-1},\ldots,a_{1,1}}.
\end{aligned}$$

This shows that the nested list

$$\begin{aligned}
&\Big( \left(a_{k,n_k}, a_{k,n_k-1}, \ldots, a_{k,1}\right), \\
&\quad \left(a_{k-1,n_{k-1}}, a_{k-1,n_{k-1}-1}, \ldots, a_{k-1,1}\right), \\
&\quad \ldots, \\
&\quad \left(a_{1,n_1}, a_{1,n_1-1}, \ldots, a_{1,1}\right) \Big)
\end{aligned}$$

is a DCD of $\sigma^{-1}$ (since each element of $X$ appears exactly once in the composite list

$$\begin{aligned}
&\big(a_{k,n_k}, a_{k,n_k-1}, \ldots, a_{k,1}, \\
&\quad a_{k-1,n_{k-1}}, a_{k-1,n_{k-1}-1}, \ldots, a_{k-1,1}, \\
&\quad \ldots, \\
&\quad a_{1,n_1}, a_{1,n_1-1}, \ldots, a_{1,1}\big)
\end{aligned}$$

(because each element of $X$ appears exactly once in the composite list (58))). Thus, the cycles of $\sigma^{-1}$ are

$$\begin{aligned}
&\left(a_{k,n_k}, a_{k,n_k-1}, \ldots, a_{k,1}\right), \\
&\left(a_{k-1,n_{k-1}}, a_{k-1,n_{k-1}-1}, \ldots, a_{k-1,1}\right), \\
&\ldots, \\
&\left(a_{1,n_1}, a_{1,n_1-1}, \ldots, a_{1,1}\right).
\end{aligned}$$

On the other hand, the cycles of $\sigma$ are

$$\begin{aligned}
&\left(a_{1,1}, a_{1,2}, \ldots, a_{1,n_1}\right), \\
&\left(a_{2,1}, a_{2,2}, \ldots, a_{2,n_2}\right), \\
&\ldots, \\
&\left(a_{k,1}, a_{k,2}, \ldots, a_{k,n_k}\right)
\end{aligned}$$

(since (57) is a DCD of $\sigma$). Comparing these, we see that the cycles of $\sigma^{-1}$ are the cycles of $\sigma$, read backwards (and listed in reverse order, but this does not matter, since the different cycles can be permuted at will anyway). This proves Proposition 3.8.7 **(a)**.

**(b)** This follows easily from part **(a)** using Proposition 3.8.6. But let us give a more direct proof:

Recall the binary relation $\overset{\sigma}{\sim}$ (from Definition 3.8.1) and the analogously defined relation $\overset{\sigma^{-1}}{\sim}$. As we know, the relation $\overset{\sigma}{\sim}$ is an equivalence relation, thus is symmetric. Hence, for any $i, j \in X$, we have the following chain of equivalences:

$$\begin{aligned}
\left(i \overset{\sigma}{\sim} j\right) &\iff \left(j \overset{\sigma}{\sim} i\right) \\
&\iff \left(j = \sigma^k(i) \text{ for some } k \in \mathbb{N}\right) \\
&\qquad \left(\text{by the definition of the relation } \overset{\sigma}{\sim}\right) \\
&\iff \left(i = \left(\sigma^k\right)^{-1}(j) \text{ for some } k \in \mathbb{N}\right) \\
&\qquad \left(\begin{array}{c}\text{because the equality } j = \sigma^k(i) \text{ can be equivalently} \\ \text{rewritten as } i = \left(\sigma^k\right)^{-1}(j)\end{array}\right) \\
&\iff \left(i = \left(\sigma^{-1}\right)^k(j) \text{ for some } k \in \mathbb{N}\right) \\
&\qquad \left(\text{since } \left(\sigma^k\right)^{-1} = \left(\sigma^{-1}\right)^k \text{ for any } k \in \mathbb{N}\right) \\
&\iff \left(i \overset{\sigma^{-1}}{\sim} j\right) \qquad \left(\text{by the definition of the relation } \overset{\sigma^{-1}}{\sim}\right).
\end{aligned}$$

In other words, the relation $\overset{\sigma}{\sim}$ and the relation $\overset{\sigma^{-1}}{\sim}$ are the same. Hence, the equivalence classes of the latter relation are the equivalence classes of the former. In other words, the orbits of $\sigma^{-1}$ are the orbits of $\sigma$ (since the orbits of $\sigma^{-1}$ are defined as the equivalence classes of $\overset{\sigma^{-1}}{\sim}$, while the orbits of $\sigma$ are defined as the equivalence classes of $\overset{\sigma}{\sim}$). This proves Proposition 3.8.7 **(b)**. □

**Definition 3.8.8.** Let $X$ be a finite set. Let $\sigma \in S_X$ be any permutation. The *reflection length* of $\sigma$ is defined to be the number

$$\operatorname{rl}\sigma := |X| - (\#\text{ of orbits of } \sigma).$$

This is a nonnegative integer, since $(\#\text{ of orbits of } \sigma) \leq |X|$. It is no larger than $|X| - 1$ (unless $X = \varnothing$), since $\sigma$ has at least one orbit.

The reflection length is closely connected to transpositions, as the following theorem ([22fco, Lecture 28, Theorem 4.3.23 **(b)**], [Conrad, "The sign of a permutation", Theorem 6.1]) shows:

**Theorem 3.8.9.** Let $X$ be a finite set. Let $\sigma \in S_X$ be any permutation. Then, $\sigma$ can be written as a product of $\operatorname{rl}\sigma$ many transpositions in $S_X$. Moreover, $\operatorname{rl}\sigma$ is the smallest number $p \in \mathbb{N}$ such that $\sigma$ can be written as a product of $p$ transpositions in $S_X$.

Thus the notation "reflection length". ("Reflection" is how geometers refer to transpositions.)

The proof of Theorem 3.8.9 boils down to the following important lemma:

**Lemma 3.8.10.** Let $X$ be a finite set. Let $\sigma \in S_X$ be any permutation. Let $i$ and $j$ be two distinct elements of $X$. Let $\tau = t_{i,j} \circ \sigma$. Then:

**(a)** If $i \overset{\sigma}{\sim} j$, then the permutation $\tau$ has 1 more orbit than $\sigma$.

**(b)** If $i \overset{\sigma}{\sim} j$, then we don't have $i \overset{\tau}{\sim} j$.

**(c)** If we don't have $i \overset{\sigma}{\sim} j$, then the permutation $\tau$ has 1 fewer orbit than $\sigma$.

**(d)** If we don't have $i \overset{\sigma}{\sim} j$, then we do have $i \overset{\tau}{\sim} j$.

*Proof.* See [22fco, Lecture 27, Lemma 4.3.13]. □

### 3.8.2. Nonstarters and starters

Next, we introduce some nonstandard but useful notation:

**Definition 3.8.11.** Let $\sigma \in S_n$.

An element $i \in [n]$ is called a *nonstarter* of $\sigma$ if the orbit of $\sigma$ that contains $i$ also contains some number smaller than $i$ (that is, if $i$ is **not** the smallest element of the orbit of $\sigma$ that contains $i$). Otherwise, $i$ is called a *starter* of $\sigma$. (Thus, $i$ is called a starter of $\sigma$ if and only if $i$ is the smallest element of the orbit of $\sigma$ that contains $i$.)

We let $\operatorname{NoSt}\sigma$ denote the set of all nonstarters of $\sigma$.

**Example 3.8.12.** The nonstarters of $\operatorname{oln}(432651) \in S_6$ are 3, 4 and 6. The starters of this permutation are thus 1, 2 and 5.

The following is easy to see:

**Remark 3.8.13.** Let $\sigma \in S_n$ be a permutation. Then:

**(a)** Any fixed point of $\sigma$ is a starter of $\sigma$.

**(b)** The number 1 is always a starter of $\sigma$, as long as $n \geq 1$.

**(c)** If $\sigma$ has no nonstarters, then $\sigma = \operatorname{id}$.

**(d)** We have $(\#\text{ of starters of } \sigma) = (\#\text{ of orbits of } \sigma)$.

**(e)** We have $|\operatorname{NoSt}\sigma| = \operatorname{rl}\sigma$.

*Proof of Remark 3.8.13.* **(a)** Let $i \in [n]$ be a fixed point of $\sigma$. Then, $\sigma(i) = i$, so that $\sigma^k(i) = i$ for each $k \in \mathbb{N}$. Thus, the orbit of $\sigma$ that contains $i$ is just the one-element set $\{i\}$. Hence, $i$ is the smallest element of this orbit, and thus a starter of $\sigma$ (by the definition of a starter). So we have shown that any fixed point of $\sigma$ is a starter of $\sigma$. This proves Remark 3.8.13 **(a)**.

**(b)** Assume that $n \geq 1$. Then, 1 is the smallest element of the orbit of $\sigma$ that contains 1 (since 1 is the smallest element of $[n]$). Hence, 1 is a starter of $\sigma$. This proves Remark 3.8.13 **(b)**.

**(c)** Assume that $\sigma$ has no nonstarters. Let $i \in [n]$. Then, $\sigma(i) \overset{\sigma}{\sim} i$ (by the definition of $\overset{\sigma}{\sim}$), and thus the orbit of $\sigma$ that contains $i$ must also contain $\sigma(i)$. If we had $\sigma(i) \neq i$, then the orbit of $\sigma$ that contains $i$ would contain at least two elements (since it would contain the two distinct elements $i$ and $\sigma(i)$), and thus it would contain at least one element that is **not** the smallest element of this orbit. This element would then be a nonstarter of $\sigma$ (by the definition of a nonstarter), which would contradict the assumption that $\sigma$ has no nonstarters. Thus, $\sigma(i) \neq i$ is impossible. Hence, $\sigma(i) = i$. Since we have proved this for each $i \in [n]$, we thus conclude that $\sigma = \operatorname{id}$. This proves Remark 3.8.13 **(c)**.

**(d)** A starter of $\sigma$ is defined to be an element of $[n]$ that is the smallest element of the orbit of $\sigma$ that contains it. Thus, each orbit of $\sigma$ contains exactly one starter of $\sigma$ (namely, its smallest element). Hence, the permutation

$\sigma$ has as many starters as it has orbits. In other words, (# of starters of $\sigma$) = (# of orbits of $\sigma$). This proves Remark 3.8.13 **(d)**.

**(e)** Recall that $\operatorname{NoSt}\sigma$ is defined as the set of all nonstarters of $\sigma$. Thus,

$$\begin{aligned}|\operatorname{NoSt}\sigma| &= (\#\text{ of nonstarters of }\sigma) = |[n]| - \underbrace{(\#\text{ of starters of }\sigma)}_{\substack{=(\#\text{ of orbits of }\sigma)\\ \text{(by Remark 3.8.13 \textbf{(d)})}}} \\ &\qquad \left(\begin{array}{c}\text{since each } i \in [n] \text{ is either a starter of } \sigma \\ \text{or a nonstarter of } \sigma\text{, but not both}\end{array}\right) \\ &= |[n]| - (\#\text{ of orbits of }\sigma) = \operatorname{rl}\sigma\end{aligned}$$

(by Definition 3.8.8). This proves Remark 3.8.13 **(e)**. □

### 3.8.3. Products of distinct jucys–murphies

We can now give an explicit formula for an arbitrary product of distinct jucys–murphies:

**Theorem 3.8.14.** Let $i_1, i_2, \ldots, i_k$ be $k$ distinct elements of $[n]$. Then,

$$\mathbf{m}_{i_1}\mathbf{m}_{i_2}\cdots\mathbf{m}_{i_k} = \sum_{\substack{w\in S_n;\\ \operatorname{NoSt} w=\{i_1,i_2,\ldots,i_k\}}} w.$$

**Remark 3.8.15.** If $i_1 = 1$, then both sides of this equality are 0, since $\mathbf{m}_1 = 0$ and since 1 is never a nonstarter of a permutation (by Remark 3.8.13 **(b)**).

**Example 3.8.16.** Let $n = 5$. Then, Theorem 3.8.14 yields

$$\mathbf{m}_2\mathbf{m}_4 = \sum_{\substack{w\in S_5;\\ \operatorname{NoSt} w=\{2,4\}}} w.$$

And indeed, it is not hard to see that both sides equal $\operatorname{cyc}_{1,4,2} + \operatorname{cyc}_{1,2,4} + t_{1,2}t_{3,4}$.

To prove Theorem 3.8.14, we need the following two lemmas:

**Lemma 3.8.17.** Let $w \in S_n$. Then:

**(a)** If $p \in [n]$ is larger than each nonstarter of $w$, then $w(p) = p$.

**(b)** If $q$ is the largest nonstarter of $w$, then $w(q) \in [q-1]$.

*Proof of Lemma 3.8.17.* In the following, the word "nonstarter" will always mean "nonstarter of $w$". Likewise, the words "starter" and "orbit" will mean "starter of $w$" or "orbit of $w$".

**(a)** Let $p \in [n]$ be larger than each nonstarter of $w$. Then, $p$ cannot be a nonstarter. In other words, $p$ is a starter. In yet other words, $p$ is the smallest element of the orbit that contains $p$ (by the definition of "starter").

But $w(p) \overset{w}{\sim} p$ (by the definition of the relation $\overset{w}{\sim}$), and therefore $w(p)$ belongs to the orbit that contains $p$. Since $p$ is the smallest element of this orbit, we thus obtain $w(p) \geq p$. In other words, $p$ is not larger than $w(p)$.

Hence, $w(p)$ is not a nonstarter (since $p$ is larger than each nonstarter, but not larger than $w(p)$). In other words, $w(p)$ is a starter.

But recall that a starter is an element that is the smallest element of the orbit that contains it. Hence, each orbit contains exactly one starter. Therefore, any two starters belonging to the same orbit must be equal. Since $w(p)$ and $p$ are two starters belonging to the same orbit (because $w(p) \overset{w}{\sim} p$), we thus conclude that $w(p) = p$. This proves Lemma 3.8.17 **(a)**.

**(b)** First, we note that $w$ is a permutation and thus injective.

Let $q$ be the largest nonstarter of $w$. If we had $w(q) = q$, then $q$ would be a fixed point of $w$ and therefore a starter (since Remark 3.8.13 **(a)** yields that any fixed point of $w$ is a starter of $w$), which would contradict the fact that $q$ is a nonstarter. Hence, we cannot have $w(q) = q$. Thus, $w(q) \neq q$.

Let $p = w(q)$. If we had $p > q$, then $p$ would be larger than each nonstarter of $w$ (since $q$ is the largest nonstarter of $w$), and thus we would have $w(p) = p$ (by Lemma 3.8.17 **(a)**); but this would yield $w(p) = p = w(q)$ and therefore $p = q$ (since $w$ is injective), which would contradict $p > q$. Hence, we cannot have $p > q$. Thus, we have $p \leq q$. In other words, $w(q) \leq q$ (since $p = w(q)$). Combined with $w(q) \neq q$, this yields $w(q) < q$. Hence, $w(q) \leq q - 1$, so that $w(q) \in [q-1]$. This proves Lemma 3.8.17 **(b)**. □

**Lemma 3.8.18.** Let $w \in S_n$. Let $p, q \in [n]$ be two elements with $p < q$. Then:

**(a)** If $w(q) = q$, then $\operatorname{NoSt}\left(t_{p,q} w\right) = (\operatorname{NoSt} w) \cup \{q\}$.

**(b)** If $w(q) = p$, then $\operatorname{NoSt}\left(t_{p,q} w\right) = (\operatorname{NoSt} w) \setminus \{q\}$.

*Proof of Lemma 3.8.18.* **(a)** Assume that $w(q) = q$. Thus, $w^k(q) = q$ for each $k \in \mathbb{N}$. Hence, the one-element set $\{q\}$ is an orbit of $w$. Let us denote this orbit by $Q$. The permutation $w$ acts on this orbit $Q$ by

$$q \overset{w}{\mapsto} q \tag{62}$$

(where the notation "$x \overset{w}{\mapsto} y$" means "$w(x) = y$").

Let $P$ be the orbit of $w$ that contains $p$. Set $m = |P|$. Then, by Remark 3.8.3 (applied to $X = [n]$ and $\sigma = w$ and $O = P$ and $a = p$), we have $w^m(p) = p$ and

$$P = \left\{ w^0(p),\ w^1(p),\ \ldots,\ w^{m-1}(p) \right\}. \tag{63}$$

The permutation $w$ acts on the orbit $P$ by

$$w^0(p) \overset{w}{\mapsto} w^1(p) \overset{w}{\mapsto} w^2(p) \overset{w}{\mapsto} \cdots \overset{w}{\mapsto} w^{m-1}(p) \overset{w}{\mapsto} w^m(p) = p = w^0(p).$$

Recall that $|P| = m$. In view of (63), we can rewrite this as

$$\left|\left\{w^0(p),\ w^1(p),\ \ldots,\ w^{m-1}(p)\right\}\right| = m.$$

Hence, the $m$ elements $w^0(p),\ w^1(p),\ \ldots,\ w^{m-1}(p)$ are distinct.

The two orbits $P$ and $Q$ are distinct[13]. Hence, none of the elements $w^0(p),\ w^1(p),\ \ldots,\ w^{m-1}(p)$ equals $q$ (since they belong to a different orbit of $w$ than $q$ does[14]). Thus, in particular, none of the elements $w^0(p),\ w^1(p),\ \ldots,\ w^{m-2}(p)$ equals $q$. Moreover, none of these elements $w^0(p),\ w^1(p),\ \ldots,\ w^{m-2}(p)$ equals $w^{m-1}(p)$ either (since the $m$ elements $w^0(p),\ w^1(p),\ \ldots,\ w^{m-1}(p)$ are distinct).

Now, let $v$ be the permutation $t_{p,q}w$. This permutation $v$ is obtained from $w$ by swapping the two values $p$ and $q$ (since it is the composition of $t_{p,q}$ with $w$). Thus, $v$ sends $q$ not to $w(q) = q$ but to $p$ instead, while sending $w^{m-1}(p)$ not to $w^m(p) = p$ but to $q$ instead; on all other inputs, $v$ acts in the exact same way as $w$. Hence, the relations

$$w^0(p) \overset{w}{\mapsto} w^1(p) \overset{w}{\mapsto} w^2(p) \overset{w}{\mapsto} \cdots \overset{w}{\mapsto} w^{m-1}(p)$$

(which we have proved above) become

$$w^0(p) \overset{v}{\mapsto} w^1(p) \overset{v}{\mapsto} w^2(p) \overset{v}{\mapsto} \cdots \overset{v}{\mapsto} w^{m-1}(p)$$

(since none of the elements $w^0(p),\ w^1(p),\ \ldots,\ w^{m-2}(p)$ equals $q$, and none of them equals $w^{m-1}(p)$ either). Combining this chain of relations with $w^{m-1}(p) \overset{v}{\mapsto} q$ (since $v$ sends $w^{m-1}(p)$ to $q$) and with $q \overset{v}{\mapsto} w^0(p)$ (since $v$ sends $q$ to $p = w^0(p)$), we obtain a single circular chain

$$q \overset{v}{\mapsto} w^0(p) \overset{v}{\mapsto} w^1(p) \overset{v}{\mapsto} w^2(p) \overset{v}{\mapsto} \cdots \overset{v}{\mapsto} w^{m-1}(p) \overset{v}{\mapsto} q.$$

Hence, the set $\left\{q,\ w^0(p),\ w^1(p),\ \ldots,\ w^{m-1}(p)\right\}$ is an orbit of $v$. In other words, $P \cup Q$ is an orbit of $v$ (since

$$\begin{aligned}\left\{q,\ w^0(p),\ w^1(p),\ \ldots,\ w^{m-1}(p)\right\} &= \underbrace{\{q\}}_{=Q} \cup \underbrace{\left\{w^0(p),\ w^1(p),\ \ldots,\ w^{m-1}(p)\right\}}_{=P} \\ &= Q \cup P = P \cup Q\end{aligned}$$

[13] *Proof.* From $p < q$, we obtain $p \neq q$, so that $p \notin \{q\} = Q$. Consider the two orbits $P$ and $Q$ of $w$. The orbit $P$ contains $p$ (by its definition), whereas the orbit $Q$ does not (since $p \notin Q$). Thus, the two orbits $P$ and $Q$ are distinct, qed.

[14] *Proof.* The two orbits $P$ and $Q$ are distinct, and therefore disjoint (since any two distinct orbits of $w$ are disjoint). Thus, none of the elements $w^0(p),\ w^1(p),\ \ldots,\ w^{m-1}(p)$ of $P$ belongs to $Q$. Hence, none of the elements $w^0(p),\ w^1(p),\ \ldots,\ w^{m-1}(p)$ of $P$ equals $q$ (since $q$ belongs to $Q$).

). Thus, when we pass from the permutation $w$ to $v$, the two orbits $P$ and $Q$ are merged into a single orbit $P \cup Q$. All the other orbits of $w$ remain orbits of $v$ (since $v$ differs from $w$ only in that the values $p$ and $q$ are swapped, and this swap clearly does not affect any orbits except for the ones that contain $p$ or $q$). In other words, the orbits of $v$ differ from the orbits of $w$ only in that the two orbits $P$ and $Q$ are merged into a single orbit $P \cup Q$.

Now recall that the starters of a permutation are the smallest elements of its orbits (by the definition of a starter). Each orbit $O$ thus produces exactly one starter, namely its smallest element $\min O$. In particular, the two distinct orbits $P$ and $Q$ of $w$ thus produce two distinct starters $\min P$ and $\min Q$ of $w$. Hence, $\min P$ and $\min Q$ are two distinct starters of $w$, that is, two distinct elements of the set $\{\text{starters of } w\}$. Therefore,

$$\min P \in \{\text{starters of } w\} \setminus \{\min Q\}.$$

Recall that when we pass from the permutation $w$ to $v$, the two orbits $P$ and $Q$ are merged into a single orbit $P \cup Q$, while all other orbits stay unchanged. Thus, when we pass from $w$ to $v$, all the starters stay unchanged, except that the two starters $\min P$ and $\min Q$ coming from the orbits $P$ and $Q$ get replaced by a single starter $\min(P \cup Q)$ coming from the new orbit $P \cup Q$ (since the starters of a permutation are the smallest elements of its orbits). Hence,

$$\{\text{starters of } v\} = (\{\text{starters of } w\} \setminus \{\min P,\ \min Q\}) \cup \{\min(P \cup Q)\}.$$

Now, recall that $Q = \{q\}$, so that $\min Q = q$. This shows that $q$ is a starter of $w$ (since $\min Q$ is a starter of $w$). In other words, $q \in \{\text{starters of } w\}$.

However, $p \in P$ (by the definition of $P$) and thus $\min P \leq p$ (since the smallest element of a set is $\leq$ to any element of this set). Hence, $\min P \leq p < q = \min Q$. Furthermore, recall that the minimum of the union of two sets is the smaller one among their individual minima. Thus,

$$\min(P \cup Q) = \min\{\min P,\ \min Q\} = \min P$$

(since $\min P < \min Q$). Thus, our above computation becomes

$$\begin{aligned}
\{\text{starters of } v\} &= \underbrace{(\{\text{starters of } w\} \setminus \{\min P,\ \min Q\})}_{=(\{\text{starters of } w\} \setminus \{\min Q\}) \setminus \{\min P\}} \cup \left\{ \underbrace{\min(P \cup Q)}_{=\min P} \right\} \\
&= ((\{\text{starters of } w\} \setminus \{\min Q\}) \setminus \{\min P\}) \cup \{\min P\} \\
&= \{\text{starters of } w\} \setminus \{\min Q\}
\end{aligned}$$

(since $\min P \in \{\text{starters of } w\} \setminus \{\min Q\}$). Since $\min Q = q$, we can rewrite this as

$$\{\text{starters of } v\} = \{\text{starters of } w\} \setminus \{q\}.$$

Now, the definition of $\operatorname{NoSt} w$ yields

$$\operatorname{NoSt} w = \{\text{nonstarters of } w\} = [n] \setminus \{\text{starters of } w\}$$

(since the nonstarters of $w$ are the elements $i \in [n]$ that are not starters of $w$). Similarly,

$$\begin{aligned}
\operatorname{NoSt} v &= [n] \setminus \underbrace{\{\text{starters of } v\}}_{=\{\text{starters of } w\} \setminus \{q\}} \\
&= [n] \setminus (\{\text{starters of } w\} \setminus \{q\}) \\
&= \underbrace{([n] \setminus \{\text{starters of } w\})}_{=\operatorname{NoSt} w} \cup \{q\} \qquad \left( \begin{array}{c} \text{since } q \in \{\text{starters of } w\} \\ \text{and } \{\text{starters of } w\} \subseteq [n] \end{array} \right) \\
&= (\operatorname{NoSt} w) \cup \{q\}.
\end{aligned}$$

In other words, $\operatorname{NoSt}\left(t_{p,q} w\right) = (\operatorname{NoSt} w) \cup \{q\}$ (since $v = t_{p,q} w$). This proves Lemma 3.8.18 **(a)**.

**(b)** Assume that $w(q) = p$. Hence, $\left(t_{p,q} w\right)(q) = t_{p,q}\left(\underbrace{w(q)}_{=p}\right) = t_{p,q}(p) = q$.

Hence, $q$ is a fixed point of $t_{p,q} w$, and thus is a starter of $t_{p,q} w$ (since Remark 3.8.13 **(a)** yields that any fixed point of $t_{p,q} w$ is a starter of $t_{p,q} w$). Thus, $q$ is not a nonstarter of $t_{p,q} w$. In other words, $q \notin \operatorname{NoSt}\left(t_{p,q} w\right)$.

We have $\left(t_{p,q} w\right)(q) = q$. Thus, Lemma 3.8.18 **(a)** (applied to $t_{p,q} w$ instead of $w$) yields $\operatorname{NoSt}\left(t_{p,q} t_{p,q} w\right) = \left(\operatorname{NoSt}\left(t_{p,q} w\right)\right) \cup \{q\}$. In view of $\underbrace{t_{p,q} t_{p,q}}_{=t_{p,q}^2 = \mathrm{id}} w = w$, we can rewrite this as

$$\operatorname{NoSt} w = \left(\operatorname{NoSt}\left(t_{p,q} w\right)\right) \cup \{q\}.$$

Hence,

$$(\operatorname{NoSt} w) \setminus \{q\} = \left(\left(\operatorname{NoSt}\left(t_{p,q} w\right)\right) \cup \{q\}\right) \setminus \{q\} = \operatorname{NoSt}\left(t_{p,q} w\right)$$

(since $q \notin \operatorname{NoSt}\left(t_{p,q} w\right)$). In other words, $\operatorname{NoSt}\left(t_{p,q} w\right) = (\operatorname{NoSt} w) \setminus \{q\}$. This proves Lemma 3.8.18 **(b)**. □

*Proof of Theorem 3.8.14.* We induct on $k$.

*Base case:* We must prove Theorem 3.8.14 in the case when $k = 0$. In this case, the product $\mathbf{m}_{i_1} \mathbf{m}_{i_2} \cdots \mathbf{m}_{i_k}$ is an empty product, and thus equals 1. On the other hand, the sum $\sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}}} w$ equals id in this case, because its only addend is id (since the only permutation $w \in S_n$ that has no nonstarters is id (by Remark 3.8.13 **(c)**)). Thus, in the case when $k = 0$, the claim of Theorem 3.8.14 boils down to $1 = \mathrm{id}$, which is true. This completes the base case.

*Induction step:* We proceed from $k-1$ to $k$. Thus, we fix a positive integer $k$, and we set out to prove Theorem 3.8.14 for $k$, assuming (as the induction hypothesis) that it has already been proved for $k-1$.

So let $i_1, i_2, \ldots, i_k$ be $k$ distinct elements of $[n]$. We must prove that

$$\mathbf{m}_{i_1}\mathbf{m}_{i_2}\cdots\mathbf{m}_{i_k} = \sum_{\substack{w\in S_n; \\ \operatorname{NoSt} w = \{i_1,i_2,\ldots,i_k\}}} w. \tag{64}$$

This claim does not change if we permute the elements $i_1, i_2, \ldots, i_k$ (since the jucys–murphies commute). Thus, we can WLOG assume that $i_1 \leq i_2 \leq \cdots \leq i_k$. Assume this. Hence, $i_1 < i_2 < \cdots < i_k$ (since $i_1, i_2, \ldots, i_k$ are distinct).

Our induction hypothesis yields

$$\mathbf{m}_{i_1}\mathbf{m}_{i_2}\cdots\mathbf{m}_{i_{k-1}} = \sum_{\substack{w\in S_n; \\ \operatorname{NoSt} w = \{i_1,i_2,\ldots,i_{k-1}\}}} w. \tag{65}$$

Now,

$$\begin{aligned}
\mathbf{m}_{i_1}\mathbf{m}_{i_2}\cdots\mathbf{m}_{i_k} &= \left(\mathbf{m}_{i_1}\mathbf{m}_{i_2}\cdots\mathbf{m}_{i_{k-1}}\right)\mathbf{m}_{i_k} \\
&= \mathbf{m}_{i_k}\left(\mathbf{m}_{i_1}\mathbf{m}_{i_2}\cdots\mathbf{m}_{i_{k-1}}\right) \qquad \text{(since the jucys–murphies commute)} \\
&= \mathbf{m}_{i_k} \sum_{\substack{w\in S_n; \\ \operatorname{NoSt} w = \{i_1,i_2,\ldots,i_{k-1}\}}} w \qquad \text{(by (65))} \\
&= \sum_{p=1}^{i_k-1} t_{p,i_k} \sum_{\substack{w\in S_n; \\ \operatorname{NoSt} w = \{i_1,i_2,\ldots,i_{k-1}\}}} w \\
&\qquad \left( \begin{array}{c} \text{since the definition of } \mathbf{m}_{i_k} \\ \text{says } \mathbf{m}_{i_k} = t_{1,i_k} + t_{2,i_k} + \cdots + t_{i_k-1,i_k} = \sum\limits_{p=1}^{i_k-1} t_{p,i_k} \end{array} \right) \\
&= \sum_{p=1}^{i_k-1} \sum_{\substack{w\in S_n; \\ \operatorname{NoSt} w = \{i_1,i_2,\ldots,i_{k-1}\}}} t_{p,i_k} w. \qquad (66)
\end{aligned}$$

Let us now show that each $p \in [i_k - 1]$ satisfies

$$\sum_{\substack{w\in S_n; \\ \operatorname{NoSt} w = \{i_1,i_2,\ldots,i_{k-1}\}}} t_{p,i_k} w = \sum_{\substack{w\in S_n; \\ \operatorname{NoSt} w = \{i_1,i_2,\ldots,i_k\}; \\ w(i_k)=p}} w. \tag{67}$$

*Proof of (67).* Let $p \in [i_k - 1]$. Thus, $p \leq i_k - 1 < i_k$. Define the two sets

$$U = \{w \in S_n \mid \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\} \text{ and } w(i_k) = p\} \qquad \text{and}$$
$$V = \{w \in S_n \mid \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_{k-1}\}\}.$$

We shall now show that the maps

$$\begin{aligned}\Phi : V &\to U, \\ w &\mapsto t_{p,i_k} w\end{aligned}$$

and

$$\begin{aligned}\Psi : U &\to V, \\ w &\mapsto t_{p,i_k} w\end{aligned}$$

are well-defined.

*Well-definedness of $\Phi$:* In order to show that $\Phi$ is well-defined, we must prove that $t_{p,i_k} w \in U$ for each $w \in V$. So let $w \in V$ be arbitrary. Thus, $w \in S_n$ and $\operatorname{NoSt} w = \{i_1, i_2, \ldots, i_{k-1}\}$ (by the definition of $V$). The latter equality shows that the nonstarters of $w$ are $i_1, i_2, \ldots, i_{k-1}$. Hence, $i_k$ is larger than each nonstarter of $w$ (since $i_1 < i_2 < \cdots < i_k$). Thus, Lemma 3.8.17 **(a)** (applied to $i_k$ instead of $p$) shows that $w(i_k) = i_k$. Hence, Lemma 3.8.18 **(a)** (applied to $q = i_k$) yields that

$$\operatorname{NoSt}\left(t_{p,i_k} w\right) = \underbrace{\left(\operatorname{NoSt} w\right)}_{=\{i_1,i_2,\ldots,i_{k-1}\}} \cup \{i_k\} = \{i_1, i_2, \ldots, i_{k-1}\} \cup \{i_k\} = \{i_1, i_2, \ldots, i_k\}.$$

Furthermore, $\left(t_{p,i_k} w\right)(i_k) = t_{p,i_k}\left(\underbrace{w(i_k)}_{=i_k}\right) = t_{p,i_k}(i_k) = p$. These two equalities, combined, show that $t_{p,i_k} w \in U$ (by the definition of $U$). Thus, we have shown that $\Phi$ is well-defined.

*Well-definedness of $\Psi$:* In order to show that $\Psi$ is well-defined, we must prove that $t_{p,i_k} w \in V$ for each $w \in U$. So let $w \in U$ be arbitrary. Thus, $w \in S_n$ and $\operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}$ and $w(i_k) = p$ (by the definition of $U$). The equality $\operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}$ shows that the nonstarters of $w$ are $i_1, i_2, \ldots, i_k$. Hence, $i_k$ is the largest nonstarter of $w$ (since $i_1 < i_2 < \cdots < i_k$). Thus, Lemma 3.8.18 **(b)** (applied to $i_k$ instead of $q$) shows that

$$\begin{aligned}\operatorname{NoSt}\left(t_{p,i_k} w\right) &= \underbrace{\left(\operatorname{NoSt} w\right)}_{=\{i_1,i_2,\ldots,i_k\}} \setminus \{i_k\} \qquad (\text{since } w(i_k) = p) \\ &= \{i_1, i_2, \ldots, i_k\} \setminus \{i_k\} = \{i_1, i_2, \ldots, i_{k-1}\}\end{aligned}$$

(since $i_1, i_2, \ldots, i_k$ are distinct). Hence, $t_{p,i_k} w \in V$ (by the definition of $V$). Thus, we have shown that $\Psi$ is well-defined.

The maps $\Phi$ and $\Psi$ satisfy $\Phi \circ \Psi = \mathrm{id}$ (since the definitions of $\Phi$ and $\Psi$ yield $\Phi(\Psi(w)) = t_{p,i_k}\left(t_{p,i_k} w\right) = \underbrace{t_{p,i_k} t_{p,i_k}}_{=t_{p,i_k}^2 = \mathrm{id}} w = w$ for each $w \in U$) and $\Psi \circ \Phi = \mathrm{id}$ (for

similar reasons). Thus, they are mutually inverse. Hence, the map

$$\Phi : V \to U,$$
$$w \mapsto t_{p,i_k} w$$

is invertible (with inverse $\Psi$), and therefore is a bijection. Hence, we can substitute $t_{p,i_k} w$ for $w$ in the sum $\sum_{w \in U} w$, thus obtaining the equality

$$\sum_{w \in U} w = \sum_{w \in V} t_{p,i_k} w.$$

Recalling the definitions of $U$ and $V$, we can rewrite the two summation signs $\sum_{w \in U}$ and $\sum_{w \in V}$ in this equality as $\sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}; \\ w(i_k) = p}}$ and $\sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_{k-1}\}}}$, respectively. Thus, this equality rewrites as

$$\sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}; \\ w(i_k) = p}} w = \sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_{k-1}\}}} t_{p,i_k} w.$$

Hence, (67) is proved. □

However, it is easy to see that if a permutation $w \in S_n$ satisfies $\operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}$, then $w(i_k) \in [i_k - 1]$ [15]. Hence, we can split the sum $\sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}}} w$ according to the value of $w(i_k)$ as follows:

$$\sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}}} w = \underbrace{\sum_{p \in [i_k - 1]}}_{= \sum_{p=1}^{i_k - 1}} \underbrace{\sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}; \\ w(i_k) = p}} w}_{= \underbrace{\sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_{k-1}\}}} t_{p,i_k} w}_{\text{(by (67))}}} = \sum_{p=1}^{i_k - 1} \sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_{k-1}\}}} t_{p,i_k} w.$$

Comparing this with (66), we find

$$\mathbf{m}_{i_1} \mathbf{m}_{i_2} \cdots \mathbf{m}_{i_k} = \sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}}} w.$$

Thus, (64) is proved. This completes the induction step. Thus, Theorem 3.8.14 is proved. □

[15] *Proof.* Let $w \in S_n$ be a permutation that satisfies $\operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}$. The equality $\operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}$ shows that the nonstarters of $w$ are $i_1, i_2, \ldots, i_k$. Hence, $i_k$ is the largest nonstarter of $w$ (since $i_1 < i_2 < \cdots < i_k$). Hence, Lemma 3.8.17 **(b)** (applied to $q = i_k$) yields that $w(i_k) \in [i_k - 1]$, qed.

**Exercise 3.8.1.** [1] Reprove Theorem 3.4.3 using Theorem 3.8.14.

But we can get more mileage out of Theorem 3.8.14. Indeed, recall the following definition (a consequence of [21s, Definition 7.1.9 **(a)**]):

**Definition 3.8.19.** Let $u_1, u_2, \ldots, u_n$ be any $n$ commuting elements of a ring. Then, we define

$$e_k(u_1, u_2, \ldots, u_n) := \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} u_{i_1} u_{i_2} \cdots u_{i_k}$$

(the sum ranges over all $k$-tuples $(i_1, i_2, \ldots, i_k) \in [n]^k$ satisfying $i_1 < i_2 < \cdots < i_k$). In other words, we define $e_k(u_1, u_2, \ldots, u_n)$ to be the sum of all products of $k$ among the $n$ inputs $u_1, u_2, \ldots, u_n$.

This $e_k(u_1, u_2, \ldots, u_n)$ is called the *$k$-th elementary symmetric polynomial* in the inputs $u_1, u_2, \ldots, u_n$.

For example,

$$\begin{aligned}
e_0(u_1, u_2, \ldots, u_n) &= 1; \\
e_1(u_1, u_2, \ldots, u_n) &= u_1 + u_2 + \cdots + u_n; \\
e_2(u_1, u_2, \ldots, u_n) &= u_1u_2 + u_1u_3 + \cdots + u_1u_n \\
&\quad + u_2u_3 + \cdots + u_2u_n \\
&\quad + \cdots \\
&\quad + u_{n-1}u_n; \\
e_n(u_1, u_2, \ldots, u_n) &= u_1u_2 \cdots u_n; \\
e_k(u_1, u_2, \ldots, u_n) &= 0 \qquad \text{for } k > n.
\end{aligned}$$

The best-known example of $e_k(u_1, u_2, \ldots, u_n)$ is obtained when the inputs $u_1, u_2, \ldots, u_n$ are the $n$ indeterminates $x_1, x_2, \ldots, x_n$ of the polynomial ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$. In this case, we obtain the polynomial $e_k(x_1, x_2, \ldots, x_n)$, which is known as the *$k$-th elementary symmetric polynomial* in the $n$ indeterminates $x_1, x_2, \ldots, x_n$. This polynomial is the $e_k$ defined in [21s, Definition 7.1.9 **(a)**], and has many important properties (figuring, e.g., in Vieta's formulas). In the general case, $e_k(u_1, u_2, \ldots, u_n)$ can be obtained from this polynomial $e_k$ by substituting $u_i$ for each $x_i$.

The jucys–murphies $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n$ commute (by Theorem 3.4.2); thus, we can apply the elementary symmetric polynomials to them. We obtain something rather nice ([Jucys74, (8)]):

**Corollary 3.8.20.** For each $k \in \mathbb{N}$, we have

$$e_k(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n) = \sum_{\substack{w \in S_n; \\ w \text{ has exactly } n-k \text{ orbits}}} w = \sum_{\substack{w \in S_n; \\ \operatorname{rl} w = k}} w.$$

*Proof.* For each $w \in S_n$, we have

$$\begin{aligned}
|\operatorname{NoSt} w| &= (\#\text{ of nonstarters of } w) \qquad (\text{by the definition of } \operatorname{NoSt} w) \\
&= n - \underbrace{(\#\text{ of starters of } w)}_{\substack{=(\#\text{ of orbits of } w) \\ (\text{by Remark 3.8.13 } \textbf{(d)})}} \\
&\qquad \left( \begin{array}{c} \text{since each } i \in [n] \text{ is either a starter} \\ \text{or a nonstarter, but not both} \end{array} \right) \\
&= \underbrace{n}_{=|[n]|} - (\#\text{ of orbits of } w) \qquad (68) \\
&= |[n]| - (\#\text{ of orbits of } w) \\
&= \operatorname{rl} w \qquad (69)
\end{aligned}$$

(by Definition 3.8.8).

By Definition 3.8.19, we have

$$\begin{aligned}
&e_k(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n) \\
&= \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} \mathbf{m}_{i_1} \mathbf{m}_{i_2} \cdots \mathbf{m}_{i_k} \\
&= \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} \; \sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = \{i_1, i_2, \ldots, i_k\}}} w \qquad (\text{by Theorem 3.8.14}) \\
&= \underbrace{\sum_{\substack{I \subseteq [n]; \\ |I| = k}} \; \sum_{\substack{w \in S_n; \\ \operatorname{NoSt} w = I}} w}_{= \sum\limits_{\substack{w \in S_n; \\ |\operatorname{NoSt} w| = k}}} \qquad \left( \begin{array}{c} \text{here, we have substituted } I \text{ for } \{i_1, i_2, \ldots, i_k\} \\ \text{in the sum, since each } k\text{-element subset } I \\ \text{of } [n] \text{ can be written as } \{i_1, i_2, \ldots, i_k\} \\ \text{for a unique } k\text{-tuple } (i_1, i_2, \ldots, i_k) \in [n]^k \\ \text{satisfying } i_1 < i_2 < \cdots < i_k \end{array} \right) \\
&= \sum_{\substack{w \in S_n; \\ |\operatorname{NoSt} w| = k}} w = \sum_{\substack{w \in S_n; \\ \operatorname{rl} w = k}} w \qquad (\text{by (69)}) \\
&= \sum_{\substack{w \in S_n; \\ |\operatorname{NoSt} w| = k}} w = \sum_{\substack{w \in S_n; \\ n - (\#\text{ of orbits of } w) = k}} w \qquad (\text{by (68)}) \\
&= \sum_{\substack{w \in S_n; \\ (\#\text{ of orbits of } w) = n-k}} w \qquad \left( \begin{array}{c} \text{since the condition ``} n - (\#\text{ of orbits of } w) = k\text{''} \\ \text{is equivalent to `` } (\#\text{ of orbits of } w) = n - k\text{''} \end{array} \right) \\
&= \sum_{\substack{w \in S_n; \\ w \text{ has exactly } n-k \text{ orbits}}} w.
\end{aligned}$$

Thus, Corollary 3.8.20 is proved. □

**Corollary 3.8.21.** If $n \geq 1$, then

$$\mathbf{m}_2\mathbf{m}_3\cdots\mathbf{m}_n = \sum_{\substack{w\in S_n;\\ w \text{ has exactly 1 cycle}}} w.$$

*Proof.* Assume that $n \geq 1$. Then, Corollary 3.8.20 (applied to $k = n-1$) yields

$$e_{n-1}(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n) = \sum_{\substack{w\in S_n;\\ w \text{ has exactly 1 cycle}}} w. \tag{70}$$

Now, recall that $e_{n-1}(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n)$ is the sum of all products of $n-1$ among the $n$ jucys–murphies $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n$ (by Definition 3.8.19). However, all such products except for $\mathbf{m}_2\mathbf{m}_3\cdots\mathbf{m}_n$ are 0 (since they contain the factor $\mathbf{m}_1$, which is 0). Thus, $e_{n-1}(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n) = \mathbf{m}_2\mathbf{m}_3\cdots\mathbf{m}_n$. Comparing this with (70), we obtain

$$\mathbf{m}_2\mathbf{m}_3\cdots\mathbf{m}_n = \sum_{\substack{w\in S_n;\\ w \text{ has exactly 1 cycle}}} w.$$

This proves Corollary 3.8.21. □

**Question 3.8.22.** Can we generalize Theorem 3.8.14 to non-distinct $i_1, i_2, \ldots, i_k$ in some way? In other words, is there an explicit formula for a product $\mathbf{m}_{i_1}\mathbf{m}_{i_2}\cdots\mathbf{m}_{i_k}$ of non-distinct jucys–murphies? In particular, is there a formula for $\mathbf{m}_i^k$ ?

Be ready to see the same permutation multiple times in such a formula. For instance,

$$\mathbf{m}_3^2 = (t_{1,3} + t_{2,3})^2 = \underbrace{2}_{=2\,\mathrm{id}} + \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2}.$$

The following corollary is just a neat way to repack Corollary 3.8.20:

**Corollary 3.8.23.** Let $x \in \mathbf{k}$. Then,

$$\begin{aligned}&(1 + x\mathbf{m}_1)(1 + x\mathbf{m}_2)\cdots(1 + x\mathbf{m}_n)\\ &= \sum_{k\in\mathbb{N}} x^k e_k(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n) = \sum_{k\in\mathbb{N}} x^k \sum_{\substack{w\in S_n;\\ \mathrm{rl}\, w = k}} w = \sum_{w\in S_n} x^{\mathrm{rl}\, w} w.\end{aligned}$$

For instance, for $n = 3$, this is saying that

$$\begin{aligned}&(1 + x\mathbf{m}_1)(1 + x\mathbf{m}_2)(1 + x\mathbf{m}_3)\\ &= 1 + xt_{1,2} + xt_{1,3} + xt_{2,3} + x^2\,\mathrm{cyc}_{1,2,3} + x^2\,\mathrm{cyc}_{1,3,2}.\end{aligned}$$

*Proof of Corollary 3.8.23.* It is well-known that any $n$ elements $a_1, a_2, \ldots, a_n$ of a ring $R$ satisfy the equality

$$(1+a_1)(1+a_2)\cdots(1+a_n) = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} a_{i_1} a_{i_2} \cdots a_{i_k}.$$

(Indeed, this equality can be obtained by expanding the left hand side.) Applying this equality to $R = \mathbf{k}[S_n]$ and $a_i = x\mathbf{m}_i$, we obtain

$$\begin{aligned}
&(1+x\mathbf{m}_1)(1+x\mathbf{m}_2)\cdots(1+x\mathbf{m}_n) \\
&= \underbrace{\sum_{1 \le i_1 < i_2 < \cdots < i_k \le n}}_{= \sum\limits_{k \in \mathbb{N}} \ \sum\limits_{1 \le i_1 < i_2 < \cdots < i_k \le n}} \underbrace{\left(x\mathbf{m}_{i_1}\right)\left(x\mathbf{m}_{i_2}\right)\cdots\left(x\mathbf{m}_{i_k}\right)}_{= x^k \mathbf{m}_{i_1}\mathbf{m}_{i_2}\cdots\mathbf{m}_{i_k}} \\
&= \sum_{k \in \mathbb{N}} \ \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} x^k \mathbf{m}_{i_1}\mathbf{m}_{i_2}\cdots\mathbf{m}_{i_k} \\
&= \sum_{k \in \mathbb{N}} x^k \underbrace{\sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} \mathbf{m}_{i_1}\mathbf{m}_{i_2}\cdots\mathbf{m}_{i_k}}_{= e_k(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n) \text{ (by the definition of } e_k(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n))} \\
&= \sum_{k \in \mathbb{N}} x^k \underbrace{e_k(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n)}_{= \sum\limits_{\substack{w \in S_n; \\ \operatorname{rl} w = k}} w \text{ (by Corollary 3.8.20)}} = \sum_{k \in \mathbb{N}} x^k \sum_{\substack{w \in S_n; \\ \operatorname{rl} w = k}} w \\
&= \sum_{k \in \mathbb{N}} \ \sum_{\substack{w \in S_n; \\ \operatorname{rl} w = k}} \underbrace{x^k}_{= x^{\operatorname{rl} w} \text{ (since } k = \operatorname{rl} w)} w = \underbrace{\sum_{k \in \mathbb{N}} \ \sum_{\substack{w \in S_n; \\ \operatorname{rl} w = k}}}_{= \sum\limits_{w \in S_n}} x^{\operatorname{rl} w} w = \sum_{w \in S_n} x^{\operatorname{rl} w} w.
\end{aligned}$$

Thus, Corollary 3.8.23 is proved. □

## 3.9. Conjugacy classes and the center

### 3.9.1. The general case

Let us now recall some more abstract algebra, namely the definition of a *center* for groups and for rings:

**Definition 3.9.1. (a)** The *center* of a group $G$ is defined to be the subset

$$Z(G) := \{x \in G \mid xy = yx \text{ for all } y \in G\} \text{ of } G.$$

This is the set of all elements of $G$ that commute with all elements of $G$. It is always an abelian subgroup of $G$. If $G$ itself is abelian, then $Z(G) = G$.

**(b)** The *center* of a ring $R$ is defined to be the subset

$$Z(R) := \{x \in R \mid xy = yx \text{ for all } y \in R\} \text{ of } R.$$

This is the set of all elements of $R$ that commute with all elements of $R$. It is always a commutative subring of $R$. Moreover, if $R$ is a $\mathbf{k}$-algebra, then this center is also a $\mathbf{k}$-subalgebra of $R$. If $R$ itself is commutative, then $Z(R) = R$.

The two kinds of centers are clearly analogous, and one might wonder how they interact in the case of a group algebra. It is easy to see that if $g$ is an element of the center $Z(G)$ of a group $G$, then the corresponding standard basis vector $g = e_g \in \mathbf{k}[G]$ belongs to the center $Z(\mathbf{k}[G])$ of the group algebra $\mathbf{k}[G]$ (this is by linearity). Taking linear combinations of such basis vectors yields further elements of $Z(\mathbf{k}[G])$ (since $Z(\mathbf{k}[G])$ is a $\mathbf{k}$-subalgebra). However, not every element of $Z(\mathbf{k}[G])$ is obtained this way (in general). For example:

**Proposition 3.9.2.** The sum

$$\mathbf{m}_1 + \mathbf{m}_2 + \cdots + \mathbf{m}_n = \sum_{i<j} t_{i,j}$$

(that is, the sum of all transpositions in $S_n$) belongs to the center of $\mathbf{k}[S_n]$.

We will soon prove this. Note that none of the transpositions $t_{i,j}$ itself belongs to $Z(S_n)$ (unless $n \leq 2$).

Let us first study the center of $\mathbf{k}[G]$ for arbitrary $G$:

**Lemma 3.9.3.** Let $G$ be a group. Let $\mathbf{a} \in \mathbf{k}[G]$. Then, $\mathbf{a} \in Z(\mathbf{k}[G])$ if and only if each $w \in G$ satisfies $w\mathbf{a}w^{-1} = \mathbf{a}$.

*Proof.* $\Longrightarrow$: Assume that $\mathbf{a} \in Z(\mathbf{k}[G])$. Then, each $w \in G$ satisfies $w\mathbf{a} = \mathbf{a}w$ and therefore $\underbrace{w\mathbf{a}}_{=\mathbf{a}w} w^{-1} = \mathbf{a}\underbrace{ww^{-1}}_{=1} = \mathbf{a}$. In other words, each $w \in G$ satisfies $w\mathbf{a}w^{-1} = \mathbf{a}$. This proves the "$\Longrightarrow$" direction of Lemma 3.9.3.

$\Longleftarrow$: Assume that each $w \in G$ satisfies $w\mathbf{a}w^{-1} = \mathbf{a}$. Thus, each $w \in G$ satisfies

$$w\mathbf{a} = \mathbf{a}w \tag{71}$$

(since multiplying the equality $w\mathbf{a}w^{-1} = \mathbf{a}$ by $w$ on the right yields $w\mathbf{a}w^{-1}w = \mathbf{a}w$, thus $\mathbf{a}w = w\mathbf{a}\underbrace{w^{-1}w}_{=1} = w\mathbf{a}$). By linearity, this entails that any $\mathbf{k}$-linear combination $\sum_{w\in G} \lambda_w w$ of elements of $G$ (with $\lambda_w \in \mathbf{k}$ for all $w \in G$) satisfies

$$\left(\sum_{w\in G} \lambda_w w\right)\mathbf{a} = \sum_{w\in G} \lambda_w \underbrace{w\mathbf{a}}_{\substack{=\mathbf{a}w \\ \text{(by (71))}}} = \sum_{w\in G} \lambda_w \mathbf{a}w = \mathbf{a}\left(\sum_{w\in G} \lambda_w w\right).$$

In other words, $\mathbf{a}$ commutes with any $\mathbf{k}$-linear combination $\sum_{w \in G} \lambda_w w$ of elements of $G$. But this is just saying that $\mathbf{a}$ commutes with every element of $\mathbf{k}[G]$ (since every element of $\mathbf{k}[G]$ is such a $\mathbf{k}$-linear combination). In other words, $\mathbf{a} \in Z(\mathbf{k}[G])$. This proves the "$\Longleftarrow$" direction of Lemma 3.9.3. $\square$

Using Lemma 3.9.3, we can now prove Proposition 3.9.2:

*Proof of Proposition 3.9.2.* Let $T = \left\{ t_{i,j} \mid i,j \in [n] \text{ with } i < j \right\}$ be the set of all transpositions in $S_n$. Then,

$$\sum_{\sigma \in T} \sigma = \sum_{i<j} t_{i,j}$$

(since each transposition in $S_n$ can be written as $t_{i,j}$ for a unique pair of integers $i,j \in [n]$ with $i < j$).

Set $\mathbf{a} = \sum_{\sigma \in T} \sigma$. Then, $\mathbf{a}$ is the sum of all transpositions in $S_n$. Moreover,

$$\begin{aligned}
\mathbf{a} = \sum_{\sigma \in T} \sigma = \sum_{i<j} t_{i,j} &= \sum_{j=1}^{n} \underbrace{\sum_{i=1}^{j-1} t_{i,j}}_{\substack{=\mathbf{m}_j \\ \text{(by the definition of } \mathbf{m}_j\text{)}}} = \sum_{j=1}^{n} \mathbf{m}_j \\
&= \mathbf{m}_1 + \mathbf{m}_2 + \cdots + \mathbf{m}_n.
\end{aligned}$$

We shall now show that $\mathbf{a} \in Z(\mathbf{k}[S_n])$.

Indeed, let $w \in S_n$. For any transposition $\sigma \in T$, the permutation $w\sigma w^{-1}$ is again a transposition (since (5) shows that $wt_{i,j}w^{-1} = t_{w(i),w(j)}$ for any distinct elements $i,j \in [n]$). In other words, $w\sigma w^{-1} \in T$ for any $\sigma \in T$. Hence, the map

$$\begin{aligned}
&T \to T, \\
&\sigma \mapsto w\sigma w^{-1}
\end{aligned}$$

is well-defined. Similarly, the map

$$\begin{aligned}
&T \to T, \\
&\sigma \mapsto w^{-1}\sigma w
\end{aligned}$$

is well-defined. These two maps are easily seen to be mutually inverse (because for each $\sigma \in T$, we have $w^{-1}\left(w\sigma w^{-1}\right)w = \underbrace{w^{-1}w}_{=1}\sigma\underbrace{w^{-1}w}_{=1} = \sigma$ and similarly $w\left(w^{-1}\sigma w\right)w^{-1} = \sigma$), and thus are bijections. Hence, in particular, the map

$$\begin{aligned}
&T \to T, \\
&\sigma \mapsto w\sigma w^{-1}
\end{aligned}$$

is a bijection. Thus, we can substitute $w\sigma w^{-1}$ for $\sigma$ in the sum $\sum_{\sigma\in T}\sigma$. We thus obtain

$$\sum_{\sigma\in T}\sigma=\sum_{\sigma\in T}w\sigma w^{-1}=w\left(\sum_{\sigma\in T}\sigma\right)w^{-1}.$$

In view of $\mathbf{a}=\sum_{\sigma\in T}\sigma$, we can rewrite this as $\mathbf{a}=w\mathbf{a}w^{-1}$. In other words, $w\mathbf{a}w^{-1}=\mathbf{a}$.

Forget now that we fixed $w$. We thus have shown that each $w\in S_n$ satisfies $w\mathbf{a}w^{-1}=\mathbf{a}$. Hence, by Lemma 3.9.3 (applied to $G=S_n$), we conclude that $\mathbf{a}\in Z\left(\mathbf{k}\left[S_n\right]\right)$. In other words, $\mathbf{a}$ belongs to the center of $\mathbf{k}\left[S_n\right]$. In other words,

$$\mathbf{m}_1+\mathbf{m}_2+\cdots+\mathbf{m}_n=\sum_{i<j}t_{i,j}$$

belongs to the center of $\mathbf{k}\left[S_n\right]$ (since $\mathbf{a}=\mathbf{m}_1+\mathbf{m}_2+\cdots+\mathbf{m}_n=\sum_{i<j}t_{i,j}$). This proves Proposition 3.9.2. □

Let us try to describe $Z\left(\mathbf{k}\left[G\right]\right)$ more explicitly. For this we recall some classical notations from group theory:

**Definition 3.9.4.** Let $G$ be a group.

**(a)** Two elements $x$ and $y$ of $G$ are said to be *conjugate* if there exists a $g\in G$ such that $y=gxg^{-1}$. In this case, we write $x\sim y$.

**(b)** The binary relation $\sim$ on $G$ is an equivalence relation, and is known as *conjugacy*. Its equivalence classes are called the *conjugacy classes* of $G$. Thus, each element of $G$ belongs to exactly one conjugacy class.

**(c)** If $C$ is a finite conjugacy class of $G$, then we define $\mathbf{z}_C$ to be the sum $\sum_{c\in C}c$ in $\mathbf{k}\left[G\right]$. This is called a *conjugacy class sum* of $G$. (We require $C$ to be finite, since infinite sums do not make sense in $\mathbf{k}\left[G\right]$. However, this will not be important for us in the case $G=S_n$, since $S_n$ is finite.)

Here are some examples:

- If $G$ is an abelian group, then conjugacy in $G$ is just equality (i.e., we have $x\sim y$ if and only if $x=y$), so all the conjugacy classes are singletons $\{x\}$ (with $x\in G$), and thus the conjugacy class sums are just the standard basis vectors $x=e_x$ of $\mathbf{k}\left[G\right]$.

  More generally, if $x$ is any element of the center of any group $G$, then its conjugacy class is the singleton $\{x\}$ (since $gxg^{-1}=x$ for all $g\in G$).

- The conjugacy classes of the symmetric group $S_3$ are

$$\{\mathrm{id}\},\qquad\{t_{1,2},\ t_{1,3},\ t_{2,3}\},\qquad\left\{\mathrm{cyc}_{1,2,3},\ \mathrm{cyc}_{1,3,2}\right\}.$$

Thus, the conjugacy class sums of $S_3$ are

$$\mathrm{id} = 1, \qquad t_{1,2} + t_{1,3} + t_{2,3}, \qquad \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2}\,.$$

See [Conrad, "Conjugation in a group"] for more examples. You might notice that in all these examples, the conjugacy class sums all belong to $Z(\mathbf{k}[G])$. This does generalize, and even more is true:

**Theorem 3.9.5.** Let $G$ be a group. Then, the center $Z(\mathbf{k}[G])$ of the group algebra $\mathbf{k}[G]$ is the $\mathbf{k}$-linear span of the conjugacy class sums. Moreover, these sums are $\mathbf{k}$-linearly independent, so they form a basis of $Z(\mathbf{k}[G])$.

*Proof of Theorem 3.9.5.* We need to prove the following three claims:

*Claim 1:* Each conjugacy class sum $\mathbf{z}_C$ belongs to $Z(\mathbf{k}[G])$.

*Claim 2:* The conjugacy class sums span the $\mathbf{k}$-module $Z(\mathbf{k}[G])$.

*Claim 3:* The conjugacy class sums are $\mathbf{k}$-linearly independent.

*Proof of Claim 1.* Let $C$ be a finite conjugacy class of $G$. We must prove that $\mathbf{z}_C \in Z(\mathbf{k}[G])$.

We will prove this using Lemma 3.9.3. Recall that $\mathbf{z}_C = \sum_{c \in C} c$ (by the definition of $\mathbf{z}_C$).

Let $w \in G$. Then, $wcw^{-1} \in C$ for each $c \in C$ (since the element $wcw^{-1}$ is clearly conjugate to $c$, and thus belongs to the same conjugacy class as $c$ does, but the latter class is $C$ of course). Hence, the map

$$\begin{aligned} C &\to C, \\ c &\mapsto wcw^{-1} \end{aligned}$$

is well-defined. Similarly, the map

$$\begin{aligned} C &\to C, \\ c &\mapsto w^{-1}cw \end{aligned}$$

is well-defined. These two maps are mutually inverse (since each $c \in C$ satisfies $w^{-1}\left(wcw^{-1}\right)w = \underbrace{w^{-1}w}_{=1} c \underbrace{w^{-1}w}_{=1} = c$ and similarly $w\left(w^{-1}cw\right)w^{-1} = c$), and thus are bijections. In particular, the map

$$\begin{aligned} C &\to C, \\ c &\mapsto wcw^{-1} \end{aligned}$$

is a bijection. Hence, we can substitute $wcw^{-1}$ for $c$ in the sum $\sum_{c \in C} c$. We thus obtain

$$\sum_{c \in C} c = \sum_{c \in C} wcw^{-1} = w \left( \sum_{c \in C} c \right) w^{-1}.$$

In view of $\mathbf{z}_C = \sum_{c \in C} c$, we can rewrite this as $\mathbf{z}_C = w \mathbf{z}_C w^{-1}$. In other words, $w \mathbf{z}_C w^{-1} = \mathbf{z}_C$.

Forget that we fixed $w$. We thus have shown that each $w \in G$ satisfies $w \mathbf{z}_C w^{-1} = \mathbf{z}_C$. According to Lemma 3.9.3 (applied to $\mathbf{a} = \mathbf{z}_C$), we thus conclude that $\mathbf{z}_C \in Z(\mathbf{k}[G])$. This proves Claim 1. □

*Proof of Claim 2.* Claim 1 shows that these conjugacy class sums $\mathbf{z}_C$ all belong to $Z(\mathbf{k}[G])$. It remains to show that each $\mathbf{a} \in Z(\mathbf{k}[G])$ is a $\mathbf{k}$-linear combination of these sums. So let us do this.

Let $\mathbf{a} \in Z(\mathbf{k}[G])$ be arbitrary. Recall that $\mathbf{k}[G] = \mathbf{k}^{(G)}$ as a $\mathbf{k}$-module (by the definition of a monoid algebra). Hence, $\mathbf{a} \in Z(\mathbf{k}[G]) \subseteq \mathbf{k}[G] = \mathbf{k}^{(G)} \subseteq \mathbf{k}^G$. Hence, $\mathbf{a}$ is a family $(\alpha_w)_{w \in G}$ of elements of $\mathbf{k}$. Consider these elements. Thus, $\mathbf{a} = (\alpha_w)_{w \in G} = \sum_{w \in G} \alpha_w w$ (as we know from Subsection 3.1.5). Moreover, all but finitely many $w \in G$ satisfy $\alpha_w = 0$ (since $(\alpha_w)_{w \in G} = \mathbf{a} \in \mathbf{k}^{(G)}$). In other words, only finitely many $w \in G$ satisfy $\alpha_w \neq 0$. We now make the following observations:

- Any two conjugate elements $x$ and $y$ of $G$ satisfy

$$\alpha_x = \alpha_y. \tag{72}$$

[*Proof:* Let $x$ and $y$ be two conjugate elements of $G$. Thus, $y = gxg^{-1}$ for some $g \in G$. Consider this $g$. Now, $\mathbf{a} \in Z(\mathbf{k}[G])$, so that each $w \in G$ satisfies $w\mathbf{a}w^{-1} = \mathbf{a}$ (by Lemma 3.9.3). Applying this to $w = g$, we obtain $g\mathbf{a}g^{-1} = \mathbf{a} = \sum_{w \in G} \alpha_w w$. Comparing this with

$$\begin{aligned} g\mathbf{a}g^{-1} &= g \left( \sum_{w \in G} \alpha_w w \right) g^{-1} \qquad \left( \text{since } \mathbf{a} = \sum_{w \in G} \alpha_w w \right) \\ &= \sum_{w \in G} \alpha_w gwg^{-1} \\ &= \sum_{w \in G} \alpha_{g^{-1}wg} \underbrace{gg^{-1}}_{=1} w \underbrace{gg^{-1}}_{=1} \qquad \left( \begin{array}{c} \text{here, we have substituted } gwg^{-1} \\ \text{for } w \text{ in the sum, since the} \\ \text{map } G \to G,\ w \mapsto gwg^{-1} \\ \text{is a bijection} \end{array} \right) \\ &= \sum_{w \in G} \alpha_{g^{-1}wg} w, \end{aligned}$$

we obtain

$$\sum_{w \in G} \alpha_w w = \sum_{w \in G} \alpha_{g^{-1}wg} w.$$

Now, the left hand side of this equality can be rewritten as the family $(\alpha_w)_{w \in G}$ (since $\sum_{w \in G} \alpha_w w = (\alpha_w)_{w \in G}$), whereas the right hand side can be rewritten as the family $\left(\alpha_{g^{-1}wg}\right)_{w \in G}$ (for similar reasons). Thus, this equality rewrites as

$$(\alpha_w)_{w \in G} = \left(\alpha_{g^{-1}wg}\right)_{w \in G}.$$

In other words, $\alpha_w = \alpha_{g^{-1}wg}$ for each $w \in G$. Applying this to $w = y$, we obtain $\alpha_y = \alpha_{g^{-1}yg}$. Since $g^{-1} \underbrace{y}_{=gxg^{-1}} g = \underbrace{g^{-1}g}_{=1} x \underbrace{g^{-1}g}_{=1} = x$, we can rewrite this as $\alpha_y = \alpha_x$. This proves (72).]

- For each conjugacy class $C$ of $G$, there exists an element $\beta_C$ of $\mathbf{k}$ such that

$$\text{all } w \in C \text{ satisfy } \alpha_w = \beta_C. \tag{73}$$

[*Proof of (73):* Let $C$ be a conjugacy class of $G$. Pick any element $u$ of $C$, and set $\beta_C := \alpha_u$. We shall now show that all $w \in C$ satisfy $\alpha_w = \beta_C$. Indeed, for any $w \in C$, the elements $w$ and $u$ of $G$ are conjugate (since they both belong to the same conjugacy class, namely $C$), and thus satisfy $\alpha_w = \alpha_u$ (by (72), applied to $x = w$ and $y = u$) and therefore $\alpha_w = \alpha_u = \beta_C$. Thus, we have found an element $\beta_C$ of $\mathbf{k}$ such that all $w \in C$ satisfy $\alpha_w = \beta_C$. This proves (73).]

- If $C$ is an infinite conjugacy class of $G$, then the element $\beta_C$ of $\mathbf{k}$ that satisfies (73) is given by

$$\beta_C = 0. \tag{74}$$

[*Proof:* Let $C$ be an infinite conjugacy class of $G$. Then, there are infinitely many $w \in C$. If we had $\beta_C \neq 0$, then all these infinitely many $w \in C$ would satisfy $\alpha_w \neq 0$ (since (73) would yield $\alpha_w = \beta_C \neq 0$), which would contradict the fact that only finitely many $w \in G$ satisfy $\alpha_w \neq 0$. Hence, we cannot have $\beta_C \neq 0$. Thus, $\beta_C = 0$. This proves (74).]

For each conjugacy class $C$ of $G$, let us pick an element $\beta_C$ of $\mathbf{k}$ that satisfies (73). (We have already shown that such a $\beta_C$ exists.)

Now, recall that the conjugacy classes of $G$ are the equivalence classes of an equivalence relation, and thus are disjoint nonempty subsets of $G$ whose union

is $G$. Thus, we can split the sum $\sum_{w \in G} \alpha_w w$ as follows:

$$\begin{aligned}
\sum_{w \in G} \alpha_w w &= \sum_{\substack{C \text{ is a conjugacy} \\ \text{class of } G}} \sum_{w \in C} \underbrace{\alpha_w}_{\substack{=\beta_C \\ \text{(by (73))}}} w = \sum_{\substack{C \text{ is a conjugacy} \\ \text{class of } G}} \sum_{w \in C} \beta_C w \\
&= \sum_{\substack{C \text{ is a finite conjugacy} \\ \text{class of } G}} \underbrace{\sum_{w \in C} \beta_C w}_{= \beta_C \sum_{w \in C} w} + \sum_{\substack{C \text{ is an infinite conjugacy} \\ \text{class of } G}} \sum_{w \in C} \underbrace{\beta_C}_{\substack{=0 \\ \text{(by (74))}}} w \\
&= \sum_{\substack{C \text{ is a finite conjugacy} \\ \text{class of } G}} \beta_C \underbrace{\sum_{w \in C} w}_{\substack{= \sum_{c \in C} c = \mathbf{z}_C \\ \text{(by the definition} \\ \text{of } \mathbf{z}_C)}} + \underbrace{\sum_{\substack{C \text{ is an infinite conjugacy} \\ \text{class of } G}} \sum_{w \in C} 0w}_{=0} \\
&= \sum_{\substack{C \text{ is a finite conjugacy} \\ \text{class of } G}} \beta_C \mathbf{z}_C.
\end{aligned}$$

This is clearly a $\mathbf{k}$-linear combination of the conjugacy class sums $\mathbf{z}_C$. Hence, we have shown that $\sum_{w \in G} \alpha_w w$ is a $\mathbf{k}$-linear combination of the conjugacy class sums $\mathbf{z}_C$. In other words, $\mathbf{a}$ is a $\mathbf{k}$-linear combination of the conjugacy class sums $\mathbf{z}_C$ (since $\mathbf{a} = \sum_{w \in G} \alpha_w w$). This completes our proof of Claim 2. □

*Proof of Claim 3.* This follows easily from the fact that the conjugacy classes $C$ are nonempty and disjoint. Here are the details:

We must prove that the family $(\mathbf{z}_C)_{C \in \{\text{finite conjugacy classes}\}}$ is $\mathbf{k}$-linearly independent. In other words, we must prove that if we have an essentially finite family of scalars $(\beta_C)_{C \in \{\text{finite conjugacy classes}\}}$ that satisfies

$$\sum_{\substack{C \text{ is a finite} \\ \text{conjugacy class}}} \beta_C \mathbf{z}_C = 0, \tag{75}$$

then all these scalars $\beta_C$ are 0. So let us assume that we have an essentially finite family of scalars $(\beta_C)_{C \in \{\text{finite conjugacy classes}\}}$ that satisfies (75). We must prove that all these scalars $\beta_C$ are 0.

Let us extend the definition of $\beta_C$ to **all** conjugacy classes $C$ (not just the finite ones), by setting

$$\beta_C := 0 \qquad \text{for all infinite conjugacy classes } C. \tag{76}$$

For any $g \in G$, let $K(g)$ denote the conjugacy class of $G$ that contains $g$. Thus, if $g$ is an element of a conjugacy class $C$, then

$$C = K(g). \tag{77}$$

From (75), we obtain

$$0 = \sum_{\substack{C \text{ is a finite} \\ \text{conjugacy class}}} \beta_C \underbrace{\mathbf{z}_C}_{\substack{= \sum_{c \in C} c \\ \text{(by the definition of } \mathbf{z}_C\text{)}}} = \sum_{\substack{C \text{ is a finite} \\ \text{conjugacy class}}} \beta_C \sum_{c \in C} c$$

$$= \sum_{\substack{C \text{ is a finite} \\ \text{conjugacy class}}} \beta_C \sum_{g \in C} g = \sum_{\substack{C \text{ is a finite} \\ \text{conjugacy class}}} \sum_{g \in C} \beta_C g$$

$$= \sum_{\substack{C \text{ is a} \\ \text{conjugacy class}}} \sum_{g \in C} \underbrace{\beta_C}_{\substack{= \beta_{K(g)} \\ \text{(since } C = K(g) \\ \text{(by (77)))}}} g$$

$$\left( \begin{array}{c} \text{here, we have extended the range of the outer sum} \\ \text{from } \{\text{finite conjugacy classes}\} \text{ to } \{\text{all conjugacy classes}\}, \\ \text{which does not change the sum} \\ \text{because (76) shows that all the new addends are } 0 \end{array} \right)$$

$$= \underbrace{\sum_{\substack{C \text{ is a} \\ \text{conjugacy class}}} \sum_{g \in C}}_{\substack{= \sum_{g \in G} \\ \text{(since each } g \in G \\ \text{belongs to exactly one} \\ \text{conjugacy class)}}} \beta_{K(g)} g = \sum_{g \in G} \beta_{K(g)} g = \left( \beta_{K(g)} \right)_{g \in G}$$

(since $\sum_{g \in G} \alpha_g g = \left( \alpha_g \right)_{g \in G}$ for any essentially finite family $\left( \alpha_g \right)_{g \in G}$ of scalars). In other words,

$$\left( \beta_{K(g)} \right)_{g \in G} = 0 = (0)_{g \in G}.$$

In other words, $\beta_{K(g)} = 0$ for each $g \in G$.

Now, let $C$ be any finite conjugacy class. Then, $C$ is nonempty, so that there exists some $w \in C$. Consider this $w$. Then, $C = K(w)$ (by (77), applied to $g = w$). But we showed that $\beta_{K(g)} = 0$ for each $g \in G$. Applying this to $g = w$, we obtain $\beta_{K(w)} = 0$. In other words, $\beta_C = 0$ (since $C = K(w)$).

Forget that we fixed $C$. We thus have shown that $\beta_C = 0$ for all finite conjugacy classes $C$. In other words, all the scalars $\beta_C$ are 0. This completes the proof of Claim 3. □

Having proved all three claims, we are now done proving Theorem 3.9.5. □

**Exercise 3.9.1.** Let $G$ be a finite group. Prove the following:

**(a)** $\boxed{1}$ For each $\mathbf{a} \in \mathbf{k}[G]$, we have $\sum_{w \in G} w \mathbf{a} w^{-1} \in Z(\mathbf{k}[G])$.

**(b)** $\boxed{2}$ For each $\mathbf{a}, \mathbf{b} \in \mathbf{k}[G]$, we have $\sum_{w \in G} w \mathbf{a} \mathbf{b} w^{-1} = \sum_{w \in G} w \mathbf{b} \mathbf{a} w^{-1}$.

**Exercise 3.9.2.** Let $G$ be a group. For any $\mathbf{a} = \left(\alpha_g\right)_{g \in G} \in \mathbf{k}\left[G\right]$ and any element $h \in G$, we let $\left[h\right]\mathbf{a}$ denote the scalar $\alpha_h$ (that is, the $h$-th entry of the family $\mathbf{a} = \left(\alpha_g\right)_{g \in G}$); this is also known as the *$h$-coefficient* of $\mathbf{a}$.

**(a)** $\boxed{1}$ Prove that every $\mathbf{a}, \mathbf{b} \in \mathbf{k}\left[G\right]$ satisfy $\left[1\right]\left(\mathbf{ab}\right) = \left[1\right]\left(\mathbf{ba}\right)$ (even if $\mathbf{ab}$ and $\mathbf{ba}$ are not necessarily equal). Here, as usual, 1 denotes the neutral element of $G$.

**(b)** $\boxed{1}$ Let $h \in G$. Prove that the equality $\left[h\right]\left(\mathbf{ab}\right) = \left[h\right]\left(\mathbf{ba}\right)$ holds for all $\mathbf{a}, \mathbf{b} \in \mathbf{k}\left[G\right]$ if and only if $h \in Z\left(G\right)$ or the ring $\mathbf{k}$ is trivial (i.e., we have $0 = 1$ in $\mathbf{k}$).

### 3.9.2. Symmetric groups

So much for conjugacy in arbitrary groups. Now, let us find out what conjugacy means in the symmetric group $S_n$ and, more generally, in $S_X$ for $X$ finite.

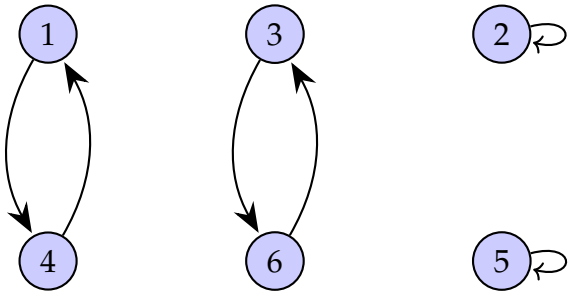
Recall that a *partition* (or, to be more precise, *integer partition*) means a weakly decreasing finite tuple of positive integers. For instance, $\left(5, 2, 2, 1\right)$ is a partition, whereas $\left(5, 1, 2\right)$ and $\left(5, 2, 0\right)$ are not. The *size* of a partition $\lambda = \left(\lambda_1, \lambda_2, \ldots, \lambda_k\right)$ is defined to be the sum $\lambda_1 + \lambda_2 + \cdots + \lambda_k \in \mathbb{N}$. A *partition of $n$* (for a given number $n \in \mathbb{Z}$) means a partition having size $n$. See [21s, Chapter 4] for more about partitions.

**Definition 3.9.6.** Let $X$ be a finite set. Let $w \in S_X$. Then, the *cycle type* of $w$ is the partition whose entries are the sizes of the orbits of $w$ (from largest to smallest). Of course, you can replace "sizes of the orbits" by "lengths of the cycles" in this definition without changing the outcome (because of Proposition 3.8.6).

We denote the cycle type of $w$ by $\operatorname{type} w$.

**Example 3.9.7.** Let $w = \operatorname{oln}\left(432651\right) \in S_6$. Then, $\operatorname{type} w = \left(3, 2, 1\right)$, since the orbits of $w$ are the sets $\left\{1, 4, 6\right\}$, $\left\{2, 3\right\}$ and $\left\{5\right\}$ with respective sizes 3, 2 and 1. (This $w$ is the permutation we analyzed in Example 3.8.2.)

**Example 3.9.8.** Let $w = \operatorname{oln}\left(426153\right) \in S_6$. Then, the cycle digraph of $w$ looks as follows:



Thus, the orbits of $w$ are the sets $\{1,4\}$, $\{2\}$, $\{3,6\}$, $\{5\}$ with respective sizes 2, 1, 2 and 1. Hence, the cycle type of $w$ is $\operatorname{type} w = (2,2,1,1)$ (don't forget to arrange the entries in decreasing order).

Note that the cycle type of a permutation $w \in S_X$ is always a partition of $|X|$.

Using cycle types, we can now characterize conjugacy in a symmetric group $S_X$ (for finite $X$): Two permutations $u, v \in S_X$ are conjugate if and only if they have the same cycle type. In other words:

**Theorem 3.9.9.** Let $X$ be a finite set. Let $u, v \in S_X$. Then, $u \sim v$ in $S_X$ if and only if $\operatorname{type} u = \operatorname{type} v$.

*Proof sketch.* $\Longrightarrow$: Assume that $u \sim v$ in $S_X$. Then, $u = pvp^{-1}$ for some $p \in S_X$ (by the definition of conjugacy). Consider this $p$. We must show that $\operatorname{type} u = \operatorname{type} v$.

Intuitively, this is pretty clear: The permutation $u = pvp^{-1}$ is obtained from $v$ by "relabelling" the elements $x$ of $X$ as $p(x)$, and it is obvious that the sizes of the orbits do not change under this "relabelling" operation. Thus, $\operatorname{type} u = \operatorname{type} v$.

To formalize this argument, we can argue as follows: Note that $p$ is a permutation, and thus an injective map; hence, any subset $U$ of $X$ satisfies

$$|p(U)| = |U| \tag{78}$$

(where $p(U)$ means the set $\{p(x) \mid x \in U\}$).

Now, if $O$ is an orbit of $v$, then $p(O)$ is an orbit of $pvp^{-1}$ (this is straightforward to check: just show that two elements $x, y \in X$ satisfy $x \overset{v}{\sim} y$ if and only if $p(x) \overset{pvp^{-1}}{\sim} p(y)$). Since $p$ is bijective, this construction yields all orbits of $pvp^{-1}$. Thus, the orbits of $pvp^{-1}$ are just the images of the orbits of $v$ under application of the map $p$. Since this application does not change the size of an orbit (by (78)), we thus conclude that the orbits of $pvp^{-1}$ have the same sizes as the orbits of $v$. In other words, $\operatorname{type}\left(pvp^{-1}\right) = \operatorname{type} v$. Since $pvp^{-1} = u$, we can rewrite this as $\operatorname{type} u = \operatorname{type} v$. This proves the "$\Longrightarrow$" direction of Theorem 3.9.9.

$\Longleftarrow$: Assume that $\operatorname{type} u = \operatorname{type} v$. Let us write the partition $\operatorname{type} u = \operatorname{type} v$ as $(n_1, n_2, \ldots, n_k)$. Thus, the cycle type of $u$ is $\operatorname{type} u = (n_1, n_2, \ldots, n_k)$. In other words, the cycles of $u$ have lengths $n_1, n_2, \ldots, n_k$ (since the cycle type of a permutation consists of the lengths of its cycles). Hence, the permutation $u$ has a DCD

$$u = \operatorname{cyc}_{a_{1,1}, a_{1,2}, \ldots, a_{1,n_1}} \circ \operatorname{cyc}_{a_{2,1}, a_{2,2}, \ldots, a_{2,n_2}} \circ \cdots \circ \operatorname{cyc}_{a_{k,1}, a_{k,2}, \ldots, a_{k,n_k}}$$

(since disjoint cycles commute and thus can be arranged in any order). Similarly, the permutation $v$ has a DCD

$$v = \operatorname{cyc}_{b_{1,1}, b_{1,2}, \ldots, b_{1,n_1}} \circ \operatorname{cyc}_{b_{2,1}, b_{2,2}, \ldots, b_{2,n_2}} \circ \cdots \circ \operatorname{cyc}_{b_{k,1}, b_{k,2}, \ldots, b_{k,n_k}}.$$

Now, let $p$ be the permutation in $S_X$ that sends each $a_{i,j}$ to the corresponding $b_{i,j}$. (This permutation exists because the definition of a DCD guarantees that each element of $X$ appears exactly once in the composite list

$$\begin{aligned}&(a_{1,1}, a_{1,2}, \ldots, a_{1,n_1},\\ &\ a_{2,1}, a_{2,2}, \ldots, a_{2,n_2},\\ &\ \ldots,\\ &\ a_{k,1}, a_{k,2}, \ldots, a_{k,n_k})\end{aligned}$$

and exactly once in the composite list

$$\begin{aligned}&(b_{1,1}, b_{1,2}, \ldots, b_{1,n_1},\\ &\ b_{2,1}, b_{2,2}, \ldots, b_{2,n_2},\\ &\ \ldots,\\ &\ b_{k,1}, b_{k,2}, \ldots, b_{k,n_k}).\end{aligned}$$

) Then, it is easy to see that $v = pup^{-1}$ (indeed, just check that both maps $v$ and $pup^{-1}$ send each $b_{i,j}$ to $b_{i,j+1}$, where $b_{i,n_i+1}$ is defined to be $b_{i,1}$ [16]). Thus, $v \sim u$, so that $u \sim v$. This proves the "$\Longleftarrow$" direction of Theorem 3.9.9. $\square$

**Corollary 3.9.10.** Let $X$ be a finite set.

**(a)** For each partition $\lambda$ of $|X|$, the set

$$C_\lambda(X) := \{w \in S_X \mid \operatorname{type} w = \lambda\}$$

(consisting of all permutations in $S_X$ having cycle type $\lambda$) is a conjugacy class of $S_X$.

**(b)** Each conjugacy class of $S_X$ has the form $C_\lambda(X)$ for some partition $\lambda$ of $|X|$.

**(c)** The # of conjugacy classes of $S_X$ equals the # of partitions of $|X|$.

[16] Here are the details: Set $b_{i,n_i+1} := b_{i,1}$ and $a_{i,n_i+1} := a_{i,1}$ for each $i \in [k]$. Then, for each $i \in [k]$ and $j \in [n_i]$, we have $u(a_{i,j}) = a_{i,j+1}$ (since $u = \operatorname{cyc}_{a_{1,1},a_{1,2},\ldots,a_{1,n_1}} \circ \operatorname{cyc}_{a_{2,1},a_{2,2},\ldots,a_{2,n_2}} \circ \cdots \circ \operatorname{cyc}_{a_{k,1},a_{k,2},\ldots,a_{k,n_k}}$) and $v(b_{i,j}) = b_{i,j+1}$ (for similar reasons) and $p(a_{i,j}) = b_{i,j}$ (by the definition of $p$) and $p(a_{i,j+1}) = b_{i,j+1}$ (by the definition of $p$, since $b_{i,n_i+1} = b_{i,1}$ and $a_{i,n_i+1} = a_{i,1}$), so that

$$\left(pup^{-1}\right)(b_{i,j}) = p\left(u\left(\underbrace{p^{-1}(b_{i,j})}_{\substack{=a_{i,j}\\ (\text{since } p(a_{i,j})=b_{i,j})}}\right)\right) = p\left(\underbrace{u(a_{i,j})}_{=a_{i,j+1}}\right) = p(a_{i,j+1}) = b_{i,j+1} = v(b_{i,j}).$$

Since each element of $X$ has the form $b_{i,j}$ for some $i$ and $j$, we thus have shown that $\left(pup^{-1}\right)(x) = v(x)$ for each $x \in X$. In other words, $pup^{-1} = v$, qed.

*Proof sketch.* **(a)** Let $\lambda$ be a partition of $|X|$. Write $\lambda$ as $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$, so that $\lambda_1 + \lambda_2 + \cdots + \lambda_k = (\text{size of } \lambda) = |X|$. Hence, we can subdivide the set $X$ into $k$ disjoint subsets

$$\begin{aligned} L_1 &= \left\{x_{1,1}, x_{1,2}, \ldots, x_{1,n_1}\right\}, \\ L_2 &= \left\{x_{2,1}, x_{2,2}, \ldots, x_{2,n_2}\right\}, \\ &\cdots, \\ L_k &= \left\{x_{k,1}, x_{k,2}, \ldots, x_{k,n_k}\right\} \end{aligned}$$

of sizes $n_1, n_2, \ldots, n_k$ (respectively). Obviously, neither these $k$ subsets nor the specific elements $x_{i,j}$ are uniquely determined, but we just make some choice and stick with it. Now, the permutation

$$u := \operatorname{cyc}_{x_{1,1}, x_{1,2}, \ldots, x_{1,n_1}} \circ \operatorname{cyc}_{x_{2,1}, x_{2,2}, \ldots, x_{2,n_2}} \circ \cdots \circ \operatorname{cyc}_{x_{k,1}, x_{k,2}, \ldots, x_{k,n_k}} \in S_X$$

has cycle type $\lambda$, and thus belongs to $C_\lambda(X)$. Hence, the set $C_\lambda(X)$ is nonempty.

Moreover, $\operatorname{type} u = \lambda$ (since $u$ has cycle type $\lambda$). The conjugacy class of the group $S_X$ that contains $u$ is

$$\begin{aligned} &\{w \in S_X \mid w \sim u\} \\ &= \{w \in S_X \mid \operatorname{type} w = \operatorname{type} u\} \\ &\qquad \left( \begin{array}{c} \text{since Theorem 3.9.9 shows that} \\ \text{we have } w \sim u \text{ if and only if } \operatorname{type} w = \operatorname{type} u \end{array} \right) \\ &= \{w \in S_X \mid \operatorname{type} w = \lambda\} \qquad (\text{since } \operatorname{type} u = \lambda) \\ &= C_\lambda(X) \qquad (\text{by the definition of } C_\lambda(X)). \end{aligned}$$

This shows that $C_\lambda(X)$ is a conjugacy class of $S_X$. This proves Corollary 3.9.10 **(a)**.

**(b)** Corollary 3.9.10 **(a)** shows that $C_\lambda(X)$ is a conjugacy class of $S_X$ whenever $\lambda$ is a partition of $|X|$. Clearly, these conjugacy classes $C_\lambda(X)$ for different partitions $\lambda$ are disjoint (since a permutation $w \in S_X$ cannot have two different cycle types at the same time) and thus distinct. Moreover, these classes $C_\lambda(X)$ cover the entire set $S_X$ (since each permutation $w \in S_X$ has cycle type $\lambda$ for some partition $\lambda$ of $|X|$, and thus belongs to the corresponding class $C_\lambda(X)$). Hence, they are all the conjugacy classes of $S_X$. In particular, each conjugacy class of $S_X$ has the form $C_\lambda(X)$ for some partition $\lambda$ of $|X|$. This proves Corollary 3.9.10 **(b)**.

**(c)** Combining parts **(a)** and **(b)** of Corollary 3.9.10, we see that the conjugacy classes of $S_X$ are precisely the sets $C_\lambda(X)$ for all partitions $\lambda$ of $|X|$. But the latter sets $C_\lambda(X)$ are distinct (as we saw in the above proof of part **(b)**), and thus their # equals the # of partitions of $|X|$. Hence, we conclude that the # of conjugacy classes of $S_X$ equals the # of partitions of $|X|$. This proves Corollary 3.9.10 **(c)**. □

For example:

- The partitions of 3 are $(1,1,1)$, $(2,1)$ and $(3)$. Hence, Corollary 3.9.10 **(b)** shows that the conjugacy classes of $S_3 = S_{[3]}$ are

$$\begin{aligned} C_{(1,1,1)}([3]) &= \{w \in S_3 \mid \operatorname{type} w = (1,1,1)\} = \{\mathrm{id}\}, \\ C_{(2,1)}([3]) &= \{w \in S_3 \mid \operatorname{type} w = (2,1)\} = \{t_{1,2},\ t_{1,3},\ t_{2,3}\}, \\ C_{(3)}([3]) &= \{w \in S_3 \mid \operatorname{type} w = (3)\} = \left\{\mathrm{cyc}_{1,2,3},\ \mathrm{cyc}_{1,3,2}\right\}. \end{aligned}$$

- The partitions of 4 are $(1,1,1,1)$, $(2,1,1)$, $(2,2)$, $(3,1)$ and $(4)$. Hence, Corollary 3.9.10 **(b)** shows that the conjugacy classes of $S_4 = S_{[4]}$ are

$$\begin{aligned} C_{(1,1,1,1)}([4]) &= \{w \in S_4 \mid \operatorname{type} w = (1,1,1,1)\} = \{\mathrm{id}\}, \\ C_{(2,1,1)}([4]) &= \{w \in S_4 \mid \operatorname{type} w = (2,1,1)\} = \left\{t_{i,j} \mid i < j\right\} \\ &= \{\text{transpositions in } S_4\}, \\ C_{(2,2)}([4]) &= \{w \in S_4 \mid \operatorname{type} w = (2,2)\} = \left\{t_{i,j}t_{u,v} \mid i,j,u,v \text{ distinct}\right\} \\ &= \{w \in S_4 \mid w \text{ is an involution and has no fixed points}\}, \\ C_{(3,1)}([4]) &= \{w \in S_4 \mid \operatorname{type} w = (3,1)\} = \left\{\mathrm{cyc}_{i,j,k} \mid i,j,k \text{ distinct}\right\}, \\ C_{(4)}([4]) &= \{w \in S_4 \mid \operatorname{type} w = (4)\} = \left\{\mathrm{cyc}_{i,j,k,\ell} \mid i,j,k,\ell \text{ distinct}\right\}. \end{aligned}$$

In general, for each $n \in \mathbb{N}$, Corollary 3.9.10 **(b)** shows that there is one conjugacy class of $S_n$ for each partition of $n$, and it consists of all permutations $w \in S_n$ whose cycle type is this partition. Thus, the center $Z(\mathbf{k}[S_n])$ is a free $\mathbf{k}$-module with a basis indexed by the partitions of $n$ (by Theorem 3.9.5).

The set of all transpositions in $S_n$ is precisely the conjugacy class

$$\left\{w \in S_n \mid \operatorname{type} w = \left(2, \underbrace{1,1,\ldots,1}_{n-2 \text{ ones}}\right)\right\}.$$

Thus, we obtain Proposition 3.9.2 as a particular case.

### 3.9.3. Centers and jucys–murphies

Having described the center of $\mathbf{k}[S_n]$ explicitly in terms of the conjugacy class sums, we can consider it well-understood, but some questions still remain: For instance, what is it **as an algebra** (i.e., how does its multiplication work)? Even the question "does a given element belong to the center?" is not completely straightforward if the coefficients of this element are not known explicitly. Thus, we have some more to say.

First, we observe something simple: The integral $\nabla$ and the sign-integral $\nabla^-$ belong to the center of $\mathbf{k}[S_n]$. This follows directly from Proposition 3.3.3, although it can also be derived from Lemma 3.9.3. The jucys–murphies $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n$ commute with each other, but (except for $\mathbf{m}_1 = 0$ and for $\mathbf{m}_2$ when $n = 2$) do not belong to the center of $\mathbf{k}[S_n]$. However, their sum $\mathbf{m}_1 + \mathbf{m}_2 + \cdots + \mathbf{m}_n$ does (as we saw in Proposition 3.9.2), and so do their elementary symmetric polynomials $e_k(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n)$ that we computed in Corollary 3.8.20 (since the number of orbits of a permutation $w$ depends only on the conjugacy class of $w$).

We can generalize this latter result somewhat. Recall that a polynomial $f \in \mathbf{k}[x_1, x_2, \ldots, x_n]$ in $n$ indeterminates $x_1, x_2, \ldots, x_n$ is said to be *symmetric* if each permutation $\sigma \in S_n$ satisfies

$$f\left(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}\right) = f.$$

The elementary symmetric polynomials $e_k(x_1, x_2, \ldots, x_n)$ are examples of symmetric polynomials, but there are others, such as the *power sum symmetric polynomials* $x_1^k + x_2^k + \cdots + x_n^k$ for all $k \in \mathbb{N}$. See [21s, Chapter 7] for an introduction to symmetric polynomials. Now we can give another characterization of $Z(\mathbf{k}[S_n])$ ([Murphy83, Theorem 1.9], [Moran92, Theorem 1'], [CSScTo10, Theorem 4.4.5]):

**Theorem 3.9.11** (Jucys–Murphy theorem)**.** The center $Z(\mathbf{k}[S_n])$ is the set of all symmetric polynomials in $n$ variables, applied to the jucys–murphies $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n$. That is,

$$Z(\mathbf{k}[S_n]) = \{f(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n) \mid f \in \mathbf{k}[x_1, x_2, \ldots, x_n] \text{ is symmetric}\}.$$

*Partial proof of Theorem 3.9.11 (sketched).* Let us show the "$\supseteq$" part (the "$\subseteq$" part is much harder). So we need to prove that

$$f(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n) \in Z(\mathbf{k}[S_n])$$

for any symmetric polynomial $f \in \mathbf{k}[x_1, x_2, \ldots, x_n]$.

A theorem of Gauss (known as the "fundamental theorem of symmetric polynomials")[17] says that each symmetric polynomial in $\mathbf{k}[x_1, x_2, \ldots, x_n]$ can be written as a polynomial in the elementary symmetric polynomials $e_1, e_2, \ldots, e_n$. For instance,

$$x_1^2 + x_2^2 + \cdots + x_n^2 = \left(\underbrace{x_1 + x_2 + \cdots + x_n}_{=e_1}\right)^2 - 2\underbrace{\sum_{i<j} x_i x_j}_{=e_2} = e_1^2 - 2e_2$$

[17] Proofs of this theorem can be found in [BluCos16, proof of Theorem 1], in [Dumas08, Theorem 1.2.1], in [Goodma16, Theorem 9.6.6], in [Smith95, §1.1] and in many other sources. (Some of these sources only state the result in the case when $\mathbf{k}$ is a field, or when $\mathbf{k} = \mathbb{C}$; but the same proof applies more generally for any commutative ring $\mathbf{k}$.)

and

$$\begin{aligned}
&x_1^3 + x_2^3 + \cdots + x_n^3 \\
&= \left( \underbrace{x_1 + x_2 + \cdots + x_n}_{=e_1} \right)^3 - 3 \underbrace{\sum_{i<j} \left( x_i^2 x_j + x_i x_j^2 \right)}_{=e_1 e_2 - 3e_3} - 6 \underbrace{\sum_{i<j<k} x_i x_j x_k}_{=e_3} \\
&= e_1^3 - 3\left(e_1 e_2 - 3e_3\right) - 6e_3 = e_1^3 - 3e_1 e_2 + 3e_3.
\end{aligned}$$

Thanks to this theorem, it suffices to show that

$$e_k\left(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n\right) \in Z\left(\mathbf{k}\left[S_n\right]\right) \qquad \text{for each } k \in \mathbb{N}$$

(since $Z\left(\mathbf{k}\left[S_n\right]\right)$ is a $\mathbf{k}$-subalgebra of $\mathbf{k}\left[S_n\right]$). But this follows from

$$e_k\left(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n\right) = \sum_{\substack{w \in S_n; \\ w \text{ has exactly } n-k \text{ orbits}}} w$$

(which is clearly preserved under conjugation). □

As a consequence, we have the following chain of subalgebras:

$$Z\left(\mathbf{k}\left[S_n\right]\right) \subseteq \mathrm{GZ}_n \subseteq \mathbf{k}\left[S_n\right], \tag{79}$$

where $\mathrm{GZ}_n$ is the Gelfand–Tsetlin subalgebra defined in Corollary 3.5.4.

## 3.10. A potpourri of other elements

In this section, I will showcase several other families of elements of $\mathbf{k}\left[S_n\right]$, stating a few properties of each. Some of these properties will be proved in later chapters.

### 3.10.1. Descent numbers and the Eulerian subalgebra

Descents are one of the simplest features of permutations:

**Definition 3.10.1.** Let $w \in S_n$ be a permutation.

**(a)** A *descent* of $w$ means a number $i \in [n-1]$ such that $w(i) > w(i+1)$.

**(b)** The *descent set* of $w$ means the set of all descents of $w$. It is called $\operatorname{Des} w$.

**(c)** The *descent number* of $w$ means the # of all descents of $w$. It is called $\operatorname{des} w$.

**Example 3.10.2.** If $w = \operatorname{oln}(516324) \in S_6$, then the descents of $w$ are $1, 3, 4$, so that $\operatorname{Des} w = \{1, 3, 4\}$ and $\operatorname{des} w = 3$.

See [22fco, Lecture 28, §4.6] and [21s, §A.4.11] for several properties of descents, and particularly for the *Eulerian numbers*, which count the permutations $w \in S_n$ with a given number of descents. We will instead take the sum of these permutations in $\mathbf{k}[S_n]$:

**Definition 3.10.3.** For any $k \in \mathbb{N}$, let us set

$$\mathbf{d}_k := \sum_{\substack{w \in S_n; \\ \operatorname{des} w = k}} w \in \mathbf{k}[S_n].$$

**Example 3.10.4.** For $n = 3$, we have

$$\begin{aligned}
&\mathbf{d}_0 = \mathrm{id} = 1; \\
&\mathbf{d}_1 = s_1 + s_2 + \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2}; \\
&\mathbf{d}_2 = t_{1,3}; \\
&\mathbf{d}_k = 0 \qquad \text{for all } k > 2.
\end{aligned}$$

More generally:

**Proposition 3.10.5.** **(a)** We have $\mathbf{d}_k = 0$ for all $k \geq n > 0$, since a permutation $w \in S_n$ has no more than $n - 1$ descents.

**(b)** We have $\mathbf{d}_0 = \mathrm{id} = 1$.

**(c)** Assume that $n > 0$. Then, $\mathbf{d}_{n-1} = w_0$, where $w_0 \in S_n$ is the permutation with OLN $(n, n-1, \ldots, 2, 1)$.

Much less obviously, the following holds:

**Theorem 3.10.6.** The $\mathbf{k}$-linear span of $\{\mathbf{d}_0, \mathbf{d}_1, \ldots, \mathbf{d}_{n-1}\}$ is a commutative $\mathbf{k}$-subalgebra of $\mathbf{k}[S_n]$. In other words, there exist coefficients $c_{i,j,k} \in \mathbf{k}$ (actually, $\in \mathbb{N}$) such that all $i, j \in \{0, 1, \ldots, n-1\}$ satisfy

$$\mathbf{d}_i \mathbf{d}_j = \mathbf{d}_j \mathbf{d}_i = c_{i,j,0} \mathbf{d}_0 + c_{i,j,1} \mathbf{d}_1 + \cdots + c_{i,j,n-1} \mathbf{d}_{n-1}.$$

This subalgebra is called the *Eulerian subalgebra* of $\mathbf{k}[S_n]$. The name harkens back to the *Eulerian numbers*, which count the permutations $w \in S_n$ having $k$ descents ([22fco, Lecture 28, Definition 4.6.3]).

**Question 3.10.7.** What are the coefficients $c_{i,j,k}$ in Theorem 3.10.6? Can they be computed explicitly?

Here is another curious property of descent-related sums ([VerTsi16, part of Corollary 2], [Rentel23, (1)]):

**Exercise 3.10.1.** $\boxed{10}$ Assume that $n \geq 1$. Let

$$\widetilde{\mathbf{d}} := \sum_{w \in S_n} \operatorname{des} w \cdot w = \sum_{k=0}^{n} k \mathbf{d}_k.$$

Prove that

$$\widetilde{\mathbf{d}}^2 = -(n-1)! \cdot \widetilde{\mathbf{d}} + \frac{1}{4}(n-1)!\,(n-1)\left(n^2 - n + 2\right)\nabla.$$

### 3.10.2. V-permutations

Let us next define a special type of permutations:

**Definition 3.10.8.** A permutation $w \in S_n$ is said to be a *V-permutation* if there exists some $k \in [n]$ such that

$$w(1) > w(2) > \cdots > w(k) < w(k+1) < w(k+2) < \cdots < w(n)$$

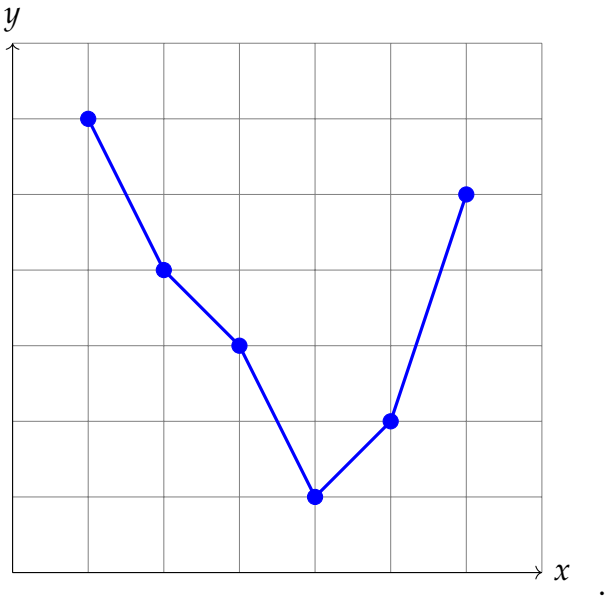
(that is, $w(i) > w(i+1)$ for each $i < k$, and $w(i) < w(i+1)$ for each $i \geq k$).

Note that the $k$ here is uniquely determined by $w$, and in fact satisfies $k = w^{-1}(1)$ (why?).

**Example 3.10.9.** The permutation $\operatorname{oln}(643125)$ is a V-permutation.

V-permutations have got their name from the fact that their "plot" (obtained by marking the points $(i, w(i))$ for all $i \in [n]$ in the Cartesian plane, and connecting adjacent points by straight lines) looks "V-shaped" (i.e., first decreasing, then increasing). For example, the "plot" of the permutation $\operatorname{oln}(643125)$ looks

as follows:




.

Some properties of V-permutations are discussed in [21s, §A.4.4]. We shall now encounter them when multiplying a certain product in $\mathbf{k}\left[S_n\right]$:

**Theorem 3.10.10.** **(a)** Let

$$\begin{aligned}\mathbf{V}_n &:= \left(1-\mathrm{cyc}_{2,1}\right)\left(1-\mathrm{cyc}_{3,2,1}\right)\left(1-\mathrm{cyc}_{4,3,2,1}\right)\cdots\left(1-\mathrm{cyc}_{n,n-1,\ldots,1}\right)\\ &= \prod_{i=2}^{n}\left(1-\mathrm{cyc}_{i,i-1,\ldots,1}\right) \in \mathbf{k}\left[S_n\right],\end{aligned}$$

where we agree to read a product $\prod_{i=2}^{n} a_i$ of noncommuting elements $a_i$ as $a_2 a_3 \cdots a_n$. Then,

$$\mathbf{V}_n = \sum_{\substack{w\in S_n \text{ is a}\\ \text{V-permutation}}} (-1)^{w^{-1}(1)-1}\, w.$$

**(b)** Moreover, if $n \geq 1$, then

$$\mathbf{V}_n^2 = n\mathbf{V}_n.$$

Hence, if $n$ is invertible in $\mathbf{k}$, then $\dfrac{1}{n}\mathbf{V}_n$ is an idempotent, known as the *Dynkin idempotent*.

Theorem 3.10.10 **(b)** is a result of Wever ([Wever47, page 571], see also [Reuten93, Theorem 8.16] and [Garsia90, Remark 2.1]).

**Remark 3.10.11.** As an illustration for the combinatorial nature of these results, let me show how Theorem 3.10.10 **(b)** can be restated as a purely combinatorial equality between two integers.

Indeed, set $\mathbf{k} = \mathbb{Z}$, and assume that $n \geq 1$. Then, Theorem 3.10.10 **(a)** yields

$$\mathbf{V}_n = \sum_{\substack{w \in S_n \text{ is a} \\ \text{V-permutation}}} (-1)^{w^{-1}(1)-1} w. \tag{80}$$

Multiplying this equality by itself, we obtain

$$\begin{aligned}
\mathbf{V}_n^2 &= \left( \sum_{\substack{w \in S_n \text{ is a} \\ \text{V-permutation}}} (-1)^{w^{-1}(1)-1} w \right) \left( \sum_{\substack{w \in S_n \text{ is a} \\ \text{V-permutation}}} (-1)^{w^{-1}(1)-1} w \right) \\
&= \left( \sum_{\substack{u \in S_n \text{ is a} \\ \text{V-permutation}}} (-1)^{u^{-1}(1)-1} u \right) \left( \sum_{\substack{v \in S_n \text{ is a} \\ \text{V-permutation}}} (-1)^{v^{-1}(1)-1} v \right) \\
&\qquad \left( \begin{array}{c} \text{here, we have renamed the indices } w \text{ and } w \\ \text{as } u \text{ and } v \text{ in the two sums} \end{array} \right) \\
&= \underbrace{\sum_{\substack{u \in S_n \text{ is a} \\ \text{V-permutation}}} \sum_{\substack{v \in S_n \text{ is a} \\ \text{V-permutation}}}}_{= \sum_{\substack{u,v \in S_n \text{ are two} \\ \text{V-permutations}}}} \underbrace{(-1)^{u^{-1}(1)-1} u (-1)^{v^{-1}(1)-1} v}_{\substack{= (-1)^{\left(u^{-1}(1)-1\right)+\left(v^{-1}(1)-1\right)} uv \\ = (-1)^{u^{-1}(1)+v^{-1}(1)} uv \\ \left(\text{since } \left(u^{-1}(1)-1\right)+\left(v^{-1}(1)-1\right) \equiv u^{-1}(1)+v^{-1}(1) \bmod 2\right)}} \\
&= \sum_{\substack{u,v \in S_n \text{ are two} \\ \text{V-permutations}}} (-1)^{u^{-1}(1)+v^{-1}(1)} uv \\
&= \sum_{w \in S_n} \sum_{\substack{u,v \in S_n \text{ are two} \\ \text{V-permutations;} \\ uv = w}} (-1)^{u^{-1}(1)+v^{-1}(1)} w.
\end{aligned}$$

Meanwhile, (80) also yields

$$\begin{aligned}
n\mathbf{V}_n &= n \sum_{\substack{w \in S_n \text{ is a} \\ \text{V-permutation}}} (-1)^{w^{-1}(1)-1} w = \sum_{\substack{w \in S_n \text{ is a} \\ \text{V-permutation}}} n (-1)^{w^{-1}(1)-1} w \\
&= \sum_{w \in S_n} \begin{cases} n (-1)^{w^{-1}(1)-1}, & \text{if } w \text{ is a V-permutation;} \\ 0, & \text{if not} \end{cases} \cdot w.
\end{aligned}$$

Now, Theorem 3.10.10 **(b)** says that the left hand sides of these two equalities are equal. Equivalently, it thus says that their right hand sides are equal.

In other words, it says that the equality

$$\sum_{w\in S_n} \sum_{\substack{u,v\in S_n \text{ are two} \\ \text{V-permutations;} \\ uv=w}} (-1)^{u^{-1}(1)+v^{-1}(1)}\, w$$
$$= \sum_{w\in S_n} \begin{cases} n\,(-1)^{w^{-1}(1)-1}, & \text{if } w \text{ is a V-permutation;} \\ 0, & \text{if not} \end{cases} \cdot w$$

holds. By comparing coefficients (see Remark 3.1.9), this equality is in turn equivalent to the statement that

$$\sum_{\substack{u,v\in S_n \text{ are two} \\ \text{V-permutations;} \\ uv=w}} (-1)^{u^{-1}(1)+v^{-1}(1)} = \begin{cases} n\,(-1)^{w^{-1}(1)-1}, & \text{if } w \text{ is a V-permutation;} \\ 0, & \text{if not} \end{cases}$$

for each $w \in S_n$. So this latter statement (an equality between two integers, explicitly given as sums) is a combinatorial restatement of Theorem 3.10.10 **(b)**. Does this restatement make the theorem easier to prove? I don't know!

**Exercise 3.10.2.** [2] Prove Theorem 3.10.10 **(a)**.

**Exercise 3.10.3.** [1] Let $w_0 \in S_n$ be the permutation $\operatorname{oln}(n, n-1, \ldots, 2, 1)$, which sends each $i \in [n]$ to $n+1-i$. Prove that $\mathbf{V}_n = (-1)^{n-1}\, \mathbf{V}_n w_0$, where $\mathbf{V}_n$ is as in Theorem 3.10.10.

**Exercise 3.10.4.** [2] For each $k \in [n]$, let $\alpha_k \in S_n$ be the permutation that sends the numbers $1, 2, \ldots, k$ to $k, k-1, \ldots, 1$ (respectively) while leaving all remaining numbers $k+1, k+2, \ldots, n$ unchanged. Let $\mathbf{V}_n$ be as in Theorem 3.10.10. Prove that

$$\mathbf{V}_n = (1-\alpha_2)\,(1+\alpha_3)\,(1-\alpha_4)\,(1+\alpha_5)\cdots\left(1-(-1)^{n-1}\alpha_n\right)$$
$$= \prod_{i=2}^{n}\left(1-(-1)^i\alpha_i\right)$$

(where the product is understood as in Theorem 3.10.10).

**Exercise 3.10.5.** [15] Prove Theorem 3.10.10 **(b)**.
[**Remark:** A direct or combinatorial proof would be particularly welcome.]

### 3.10.3. Random-to-random shuffles

We shall next define another kind of peculiar elements in $\mathbf{k}[S_n]$. They rely on the combinatorial notion of *noninversions*:

**Definition 3.10.12.** Let $i \in \mathbb{N}$. Let $w \in S_n$ be a permutation. Then, $\operatorname{noninv}_i(w)$ shall denote the # of $i$-element subsets $I$ of $[n]$ on which $w$ is increasing (i.e., for which the restriction $w|_I : I \to [n]$ is increasing). In other words, we let $\operatorname{noninv}_i(w)$ denote the # of ways to pick $i$ numbers $j_1 < j_2 < \cdots < j_i$ in $[n]$ such that $w(j_1) < w(j_2) < \cdots < w(j_i)$. This number $\operatorname{noninv}_i(w)$ is called the *$i$-noninversion number* of $w$.

**Example 3.10.13.** Let $n = 6$ and $w = \operatorname{oln}(642135) \in S_6$. Then, $\operatorname{noninv}_0(w) = 1$ (since the only 0-element subset $I$ of $[n]$ is $\varnothing$, and our permutation $w$ is increasing on $\varnothing$ for trivial reasons) and $\operatorname{noninv}_1(w) = 6$ (since there are six 1-element subsets $I$ of $[n]$, and again $w$ is increasing on each of them for trivial reasons) and $\operatorname{noninv}_2(w) = 6$ (since the 2-element subsets $I$ of $[n]$ on which $w$ is increasing are $\{2,6\}$, $\{3,5\}$, $\{3,6\}$, $\{4,5\}$, $\{4,6\}$ and $\{5,6\}$) and $\operatorname{noninv}_3(w) = 2$ (since the 3-element subsets $I$ of $[n]$ on which $w$ is increasing are $\{3,5,6\}$ and $\{4,5,6\}$) and $\operatorname{noninv}_k(w) = 0$ for all $k > 3$.

We can now define a new family of elements of $\mathbf{k}[S_n]$:

**Definition 3.10.14.** For each $k \in \{0, 1, \ldots, n\}$, we define

$$\mathbf{R}_k := \sum_{w \in S_n} \operatorname{noninv}_{n-k}(w) \cdot w \in \mathbf{k}[S_n].$$

This element $\mathbf{R}_k$ is called the *$k$-random-to-random shuffle* (for reasons we will later explain).

For example, for $n = 3$, we have

$$\mathbf{R}_0 = \sum_{w \in S_3} \operatorname{noninv}_3(w) \cdot w = \mathrm{id} = 1;$$

$$\mathbf{R}_1 = \sum_{w \in S_3} \operatorname{noninv}_2(w) \cdot w = 3\,\mathrm{id} + 2s_1 + 2s_2 + \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2};$$

$$\mathbf{R}_2 = \sum_{w \in S_3} \underbrace{\operatorname{noninv}_1(w)}_{=3} \cdot w = 3 \sum_{w \in S_3} w = 3\nabla;$$

$$\mathbf{R}_3 = \sum_{w \in S_3} \underbrace{\operatorname{noninv}_0(w)}_{=1} \cdot w = \sum_{w \in S_3} w = \nabla.$$

Some of this obviously generalizes:

- We always have $\mathbf{R}_0 = 1$ (since the only permutation $w \in S_n$ with $\operatorname{noninv}_n w \neq 0$ is $\mathrm{id} = 1$).
- We have $\mathbf{R}_n = \nabla$ (since $\operatorname{noninv}_0 w = 1$ for each $w \in S_n$).
- If $n \geq 1$, then $\mathbf{R}_{n-1} = n\nabla$ (since $\operatorname{noninv}_1 w = n$ for each $w \in S_n$).

But the other elements $\mathbf{R}_k$ with $0 < k < n-1$ are interesting. The following fact was discovered by Reiner, Saliola and Welker around 2011 ([ReSaWe11, Theorem 1.1]), and still retains its mystery:

**Theorem 3.10.15.** The $n+1$ elements $\mathbf{R}_0, \mathbf{R}_1, \ldots, \mathbf{R}_n$ commute.

Three proofs of Theorem 3.10.15 are known by now: [ReSaWe11, §6.1] gives a terse and handwavy combinatorial argument; [Lafren19, Chapitre IV] a long and computational proof; [BCGS25, Theorem 4.1] a simpler but still fairly long proof (which applies to a more general setting, in which the symmetric group algebra $\mathbf{k}[S_n]$ is replaced by the Hecke algebra $\mathcal{H}_n$).

**Question 3.10.16.** Is there a proof of Theorem 3.10.15 that is as simple as the theorem would suggest?

Another recent theorem (by Dieker, Saliola and Lafrenière) says:

**Theorem 3.10.17.** Let $k \in \{0, 1, \ldots, n\}$. Then, there exists a polynomial $P \in \mathbb{Z}[x]$ that splits over $\mathbb{Z}$ (that is, that can be written in the form $P = \prod_{r \in S} (x - r)$ for some multiset $S$ of integers) such that $P(\mathbf{R}_k) = 0$.

For $k = 1$, we can be more explicit:

$$\prod_{i=0}^{n^2} (\mathbf{R}_1 - i) = 0.$$

The original proof (in [Lafren19], building upon [DieSal18]) takes about 100 pages. Recently, a shorter proof was found by Axelrod-Freed, Brauner, Chiang, Commins, Lang and the present author (see [AFBCCL24] for the $k = 1$ case, and [BCGS25, Proposition C.1] for all other $k$). Still, the proof is technical and uses a significant amount of representation theory.

### 3.10.4. Excedance and antiexcedance elements

Here is something even stranger: an open problem (albeit one that has not been advertised very widely). We again begin with a combinatorial definition:

**Definition 3.10.18.** Let $w \in S_n$ be a permutation. Then, we define the two numbers

$$\operatorname{exc} w := (\#\text{ of } i \in [n] \text{ such that } w(i) > i) \qquad \text{and}$$
$$\operatorname{anxc} w := (\#\text{ of } i \in [n] \text{ such that } w(i) < i).$$

These numbers $\operatorname{exc} w$ and $\operatorname{anxc} w$ are called the *excedance number* and the *antiexcedance number* of $w$, respectively.

**Example 3.10.19.** Let $w = \operatorname{oln}(3524176)$. Then, $\operatorname{exc} w = 3$ and $\operatorname{anxc} w = 3$.

**Definition 3.10.20.** For any $a, b \in \mathbb{N}$, define

$$\mathbf{X}_{a,b} := \sum_{\substack{w \in S_n; \\ \operatorname{exc} w = a; \\ \operatorname{anxc} w = b}} w \in \mathbf{k}[S_n].$$

**Example 3.10.21.** For $n = 3$, we have

$$\mathbf{X}_{0,0} = \mathrm{id} = 1; \qquad \mathbf{X}_{1,2} = \mathrm{cyc}_{1,3,2}; \qquad \mathbf{X}_{2,1} = \mathrm{cyc}_{1,2,3};$$
$$\mathbf{X}_{1,1} = t_{1,2} + t_{1,3} + t_{2,3};$$

and $\mathbf{X}_{a,b} = 0$ for all remaining choices of $a, b \in \mathbb{N}$.

**Conjecture 3.10.22.** These elements $\mathbf{X}_{a,b}$ for all $a, b \in \mathbb{N}$ commute (for fixed $n$). In other words, $\mathbf{X}_{a,b}\mathbf{X}_{c,d} = \mathbf{X}_{c,d}\mathbf{X}_{a,b}$ for all $a, b, c, d \in \mathbb{N}$.

This conjecture was inspired by a conversation with Theo Douvropoulos; I have checked it for all $n \leq 8$.

We can restate this conjecture explicitly (by comparing coefficients, just as in Remark 3.10.11): It is saying that for any permutation $w \in S_n$ and any numbers $a, b, c, d \in \mathbb{N}$, we have

$$\begin{aligned}
&(\#\text{ of pairs } (u, v) \in S_n \times S_n \text{ with } w = uv \\
&\qquad \text{and } \operatorname{exc} u = a \text{ and } \operatorname{anxc} u = b \text{ and } \operatorname{exc} v = c \text{ and } \operatorname{anxc} v = d) \\
&= (\#\text{ of pairs } (u, v) \in S_n \times S_n \text{ with } w = vu \\
&\qquad \text{and } \operatorname{exc} u = a \text{ and } \operatorname{anxc} u = b \text{ and } \operatorname{exc} v = c \text{ and } \operatorname{anxc} v = d).
\end{aligned}$$

(Indeed, this equality is what we obtain if we compare the coefficients of $w$ on both sides of the equality $\mathbf{X}_{a,b}\mathbf{X}_{c,d} = \mathbf{X}_{c,d}\mathbf{X}_{a,b}$.)

For a change, the elements $\mathbf{X}_{a,b}$ are **not** (in general) annihilated by polynomials that split over $\mathbb{Z}$.

## 3.11. The antipode and the sign-twist

Now we shall define some maps on $\mathbf{k}[S_n]$. Such maps can be used to transform the above-defined elements into others, thus doubling our repository of interesting elements. Moreover, if these maps are structure-preserving (e.g., $\mathbf{k}$-linear maps or $\mathbf{k}$-algebra homomorphisms), then they can be used to carry some of the properties of the elements along, hence saving us a lot of proof work.

### 3.11.1. Morphisms

We begin by recalling the concept of a $\mathbf{k}$-algebra automorphism, and related concepts:

**Definition 3.11.1. (a)** A $\mathbf{k}$-*algebra homomorphism* (for short: $\mathbf{k}$-*algebra morphism*) is a $\mathbf{k}$-linear map $f : A \to B$ between two $\mathbf{k}$-algebras $A$ and $B$ that respects the unity (that is, satisfies $f(1_A) = 1_B$) and respects multiplication (that is, satisfies $f(aa') = f(a) f(a')$ for all $a, a' \in A$). Equivalently, it is a map that is simultaneously a $\mathbf{k}$-module morphism and a ring morphism.

**(b)** A $\mathbf{k}$-*algebra endomorphism* is a $\mathbf{k}$-algebra morphism from a $\mathbf{k}$-algebra to itself.

**(c)** A $\mathbf{k}$-*algebra isomorphism* is a $\mathbf{k}$-algebra morphism that has an inverse, which is also a $\mathbf{k}$-algebra morphism. (Actually, the "which is also a $\mathbf{k}$-algebra morphism" requirement comes for free; see [23wa, Proposition 3.11.8] for a proof.)

**(d)** A $\mathbf{k}$-*algebra automorphism* is a $\mathbf{k}$-algebra isomorphism from a $\mathbf{k}$-algebra to itself.

A ubiquitous source of $\mathbf{k}$-algebra automorphisms is conjugation: If $u$ is an invertible element of some $\mathbf{k}$-algebra $A$, then the map

$$\begin{aligned} A &\to A, \\ a &\mapsto uau^{-1} \end{aligned}$$

is a $\mathbf{k}$-algebra automorphism of $A$, known as *conjugation by* $u$. (This is easy to check – do it if you have not seen it before!) Automorphisms of $A$ obtained in this way (i.e., obtained as conjugations by invertible elements $u \in A$) are called *inner automorphisms*.

Note that any element $g$ of a group $G$ is automatically invertible in its group algebra $\mathbf{k}[G]$ (or rather, to be precise: the corresponding standard basis vector $e_g$ is invertible in $\mathbf{k}[G]$), and thus gives rise to an inner automorphism of $\mathbf{k}[G]$. There are typically many invertible elements of $\mathbf{k}[G]$ besides the standard basis vectors[18], and thus many more inner automorphisms.

But even these are often not the only automorphisms of $\mathbf{k}[G]$. For $\mathbf{k}[S_n]$ in particular, we have an important automorphism that is (almost) never an inner automorphism:

### 3.11.2. The sign-twist

[18] The simplest example is probably $-1$, but you can easily find more interesting ones. For example, if $n \geq 3$, then the element $1 + t_{1,2} + \mathrm{cyc}_{1,2,3} - \mathrm{cyc}_{1,3,2} - t_{1,3} \in \mathbf{k}[S_n]$ is invertible, with inverse $1 - t_{1,2} - \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2} + t_{1,3}$.

**Definition 3.11.2.** The *sign-twist* is the $\mathbf{k}$-linear map

$$\begin{aligned} T_{\mathrm{sign}} : \mathbf{k}\left[S_n\right] &\to \mathbf{k}\left[S_n\right], \\ w &\mapsto (-1)^w w \qquad \text{for all } w \in S_n. \end{aligned}$$

This definition should be read as follows: "The *sign-twist* is the $\mathbf{k}$-linear map $T_{\mathrm{sign}} : \mathbf{k}\left[S_n\right] \to \mathbf{k}\left[S_n\right]$ that sends each standard basis vector $w$ (with $w \in S_n$) to $(-1)^w w$". This requirement uniquely determines this map, since a $\mathbf{k}$-linear map from $\mathbf{k}\left[S_n\right]$ is uniquely determined by its values on the standard basis vectors. Explicitly, it thus is given by

$$T_{\mathrm{sign}}\left(\sum_{w \in S_n} \alpha_w w\right) = \sum_{w \in S_n} \alpha_w (-1)^w w \qquad \text{for any scalars } \alpha_w \in \mathbf{k}.$$

**Example 3.11.3.** **(a)** For $n = 3$, we have

$$\begin{aligned} &T_{\mathrm{sign}}\left(2\,\mathrm{id} + 7 s_1 - 9\,\mathrm{cyc}_{1,2,3}\right) \\ &= 2 \cdot \underbrace{(-1)^{\mathrm{id}}}_{=1} \mathrm{id} + 7 \cdot \underbrace{(-1)^{s_1}}_{=-1} s_1 - 9 \cdot \underbrace{(-1)^{\mathrm{cyc}_{1,2,3}}}_{=1} \mathrm{cyc}_{1,2,3} \\ &= 2\,\mathrm{id} - 7 s_1 - 9\,\mathrm{cyc}_{1,2,3}. \end{aligned}$$

**(b)** For each transposition $t_{i,j} \in S_n$, we have

$$T_{\mathrm{sign}}\left(t_{i,j}\right) = \underbrace{(-1)^{t_{i,j}}}_{=-1} t_{i,j} = -t_{i,j}. \tag{81}$$

Thus, for any $k \in [n]$, the jucys–murphy $\mathbf{m}_k$ satisfies

$$T_{\mathrm{sign}}\left(\mathbf{m}_k\right) = -\mathbf{m}_k$$

(since $\mathbf{m}_k$ is a sum of transpositions $t_{i,j}$).

**(c)** We have

$$T_{\mathrm{sign}}\left(\nabla\right) = \nabla^{-}. \tag{82}$$

More generally, for any subset $X$ of $[n]$, we have

$$T_{\mathrm{sign}}\left(\nabla_X\right) = \nabla_X^{-}. \tag{83}$$

This is easy to see directly from the definitions of $\nabla_X$ and $\nabla_X^{-}$. For example,

to prove (83), we can argue that every subset $X$ of $[n]$ satisfies

$$\begin{aligned} T_{\mathrm{sign}}\left(\nabla_X\right) &= T_{\mathrm{sign}}\left(\sum_{\substack{w\in S_n;\\ w(i)=i \text{ for all } i\notin X}} w\right) && \left(\text{since } \nabla_X = \sum_{\substack{w\in S_n;\\ w(i)=i \text{ for all } i\notin X}} w\right) \\ &= \sum_{\substack{w\in S_n;\\ w(i)=i \text{ for all } i\notin X}} \underbrace{T_{\mathrm{sign}}\left(w\right)}_{\substack{=(-1)^w w \\ \text{(by the definition of } T_{\mathrm{sign}}\text{)}}} && \left(\begin{array}{c}\text{since } T_{\mathrm{sign}} \text{ is}\\ \text{a } \mathbf{k}\text{-linear map}\end{array}\right) \\ &= \sum_{\substack{w\in S_n;\\ w(i)=i \text{ for all } i\notin X}} (-1)^w w = \nabla_X^- && \left(\text{by the definition of } \nabla_X^-\right). \end{aligned}$$

Thus, (83) is proved. Of course, (82) follows by applying (83) to $X = [n]$.

**(d)** For the somewhere-to-below shuffles $\mathbf{t}_k$ (defined in Definition 3.6.1), we have

$$\begin{aligned} T_{\mathrm{sign}}\left(\mathbf{t}_k\right) &= T_{\mathrm{sign}}\left(\mathrm{cyc}_k + \mathrm{cyc}_{k,k+1} + \mathrm{cyc}_{k,k+1,k+2} + \cdots + \mathrm{cyc}_{k,k+1,\ldots,n}\right) \\ &= \mathrm{cyc}_k - \mathrm{cyc}_{k,k+1} + \mathrm{cyc}_{k,k+1,k+2} \pm \cdots + (-1)^{n-k} \mathrm{cyc}_{k,k+1,\ldots,n} \end{aligned}$$

(an alternating sum, in which the signs alternate between $+$ and $-$). This follows easily from the formula $(-1)^{\mathrm{cyc}_{i_1,i_2,\ldots,i_k}} = (-1)^{k-1}$ for the sign of a cycle.

The sign-twist $T_{\mathrm{sign}}$ has some nice but simple properties. To state them, we need one more notion:

**Definition 3.11.4.** An *involution* means a map $f : X \to X$ that satisfies $f \circ f = \mathrm{id}$.

For example, all transpositions $t_{i,j} \in S_X$ for any set $X$ are involutions. But so is the sign-twist $T_{\mathrm{sign}}$, as we shall now show.

**Theorem 3.11.5. (a)** The map $T_{\mathrm{sign}} : \mathbf{k}\left[S_n\right] \to \mathbf{k}\left[S_n\right]$ is a $\mathbf{k}$-algebra automorphism.

**(b)** It is furthermore an involution (i.e., it satisfies $T_{\mathrm{sign}} \circ T_{\mathrm{sign}} = \mathrm{id}$).

*Proof of Theorem 3.11.5.* This is a fairly straightforward use of linearity. To illustrate this strategy, let me give a detailed proof. We begin with part **(b)**.

**(b)** We must show that $T_{\mathrm{sign}} \circ T_{\mathrm{sign}} = \mathrm{id}$. In other words, we must show that $T_{\mathrm{sign}}\left(T_{\mathrm{sign}}\left(\mathbf{a}\right)\right) = \mathbf{a}$ for any $\mathbf{a} \in \mathbf{k}\left[S_n\right]$.

So let $\mathbf{a} \in \mathbf{k}\left[S_n\right]$ be arbitrary. We must prove the equality $T_{\mathrm{sign}}\left(T_{\mathrm{sign}}\left(\mathbf{a}\right)\right) = \mathbf{a}$. This equality is $\mathbf{k}$-linear in $\mathbf{a}$ (since the map $T_{\mathrm{sign}}$ is $\mathbf{k}$-linear). Thus, we can

WLOG assume that $\mathbf{a}$ is a standard basis vector, i.e., has the form $\mathbf{a} = u$ for some $u \in S_n$. Assume this, and consider this $u$. From $\mathbf{a} = u$, we obtain $T_{\mathrm{sign}}(\mathbf{a}) = T_{\mathrm{sign}}(u) = (-1)^u u$ (by the definition of $T_{\mathrm{sign}}$, since $u \in S_n$). Applying the map $T_{\mathrm{sign}}$ to both sides of this equality, we find

$$\begin{aligned} T_{\mathrm{sign}}\left(T_{\mathrm{sign}}(\mathbf{a})\right) &= T_{\mathrm{sign}}\left((-1)^u u\right) \\ &= (-1)^u \underbrace{T_{\mathrm{sign}}(u)}_{=(-1)^u u} && \left(\text{since the map } T_{\mathrm{sign}} \text{ is } \mathbf{k}\text{-linear}\right) \\ &= \underbrace{(-1)^u (-1)^u}_{\substack{=\left((-1)^u\right)^2=1 \\ \text{(since } (-1)^u \text{ is either 1 or } -1, \\ \text{and thus squares to 1)}}} && u = u = \mathbf{a}. \end{aligned}$$

We thus have shown that $T_{\mathrm{sign}}\left(T_{\mathrm{sign}}(\mathbf{a})\right) = \mathbf{a}$. As explained above, this completes the proof of Theorem 3.11.5 **(b)**.

**(a)** The map $T_{\mathrm{sign}}$ is $\mathbf{k}$-linear (by definition) and respects the unity (since it sends the identity permutation $\mathrm{id} \in S_n$ to $T_{\mathrm{sign}}(\mathrm{id}) = \underbrace{(-1)^{\mathrm{id}}}_{=1} \mathrm{id} = \mathrm{id}$). We shall now show that it respects multiplication. In other words, we shall show that $T_{\mathrm{sign}}(\mathbf{ab}) = T_{\mathrm{sign}}(\mathbf{a}) \cdot T_{\mathrm{sign}}(\mathbf{b})$ for any $\mathbf{a}, \mathbf{b} \in \mathbf{k}[S_n]$.

So let $\mathbf{a}, \mathbf{b} \in \mathbf{k}[S_n]$ be arbitrary. We must prove the equality $T_{\mathrm{sign}}(\mathbf{ab}) = T_{\mathrm{sign}}(\mathbf{a}) \cdot T_{\mathrm{sign}}(\mathbf{b})$. This equality is $\mathbf{k}$-linear in $\mathbf{a}$ (since the map $T_{\mathrm{sign}}$ is $\mathbf{k}$-linear). Thus, we can WLOG assume that $\mathbf{a}$ is a standard basis vector (i.e., has the form $\mathbf{a} = u$ for some $u \in S_n$). Similarly, we can WLOG assume that $\mathbf{b}$ is a standard basis vector. Let us make both of these assumptions. Hence, $\mathbf{a} = u$ and $\mathbf{b} = v$ for some permutations $u, v \in S_n$. Consider these $u, v$. From $\mathbf{a} = u$, we obtain $T_{\mathrm{sign}}(\mathbf{a}) = T_{\mathrm{sign}}(u) = (-1)^u u$ (by the definition of $T_{\mathrm{sign}}$, since $u \in S_n$). Similarly, $T_{\mathrm{sign}}(\mathbf{b}) = (-1)^v v$.

One of the basic properties of signs shows that $(-1)^{uv} = (-1)^u (-1)^v$.

From $\mathbf{a} = u$ and $\mathbf{b} = v$, we obtain

$$\begin{aligned} T_{\mathrm{sign}}(\mathbf{ab}) &= T_{\mathrm{sign}}(uv) \\ &= \underbrace{(-1)^{uv}}_{=(-1)^u (-1)^v} uv && \left(\text{by the definition of } T_{\mathrm{sign}}, \text{ since } uv \in S_n\right) \\ &= (-1)^u (-1)^v uv = \underbrace{(-1)^u u}_{=T_{\mathrm{sign}}(\mathbf{a})} \cdot \underbrace{(-1)^v v}_{=T_{\mathrm{sign}}(\mathbf{b})} = T_{\mathrm{sign}}(\mathbf{a}) \cdot T_{\mathrm{sign}}(\mathbf{b}). \end{aligned}$$

Forget that we fixed $\mathbf{a}, \mathbf{b}$. We thus have shown that $T_{\mathrm{sign}}(\mathbf{ab}) = T_{\mathrm{sign}}(\mathbf{a}) \cdot T_{\mathrm{sign}}(\mathbf{b})$ for any $\mathbf{a}, \mathbf{b} \in \mathbf{k}[S_n]$. In other words, the map $T_{\mathrm{sign}}$ respects multiplication. Since we already know that $T_{\mathrm{sign}}$ is $\mathbf{k}$-linear and respects the unity, we conclude that $T_{\mathrm{sign}}$ is a $\mathbf{k}$-algebra morphism.

From Theorem 3.11.5 **(b)**, we know that $T_{\mathrm{sign}} \circ T_{\mathrm{sign}} = \mathrm{id}$. In other words, the map $T_{\mathrm{sign}}$ is inverse to itself. Thus, this map $T_{\mathrm{sign}}$ is invertible, and its inverse is $T_{\mathrm{sign}}$ itself. Since $T_{\mathrm{sign}}$ is a $\mathbf{k}$-algebra morphism, we thus conclude that $T_{\mathrm{sign}}$ is a $\mathbf{k}$-algebra isomorphism. Hence, $T_{\mathrm{sign}}$ is a $\mathbf{k}$-algebra automorphism (since it is a map from $\mathbf{k}[S_n]$ to itself). This proves Theorem 3.11.5 **(a)**. □

**Exercise 3.11.1.** $\boxed{1}$ Assume that $n \geq 2$ and $\mathbb{Z} \subseteq \mathbf{k}$. Prove that the sign-twist $T_{\mathrm{sign}}$ is not an inner automorphism of $\mathbf{k}[S_n]$ (that is, there exists no invertible $\mathbf{u} \in \mathbf{k}[S_n]$ such that all $\mathbf{a} \in \mathbf{k}[S_n]$ satisfy $T_{\mathrm{sign}}(\mathbf{a}) = \mathbf{u}\mathbf{a}\mathbf{u}^{-1}$).

We can use the sign-twist $T_{\mathrm{sign}}$ to turn identities in $\mathbf{k}[S_n]$ into new identities. For instance, by applying $T_{\mathrm{sign}}$ to Lemma 3.7.9, we obtain the following "negative version" of Lemma 3.7.9:

**Lemma 3.11.6.** Let $X$ be any subset of $[n]$, and let $x \in X$. Then,

$$\nabla_X^- = \nabla_{X\setminus\{x\}}^- \left(1 - \sum_{y \in X\setminus\{x\}} t_{y,x}\right).$$

*Proof.* Lemma 3.7.9 yields

$$\nabla_X = \nabla_{X\setminus\{x\}} \left(1 + \sum_{y \in X\setminus\{x\}} t_{y,x}\right).$$

Applying the map $T_{\mathrm{sign}}$ to both sides of this equality, we find

$$\begin{aligned}
T_{\mathrm{sign}}(\nabla_X) &= T_{\mathrm{sign}}\left(\nabla_{X\setminus\{x\}}\left(1 + \sum_{y \in X\setminus\{x\}} t_{y,x}\right)\right) \\
&= \underbrace{T_{\mathrm{sign}}\left(\nabla_{X\setminus\{x\}}\right)}_{\substack{=\nabla_{X\setminus\{x\}}^- \\ \text{(by (83), applied to } X\setminus\{x\} \\ \text{instead of } X\text{)}}} \left(1 + \sum_{y \in X\setminus\{x\}} \underbrace{T_{\mathrm{sign}}\left(t_{y,x}\right)}_{\substack{=-t_{y,x} \\ \text{(by (81))}}}\right) \\
&\qquad \left(\begin{array}{c}\text{since } T_{\mathrm{sign}} \text{ is a } \mathbf{k}\text{-algebra morphism} \\ \text{(by Theorem 3.11.5 } \mathbf{(a)}\text{)}\end{array}\right) \\
&= \nabla_{X\setminus\{x\}}^- \underbrace{\left(1 + \sum_{y \in X\setminus\{x\}} \left(-t_{y,x}\right)\right)}_{=1-\sum_{y \in X\setminus\{x\}} t_{y,x}} = \nabla_{X\setminus\{x\}}^- \left(1 - \sum_{y \in X\setminus\{x\}} t_{y,x}\right).
\end{aligned}$$

Comparing this with (83), we find

$$\nabla_X^- = \nabla_{X\setminus\{x\}}^- \left(1 - \sum_{y\in X\setminus\{x\}} t_{y,x}\right).$$

This proves Lemma 3.11.6. □

### 3.11.3. Anti-morphisms

There is yet another important map on $\mathbf{k}[S_n]$, called the *antipode*, which is not quite an automorphism but fairly close: It is an *algebra anti-automorphism*. Let us define this general concept:

**Definition 3.11.7. (a)** A $\mathbf{k}$-*algebra anti-homomorphism* (for short: $\mathbf{k}$-*algebra anti-morphism*) is a $\mathbf{k}$-linear map $f : A \to B$ between two $\mathbf{k}$-algebras $A$ and $B$ that respects the unity (that is, satisfies $f(1_A) = 1_B$) and reverses multiplication (that is, satisfies $f(aa') = f(a')\, f(a)$ for all $a, a' \in A$).

**(b)** A $\mathbf{k}$-*algebra anti-endomorphism* is a $\mathbf{k}$-algebra anti-morphism from a $\mathbf{k}$-algebra to itself.

**(c)** A $\mathbf{k}$-*algebra anti-isomorphism* is a $\mathbf{k}$-algebra anti-morphism that has an inverse, which is also a $\mathbf{k}$-algebra anti-morphism. (Actually, the "which is also a $\mathbf{k}$-algebra anti-morphism" requirement comes for free again.)

**(d)** A $\mathbf{k}$-*algebra anti-automorphism* is a $\mathbf{k}$-algebra anti-isomorphism from a $\mathbf{k}$-algebra to itself.

Anti-morphisms are not as frequently found as morphisms, but here is a very popular example:

**Example 3.11.8.** Let $n \in \mathbb{N}$, and consider the $\mathbf{k}$-algebra $\mathbf{k}^{n\times n}$ of all $n \times n$-matrices with entries in $\mathbf{k}$. Then, the map

$$\begin{aligned} \mathbf{k}^{n\times n} &\to \mathbf{k}^{n\times n}, \\ A &\mapsto A^T \end{aligned}$$

(which transposes every matrix) is a $\mathbf{k}$-algebra anti-automorphism of $\mathbf{k}^{n\times n}$ (and, by the way, an involution). In particular, it reverses multiplication, since $(AB)^T = B^T A^T$ for any $n \times n$-matrices $A$ and $B$.

**Remark 3.11.9.** If any one (or both) of two $\mathbf{k}$-algebras $A$ and $B$ is commutative, then there is no difference between "morphisms from $A$ to $B$" and "anti-morphisms from $A$ to $B$". Thus, the notion of an anti-morphism is interesting only in the noncommutative setting.

**Remark 3.11.10.** Anti-morphisms can be viewed as (regular) morphisms to a slightly different target. Indeed, if you are familiar with the "opposite ring" $B^{\mathrm{op}}$ of a ring $B$ (see, e.g., [23wa, Exercise 2.7.6]), then you can probably guess what the opposite $\mathbf{k}$-algebra of a $\mathbf{k}$-algebra is (essentially: it is the same $\mathbf{k}$-algebra, but the order of factors in a product is reversed, meaning that the product $b_1 b_2$ has been renamed as $b_2 b_1$). Now, a $\mathbf{k}$-algebra anti-morphism from $A$ to $B$ is nothing but a $\mathbf{k}$-algebra morphism from $A$ to the opposite $\mathbf{k}$-algebra of $B$.

**Remark 3.11.11.** Let $A$, $B$ and $C$ be three $\mathbf{k}$-algebras. If $f : B \to C$ and $g : A \to B$ are two $\mathbf{k}$-algebra anti-morphisms, then their composition $f \circ g : A \to C$ is a $\mathbf{k}$-algebra morphism (not anti-morphism). This might sound surprising, but is completely natural, since a product reversed twice reverts to its original order.

### 3.11.4. The antipode

We now define the antipode of $\mathbf{k}[S_n]$:

**Definition 3.11.12.** The *antipode* of $\mathbf{k}[S_n]$ is defined to be the $\mathbf{k}$-linear map

$$\begin{aligned} S : \mathbf{k}[S_n] &\to \mathbf{k}[S_n], \\ w &\mapsto w^{-1} \qquad \text{for all } w \in S_n. \end{aligned}$$

This is understood as in Definition 3.11.2. Thus, the map $S$ does **not** send every element of $\mathbf{k}[S_n]$ to its inverse, but only the standard basis vectors. Any other element of $\mathbf{k}[S_n]$ is written as a $\mathbf{k}$-linear combination of standard basis vectors, and then the standard basis vectors in this combination are inverted. Explicitly,

$$S\left(\sum_{w \in S_n} \alpha_w w\right) = \sum_{w \in S_n} \alpha_w w^{-1} \qquad \text{for any scalars } \alpha_w \in \mathbf{k}.$$

**Example 3.11.13.** **(a)** For $n = 3$, we have

$$\begin{aligned} S\left(2\,\mathrm{id} + 7 s_1 - 9\,\mathrm{cyc}_{1,2,3}\right) &= 2\underbrace{\mathrm{id}^{-1}}_{=\mathrm{id}} + 7\underbrace{s_1^{-1}}_{=s_1} - 9\underbrace{\left(\mathrm{cyc}_{1,2,3}\right)^{-1}}_{\substack{=\mathrm{cyc}_{1,3,2} \\ \text{(by (9))}}} \\ &= 2\,\mathrm{id} + 7 s_1 - 9\,\mathrm{cyc}_{1,3,2}. \end{aligned}$$

**(b)** For each transposition $t_{i,j} \in S_n$, we have

$$S\left(t_{i,j}\right) = t_{i,j}^{-1} = t_{i,j}.$$

Thus, for any $k \in [n]$, the jucys–murphy $\mathbf{m}_k$ satisfies

$$S(\mathbf{m}_k) = \mathbf{m}_k$$

(since $\mathbf{m}_k$ is a sum of transpositions $t_{i,j}$).

**(c)** We have

$$S(\nabla) = \nabla \qquad \text{and} \qquad S(\nabla^-) = \nabla^-.$$

Indeed, we have

$$\begin{aligned} S(\nabla^-) &= S\left(\sum_{w \in S_n} (-1)^w w\right) \qquad \left(\text{since } \nabla^- = \sum_{w \in S_n} (-1)^w w\right) \\ &= \sum_{w \in S_n} (-1)^w w^{-1} \qquad \text{(by the definition of } S\text{)} \\ &= \sum_{w \in S_n} \underbrace{(-1)^{w^{-1}}}_{=(-1)^w} \underbrace{\left(w^{-1}\right)^{-1}}_{=w} \qquad \left(\begin{array}{c}\text{here, we have substituted } w^{-1} \\ \text{for } w \text{ in the sum, because} \\ \text{the map } S_n \to S_n,\ w \mapsto w^{-1} \\ \text{is a bijection (since } S_n \text{ is a group)}\end{array}\right) \\ &= \sum_{w \in S_n} (-1)^w w = \nabla^- \end{aligned}$$

and similarly $S(\nabla) = \nabla$.

More generally, for any subset $X$ of $[n]$, we have

$$S(\nabla_X) = \nabla_X \qquad \text{and} \qquad S(\nabla_X^-) = \nabla_X^-.$$

(The proofs are analogous to the proofs of $S(\nabla) = \nabla$ and $S(\nabla^-) = \nabla^-$.)

**(d)** For the somewhere-to-below shuffles $\mathbf{t}_k$ (defined in Definition 3.6.1), we have

$$\begin{aligned} S(\mathbf{t}_k) &= S\left(\mathrm{cyc}_k + \mathrm{cyc}_{k,k+1} + \mathrm{cyc}_{k,k+1,k+2} + \cdots + \mathrm{cyc}_{k,k+1,\ldots,n}\right) \\ &= (\mathrm{cyc}_k)^{-1} + \left(\mathrm{cyc}_{k,k+1}\right)^{-1} + \left(\mathrm{cyc}_{k,k+1,k+2}\right)^{-1} + \cdots + \left(\mathrm{cyc}_{k,k+1,\ldots,n}\right)^{-1} \\ &= \mathrm{cyc}_k + \mathrm{cyc}_{k+1,k} + \mathrm{cyc}_{k+2,k+1,k} + \cdots + \mathrm{cyc}_{n,n-1,\ldots,k} \end{aligned}$$

(by (9)). I call these images $S(\mathbf{t}_k)$ the *below-to-somewhere shuffles*.

The map $S$ has almost the same nice properties as $T_{\mathrm{sign}}$ (see Theorem 3.11.5), with the exception that it is an anti-automorphism rather than an automorphism:

**Theorem 3.11.14. (a)** The map $S : \mathbf{k}[S_n] \to \mathbf{k}[S_n]$ is a $\mathbf{k}$-algebra anti-automorphism.

**(b)** It is furthermore an involution (i.e., it satisfies $S \circ S = \mathrm{id}$).

*Proof of Theorem 3.11.14.* This is analogous to Theorem 3.11.5. (Part **(b)** boils down to the fact that $\left(u^{-1}\right)^{-1} = u$ for each $u \in S_n$, whereas part **(a)** relies on the fact that $(uv)^{-1} = v^{-1}u^{-1}$ for any $u, v \in S_n$.) □

**Exercise 3.11.2.** [1] Recall the $\mathbf{R}_k$ from Definition 3.10.14. Prove that $S(\mathbf{R}_k) = \mathbf{R}_k$ for all $k \in \mathbb{N}$.

**Exercise 3.11.3.** [1] Recall the $\mathbf{X}_{a,b}$ from Definition 3.10.20. Prove that $S(\mathbf{X}_{a,b}) = \mathbf{X}_{b,a}$ for all $a, b \in \mathbb{N}$.

**Exercise 3.11.4.** [1] **(a)** Prove that the maps $S$ and $T_{\mathrm{sign}}$ commute, i.e., we have $S \circ T_{\mathrm{sign}} = T_{\mathrm{sign}} \circ S$.

**(b)** Prove that the map $S \circ T_{\mathrm{sign}}$ is an involution.

**Exercise 3.11.5.** [2] Prove that $S(\mathbf{a}) = \mathbf{a}$ for all $\mathbf{a} \in \mathrm{GZ}_n$ (where $\mathrm{GZ}_n$ is as in Corollary 3.5.4).

**Remark 3.11.15.** The antipode $S$ exists not just for $\mathbf{k}[S_n]$, but more generally for the group algebra $\mathbf{k}[G]$ of any group $G$. Indeed, Definition 3.11.12 extends verbatim to this general case (we just need to replace $S_n$ by $G$). Theorem 3.11.14 is still valid in this general case as well.

Even more generally, antipode-like maps exist for a wider class of algebraic structures known as *Hopf algebras* (see, e.g., [GriRei14, Chapter 1]). Their theory is actually the source of the name "antipode".

## 3.12. Other maps around group algebras

Next, we will introduce some very simple but useful maps from group algebras (and, more generally, monoid algebras).

### 3.12.1. The counit

We begin with a morphism from $\mathbf{k}[S_n]$ to the base ring $\mathbf{k}$:

**Definition 3.12.1.** The *counit* of $\mathbf{k}[S_n]$ is defined to be the $\mathbf{k}$-linear map

$$\begin{aligned} \varepsilon : \mathbf{k}[S_n] &\to \mathbf{k}, \\ w &\mapsto 1 \qquad \text{for all } w \in S_n. \end{aligned}$$

Thus, explicitly,

$$\varepsilon\left(\sum_{w\in S_n}\alpha_w w\right)=\sum_{w\in S_n}\alpha_w \qquad \text{for any scalars } \alpha_w\in\mathbf{k}.$$

**Example 3.12.2.** For any $k\in[n]$, the counit $\varepsilon$ sends the jucys–murphy $\mathbf{m}_k=\sum_{i=1}^{k-1}t_{i,k}$ to

$$\varepsilon\left(\mathbf{m}_k\right)=\varepsilon\left(\sum_{i=1}^{k-1}t_{i,k}\right)=\sum_{i=1}^{k-1}\underbrace{\varepsilon\left(t_{i,k}\right)}_{\substack{=1\\ \text{(since } t_{i,k}\in S_n\text{)}}} \qquad \text{(since } \varepsilon \text{ is } \mathbf{k}\text{-linear)}$$
$$=\sum_{i=1}^{k-1}1=k-1.$$

**Theorem 3.12.3.** The map $\varepsilon$ is a $\mathbf{k}$-algebra morphism.

*Proof.* This is analogous to Theorem 3.11.5 **(a)** but even easier. □

**Remark 3.12.4.** The counit $\varepsilon$ exists not just for $\mathbf{k}\left[S_n\right]$, but more generally for the monoid algebra $\mathbf{k}\left[M\right]$ of any monoid $M$. Indeed, Definition 3.12.1 extends verbatim to this general case (we just need to replace $S_n$ by $M$). Theorem 3.12.3 is still valid in this general case as well.

Simple as it is, the counit $\varepsilon$ can be quite useful in drawing combinatorial conclusions from identities in $\mathbf{k}\left[S_n\right]$. For example:

**Example 3.12.5.** Let $\mathbf{k}=\mathbb{Z}$ and $k\in\mathbb{N}$. We know from Corollary 3.8.20 that the $k$-th elementary symmetric polynomial $e_k$ applied to the jucys–murphies $\mathbf{m}_1,\mathbf{m}_2,\ldots,\mathbf{m}_n$ yields

$$e_k\left(\mathbf{m}_1,\mathbf{m}_2,\ldots,\mathbf{m}_n\right)=\sum_{\substack{w\in S_n;\\ w \text{ has exactly } n-k \text{ orbits}}} w.$$

Applying the counit $\varepsilon$ to both sides of this equality, we find

$$\varepsilon\left(e_k\left(\mathbf{m}_1,\mathbf{m}_2,\ldots,\mathbf{m}_n\right)\right)=\varepsilon\left(\sum_{\substack{w\in S_n;\\ w \text{ has exactly } n-k \text{ orbits}}} w\right).$$

Since $\varepsilon$ is a $\mathbf{k}$-algebra morphism, we can rewrite this as

$$e_k\left(\varepsilon\left(\mathbf{m}_1\right),\varepsilon\left(\mathbf{m}_2\right),\ldots,\varepsilon\left(\mathbf{m}_n\right)\right)=\sum_{\substack{w\in S_n;\\ w \text{ has exactly } n-k \text{ orbits}}}\varepsilon\left(w\right)$$

(here, we were able to rewrite $\varepsilon\left(e_k\left(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n\right)\right)$ as $e_k\left(\varepsilon\left(\mathbf{m}_1\right), \varepsilon\left(\mathbf{m}_2\right), \ldots, \varepsilon\left(\mathbf{m}_n\right)\right)$ because $\mathbf{k}$-algebra morphisms "commute" with polynomials). Since $\varepsilon\left(\mathbf{m}_i\right) = i - 1$ (by Example 3.12.2, applied to $i$ instead of $k$), this can be further rewritten as

$$\begin{aligned} & e_k\left(0, 1, \ldots, n-1\right) \\ &= \sum_{\substack{w \in S_n; \\ w \text{ has exactly } n-k \text{ orbits}}} \underbrace{\varepsilon\left(w\right)}_{\substack{=1 \\ \text{(by the definition of } \varepsilon\text{)}}} \\ &= \sum_{\substack{w \in S_n; \\ w \text{ has exactly } n-k \text{ orbits}}} 1 \\ &= \left(\#\text{ of permutations } w \in S_n \text{ that have exactly } n-k \text{ orbits}\right). \end{aligned}$$

Renaming $k$ as $n - i$, we can rewrite this as

$$e_{n-i}\left(0, 1, \ldots, n-1\right) = \left(\#\text{ of permutations } w \in S_n \text{ that have exactly } i \text{ orbits}\right).$$

This formula (which is true for each $i \in \{0, 1, \ldots, n\}$) is one of the simplest expressions for the # of permutations $w \in S_n$ that have exactly $i$ orbits. This # is known as the *unsigned Stirling number of the first kind* $c\left(n, i\right)$, and is commonly discussed in textbooks on enumerative combinatorics (see, e.g., [22fco, Lecture 27, §4.3.4] or [Stanle11, §1.3]).

**Exercise 3.12.1.** $\boxed{1}$ Let $\mathbf{a} \in \mathbf{k}\left[S_n\right]$. Prove that

$$\nabla \mathbf{a} = \mathbf{a} \nabla = \varepsilon\left(\mathbf{a}\right) \cdot \nabla.$$

For later use, let us also observe how the counit $\varepsilon$ (introduced in Definition 3.12.1) acts on an $X$-integral:

**Proposition 3.12.6.** Let $X$ be a subset of $[n]$. Then, $\varepsilon\left(\nabla_X\right) = |X|!$.

*Proof.* Define the set $S_{n,X}$ as in Proposition 3.7.4. Then, Proposition 3.7.4 **(a)** shows that $S_{n,X}$ is a subgroup of $S_n$, and is isomorphic to $S_X$. Hence, $|S_{n,X}| = |S_X|$ (since two isomorphic groups always have the same size). But we know that $|S_X| = |X|!$.

However, Proposition 3.7.4 **(b)** yields $\nabla_X = \sum_{w \in S_{n,X}} w$. Thus,

$$\varepsilon\left(\nabla_X\right) = \varepsilon\left(\sum_{w \in S_{n,X}} w\right) = \sum_{w \in S_{n,X}} \underbrace{\varepsilon\left(w\right)}_{\substack{=1 \\ \text{(by the definition of } \varepsilon\text{,} \\ \text{since } w \in S_{n,X} \subseteq S_n\text{)}}} \qquad \left(\begin{array}{c}\text{since the map } \varepsilon \\ \text{is } \mathbf{k}\text{-linear}\end{array}\right)$$

$$= \sum_{w \in S_{n,X}} 1 = \left|S_{n,X}\right| \cdot 1 = \left|S_{n,X}\right| = \left|S_X\right| = |X|!.$$

This proves Proposition 3.12.6. □

### 3.12.2. Morphisms of monoids yield morphisms of monoid algebras

Let us now explore some morphisms between different monoid algebras.

We begin with a simple construction in linear algebra:[19]

**Definition 3.12.7.** Let $M$ and $N$ be two sets, and let $f : M \to N$ be any map. Then,

$$f_* : \mathbf{k}^{(M)} \to \mathbf{k}^{(N)}$$

shall denote the $\mathbf{k}$-linear map that sends each standard basis vector $e_m \in \mathbf{k}^{(M)}$ to the standard basis vector $e_{f(m)} \in \mathbf{k}^{(N)}$. This map $f_*$ is called the *linearization* of $f$. Explicitly, it is given by

$$f_*\left(\sum_{m \in M} \alpha_m e_m\right) = \sum_{m \in M} \alpha_m e_{f(m)} \tag{84}$$

for any scalars $\alpha_m \in \mathbf{k}$.

**Proposition 3.12.8.** If the map $f$ in Definition 3.12.7 is injective, then so is $f_*$. The same holds if "injective" is replaced by "surjective" or "bijective".

*Proof.* Straightforward and left to the reader. □

**Proposition 3.12.9.** Let $M$ and $N$ be two monoids, and let $f : M \to N$ be a monoid morphism. Then the linearization $f_* : \mathbf{k}^{(M)} \to \mathbf{k}^{(N)}$ is a $\mathbf{k}$-algebra morphism from $\mathbf{k}\left[M\right]$ to $\mathbf{k}\left[N\right]$.

Using Convention 3.1.8, we can rewrite the above explicit formula (84) as

$$f_*\left(\sum_{m \in M} \alpha_m m\right) = \sum_{m \in M} \alpha_m f\left(m\right).$$

Thus, $f_*$ is the $\mathbf{k}$-linear map from $\mathbf{k}\left[M\right]$ to $\mathbf{k}\left[N\right]$ that sends each standard basis vector $m \in M$ to $f\left(m\right) \in N$.

[19] Recall the free $\mathbf{k}$-module $\mathbf{k}^{(S)}$ on a set $S$ (defined in Definition 3.1.1).

*Proof.* Straightforward and left to the reader. (Again, use linearity as in the proof of Theorem 3.11.5 **(a)**.) □

Thus, for any monoid morphism $f : M \to N$, we have constructed a $\mathbf{k}$-algebra morphism $f_* : \mathbf{k}[M] \to \mathbf{k}[N]$ between the respective monoid algebras. Not all $\mathbf{k}$-algebra morphisms between monoid algebras arise in this way, but the ones that do are particularly easy and well-behaved.

Monoid morphisms between groups are group morphisms. Here is a simple one between symmetric groups:

**Definition 3.12.10.** Let $X$ be a subset of $[n]$. Then, for each permutation $\sigma \in S_X$, we can define a permutation $\sigma \uparrow^n \in S_n$ by setting

$$(\sigma \uparrow^n)(i) = \begin{cases} \sigma(i), & \text{if } i \in X; \\ i, & \text{if } i \notin X \end{cases} \qquad \text{for each } i \in [n].$$

This permutation $\sigma \uparrow^n$ is called the *default extension* of $\sigma$ to $[n]$. In words, it can be described as "permuting the elements of $X$ like $\sigma$ does, and leaving the elements of $[n] \setminus X$ unchanged". (It is easy to check that it really is a permutation of $[n]$.)

We define the map

$$\begin{aligned} I_{X \to n} : S_X &\to S_n, \\ \sigma &\mapsto \sigma \uparrow^n, \end{aligned}$$

which sends each permutation $\sigma \in S_X$ to its default extension $\sigma \uparrow^n$. This map $I_{X \to n}$ will be called the *default embedding* of $S_X$ into $S_n$.

**Proposition 3.12.11.** Let $X$ be a subset of $[n]$. Then:

**(a)** The default embedding $I_{X \to n}$ is an injective group morphism.

**(b)** Its image is the subgroup

$$S_{n,X} = \{w \in S_n \mid w(i) = i \text{ for all } i \notin X\}$$

of $S_n$ (already defined in Proposition 3.7.4).

**(c)** The default embedding $I_{X \to n}$ sends cycles in $S_X$ to the "same" cycles in $S_n$, in the sense that it sends each cycle $\mathrm{cyc}_{i_1, i_2, \ldots, i_k} \in S_X$ to $\mathrm{cyc}_{i_1, i_2, \ldots, i_k} \in S_n$.

**(d)** The default embedding $I_{X \to n}$ preserves signs; i.e., we have

$$(-1)^\sigma = (-1)^{I_{X \to n}(\sigma)} \qquad \text{for any } \sigma \in S_X.$$

*Proof sketch.* **(a)** This is straightforward.

**(b)** This is also pretty easy. Indeed, it is clear that the image of $I_{X\to n}$ is a subset of $S_{n,X}$; it thus remains to show that all $w \in S_{n,X}$ belong to the image of $I_{X\to n}$. So let $w \in S_{n,X}$ be arbitrary. Then, Claim 1 in the proof of Proposition 3.7.4 **(a)** shows that $w(X) \subseteq X$ (so that the restriction $w\mid_X^X$ is really a well-defined map from $X$ to $X$) and that $w\mid_X^X \in S_X$. Now, it is easy to see that $w = I_{X\to n}\left(w\mid_X^X\right)$ (since $w \in S_{n,X}$ entails that $w(i) = i$ for all $i \notin X$, whereas the definition of $w\mid_X^X$ entails that $w(i) = \left(w\mid_X^X\right)(i)$ for all $i \in X$). Thus, $w$ belongs to the image of $I_{X\to n}$. This completes the proof of Proposition 3.12.11 **(b)**.

**(c)** Let $i_1, i_2, \ldots, i_k$ be $k$ distinct elements of $X$. By its definition, the cycle $\mathrm{cyc}_{i_1,i_2,\ldots,i_k} \in S_X$ sends the elements $i_1, i_2, \ldots, i_k$ to $i_2, i_3, \ldots, i_k, i_1$, while leaving all remaining elements of $X$ unchanged. Its default extension $\mathrm{cyc}_{i_1,i_2,\ldots,i_k}\uparrow^n$ thus does the same thing, but also leaves the elements of $[n]\setminus X$ unchanged. In other words, it sends the elements $i_1, i_2, \ldots, i_k$ to $i_2, i_3, \ldots, i_k, i_1$, while leaving all remaining elements of $[n]$ unchanged. But this is exactly what the cycle $\mathrm{cyc}_{i_1,i_2,\ldots,i_k} \in S_n$ does. Therefore,

$$\underbrace{\mathrm{cyc}_{i_1,i_2,\ldots,i_k}}_{\text{in } S_X}\uparrow^n = \underbrace{\mathrm{cyc}_{i_1,i_2,\ldots,i_k}}_{\text{in } S_n}.$$

Now, the definition of $I_{X\to n}$ yields

$$I_{X\to n}\left(\underbrace{\mathrm{cyc}_{i_1,i_2,\ldots,i_k}}_{\text{in } S_X}\right) = \underbrace{\mathrm{cyc}_{i_1,i_2,\ldots,i_k}}_{\text{in } S_X}\uparrow^n = \underbrace{\mathrm{cyc}_{i_1,i_2,\ldots,i_k}}_{\text{in } S_n}.$$

This proves Proposition 3.12.11 **(c)**.

**(d)** The easiest way to prove this is to recall the following two facts:

> *Claim 1:* Each permutation in $S_X$ can be written as a product of (finitely many) transpositions.
>
> *Claim 2:* If a permutation $\sigma \in S_X$ is a product of $k$ transpositions, then $(-1)^\sigma = (-1)^k$.

Claim 1 is [Grinbe15, Exercise 5.15 **(b)**], whereas Claim 2 is [Grinbe15, Exercise 5.15 **(c)**].

Now, let $\sigma \in S_X$. Then, Claim 1 shows that $\sigma$ can be written as a product of (finitely many) transpositions. In other words, there exists some transpositions $t_{i_1,j_1}, t_{i_2,j_2}, \ldots, t_{i_k,j_k} \in S_X$ such that

$$\sigma = t_{i_1,j_1} t_{i_2,j_2} \cdots t_{i_k,j_k}. \tag{85}$$

Consider these transpositions. Thus, $\sigma$ is a product of $k$ transpositions. Hence, Claim 2 yields $(-1)^\sigma = (-1)^k$. However, for any two distinct elements $u, v \in X$,

the map $I_{X\to n}$ sends the transposition $t_{u,v} \in S_X$ to the transposition $t_{u,v} \in S_n$ (since Proposition 3.12.11 **(c)** yields that $I_{X\to n}$ maps $\mathrm{cyc}_{u,v} \in S_X$ to $\mathrm{cyc}_{u,v} \in S_n$, but the transposition $t_{u,v}$ is just the cycle $\mathrm{cyc}_{u,v}$). In other words,

$$I_{X\to n}\left(t_{u,v}\right) = t_{u,v} \qquad \text{for any distinct } u, v \in X. \tag{86}$$

Now, applying the map $I_{X\to n}$ to the equality (85), we obtain

$$\begin{aligned} I_{X\to n}(\sigma) &= I_{X\to n}\left(t_{i_1,j_1} t_{i_2,j_2} \cdots t_{i_k,j_k}\right) \\ &= \underbrace{I_{X\to n}\left(t_{i_1,j_1}\right)}_{\substack{=t_{i_1,j_1} \\ \text{(by (86))}}} \underbrace{I_{X\to n}\left(t_{i_2,j_2}\right)}_{\substack{=t_{i_2,j_2} \\ \text{(by (86))}}} \cdots \underbrace{I_{X\to n}\left(t_{i_k,j_k}\right)}_{\substack{=t_{i_k,j_k} \\ \text{(by (86))}}} \\ &\qquad\qquad \left(\text{since } I_{X\to n} \text{ is a group morphism}\right) \\ &= t_{i_1,j_1} t_{i_2,j_2} \cdots t_{i_k,j_k}. \end{aligned}$$

This shows that $I_{X\to n}(\sigma)$ is a product of $k$ transpositions in $S_n$. Hence, we have $(-1)^{I_{X\to n}(\sigma)} = (-1)^k$ (by Claim 2, applied to $[n]$ instead of $X$). Comparing this with $(-1)^\sigma = (-1)^k$, we obtain $(-1)^\sigma = (-1)^{I_{X\to n}(\sigma)}$. This proves Proposition 3.12.11 **(d)**. □

**Remark 3.12.12.** Default embeddings facilitate a rather popular abuse of notation: Namely, if $X$ is a subset of $[n]$, then each permutation $\sigma \in S_X$ can be identified with its default extension $I_{X\to n}(\sigma) = \sigma \uparrow^n \in S_n$ (by abuse of notation). This identification is usually harmless, since the default embedding $I_{X\to n}$ is an injective group morphism from $S_X$ to $S_n$ (by Proposition 3.12.11 **(a)**) and thus respects multiplication (i.e., it does not matter if we compute a product $\sigma\tau$ of two permutations in $S_X$ or instead compute the product $I_{X\to n}(\sigma)\, I_{X\to n}(\tau)$ of their default extensions).[20]

As a consequence of this identification, the group $S_X$ becomes identified with its image under the morphism $I_{X\to n}$; this image is $S_{n,X}$ (by Proposition 3.12.11 **(b)**). We shall mostly avoid this identification in the following, but many authors use it, so it is worth knowing.

**Corollary 3.12.13.** Let $X$ be a subset of $[n]$. Then, the linearization of the default embedding $I_{X\to n} : S_X \to S_n$ is an injective $\mathbf{k}$-algebra morphism

$$\left(I_{X\to n}\right)_* : \mathbf{k}\left[S_X\right] \to \mathbf{k}\left[S_n\right].$$

[20] Some care needs to be taken with this identification. For example, the # of fixed points of a permutation $\sigma \in S_X$ is **not** the # of fixed points of its default extension $I_{X\to n}(\sigma)$ (unless $X = [n]$), since the latter has all the elements of $[n] \setminus X$ as additional fixed points. Thus, when counting (or listing) the fixed points, we must distinguish between $\sigma$ and $I_{X\to n}(\sigma)$. Similarly for descents and various other features.

*Proof.* Proposition 3.12.8 shows that $(I_{X\to n})_*$ is injective (since $I_{X\to n}$ is injective). Proposition 3.12.9 shows that $(I_{X\to n})_*$ is a $\mathbf{k}$-algebra morphism. □

**Definition 3.12.14.** We will refer to the injective $\mathbf{k}$-algebra morphism $(I_{X\to n})_*$ in Corollary 3.12.13 as a *default embedding of group algebras*.

**Example 3.12.15.** Let $X$ be a subset of $[n]$. Then, the $X$-integral $\nabla_X$ is

$$\begin{aligned}\nabla_X &= \sum_{w\in S_{n,X}} w \qquad \text{(by Proposition 3.7.4 \textbf{(b)})} \\ &= \sum_{u\in S_X} I_{X\to n}(u) \qquad \left(\begin{array}{c}\text{here, we have substituted } I_{X\to n}(u) \\ \text{for } w \text{ in the sum, since the} \\ \text{map } S_X \to S_{n,X},\ u \mapsto I_{X\to n}(u) \\ \text{is a bijection (by parts \textbf{(a)} and \textbf{(b)}} \\ \text{of Proposition 3.12.11)}\end{array}\right) \\ &= (I_{X\to n})_*\left(\sum_{u\in S_X} u\right). \end{aligned} \tag{87}$$

Similarly, using Proposition 3.12.11 **(d)**, we can show that

$$\nabla_X^- = (I_{X\to n})_*\left(\sum_{u\in S_X} (-1)^u u\right).$$

**Remark 3.12.16.** Let $X$ be a subset of $[n]$. If you are willing to identify each permutation $\sigma \in S_X$ with its default extension $I_{X\to n}(\sigma)$ (by the abuse of notation explained in Remark 3.12.12), then you can also identify each element $\mathbf{a} \in \mathbf{k}[S_X]$ with its image $(I_{X\to n})_*(\mathbf{a}) \in \mathbf{k}[S_n]$ under the default embedding of group algebras. Thus, $\mathbf{k}[S_X]$ becomes a $\mathbf{k}$-subalgebra of $\mathbf{k}[S_n]$, and the equality (87) can be rewritten as $\nabla_X = \sum_{u\in S_X} u$.

Default embeddings $I_{X\to n}$ are used particularly often when $X$ is itself a set of the form $[m] = \{1, 2, \ldots, m\}$ for some $m \leq n$:

**Definition 3.12.17.** Let $m$ be a nonnegative integer with $m \leq n$. Then, $[m] \subseteq [n]$. The default embedding $I_{[m]\to n} : S_{[m]} \to S_n$ will be called $I_{m\to n}$. This is an injective group morphism from $S_{[m]} = S_m$ to $S_n$.

Thus, for any nonnegative integer $m \leq n$, we have a default embedding

$$I_{m\to n} = I_{[m]\to n} : S_m \to S_n,$$

which is an injective group morphism, and thus (by Corollary 3.12.13) produces an injective $\mathbf{k}$-algebra morphism

$$(I_{m\to n})_* : \mathbf{k}[S_m] \to \mathbf{k}[S_n].$$

In terms of one-line notations, $I_{m\to n}$ is given by

$$I_{m\to n}\left(\operatorname{oln}\left(i_1,i_2,\ldots,i_m\right)\right)=\operatorname{oln}\left(i_1,i_2,\ldots,i_m,m+1,m+2,\ldots,n\right)$$

for every permutation $\operatorname{oln}\left(i_1,i_2,\ldots,i_m\right)\in S_m$.

Composing two of these default embeddings yields another:

**Proposition 3.12.18.** For any nonnegative integers $m\leq n\leq p$, we have

$$I_{n\to p}\circ I_{m\to n}=I_{m\to p}\qquad\text{and}\qquad\left(I_{n\to p}\right)_*\circ\left(I_{m\to n}\right)_*=\left(I_{m\to p}\right)_*.$$

*Proof.* Straightforward. □

**Remark 3.12.19.** Once again, things get easier if you are a bit sloppy. By Remark 3.12.12 (applied to $X=[m]$), we can identify each permutation $\sigma\in S_m$ with its default extension $I_{m\to n}\left(\sigma\right)=\sigma\uparrow^n\in S_n$ for any $m\leq n$. Likewise, by Remark 3.12.16, we can identify each element $\mathbf{a}\in\mathbf{k}\left[S_m\right]$ with its image $\left(I_{m\to n}\right)_*\left(\mathbf{a}\right)\in\mathbf{k}\left[S_n\right]$ under the default embedding of group algebras. This is an abuse of notation, but it is useful. It allows us to pretend that the default embeddings $I_{m\to n}$ and $\left(I_{m\to n}\right)_*$ are actual embeddings, i.e., that we have a chain of subgroups

$$S_0\subseteq S_1\subseteq S_2\subseteq S_3\subseteq\cdots\qquad\left(\text{note that the first }\subseteq\text{ sign is an }=\right)$$

and a chain of subalgebras

$$\mathbf{k}\left[S_0\right]\subseteq\mathbf{k}\left[S_1\right]\subseteq\mathbf{k}\left[S_2\right]\subseteq\mathbf{k}\left[S_3\right]\subseteq\cdots\qquad\left(\text{again, the first }\subseteq\text{ sign is an }=\right).$$

Keep in mind that this is just a figure of speech, and beware of the confusion that can arise when you compare elements of $\mathbf{k}\left[S_n\right]$ for different $n$'s. For example, the 3rd jucys–murphy $\mathbf{m}_3=t_{1,3}+t_{2,3}$ of $\mathbf{k}\left[S_3\right]$ agrees with the 3rd jucys–murphy $\mathbf{m}_3=t_{1,3}+t_{2,3}$ of $\mathbf{k}\left[S_4\right]$ (since the default embedding $I_{3\to 4}$ sends each $t_{i,j}\in S_3$ to $t_{i,j}\in S_4$), but the 3rd somewhere-to-below shuffle $\mathbf{t}_3=\operatorname{cyc}_3$ of $\mathbf{k}\left[S_3\right]$ does not agree with the 3rd somewhere-to-below shuffle $\mathbf{t}_3=\operatorname{cyc}_3+\operatorname{cyc}_{3,4}$ of $\mathbf{k}\left[S_4\right]$. When necessary, be explicit about what $n$ is.

Beware also that the center $Z\left(\mathbf{k}\left[S_2\right]\right)$ is not a subset of $Z\left(\mathbf{k}\left[S_3\right]\right)$, no matter how hard you abuse notation. (More generally, if $A$ is a subring of a ring $B$, then $Z\left(A\right)$ needs not be a subring of $Z\left(B\right)$.)

Default embeddings (and Proposition 3.12.11 **(d)** specifically) help us prove the following purely combinatorial lemma, which is a close relative of Proposition 3.7.13:

**Lemma 3.12.20.** Let $X_1, X_2, \ldots, X_k$ be $k$ disjoint subsets of $[n]$ such that $X_1 \cup X_2 \cup \cdots \cup X_k = [n]$.

Define a subset $K$ of $S_n$ by

$$K := \{w \in S_n \mid w(X_i) \subseteq X_i \text{ for all } i \in [k]\}.$$

For each permutation $w \in K$ and each $j \in [k]$, we define a map $w^{\langle j \rangle} : X_j \to X_j$ by setting

$$w^{\langle j \rangle}(x) = w(x) \qquad \text{for each } x \in X_j.$$

Then:

**(a)** This map $w^{\langle j \rangle}$ is well-defined and belongs to the symmetric group $S_{X_j}$ for each $w \in K$ and each $j \in [k]$.

**(b)** The map

$$K \to S_{X_1} \times S_{X_2} \times \cdots \times S_{X_k},$$
$$w \mapsto \left(w^{\langle 1 \rangle}, w^{\langle 2 \rangle}, \ldots, w^{\langle k \rangle}\right)$$

is a bijection.

**(c)** For each $w \in K$, we have

$$(-1)^w = (-1)^{w^{\langle 1 \rangle}} (-1)^{w^{\langle 2 \rangle}} \cdots (-1)^{w^{\langle k \rangle}}.$$

*Proof.* For any subset $X$ of $[n]$, define a subset $S_{n,X}$ of $S_n$ as in Proposition 3.7.4. If $X$ is any subset of $[n]$, then Proposition 3.7.4 **(a)** shows that the set $S_{n,X}$ is a subgroup of $S_n$, and is isomorphic to $S_X$ via the group isomorphism

$$S_{n,X} \to S_X,$$
$$w \mapsto w|_X^X$$

(where $w|_X^X$ means the restriction of $w$ to $X$, regarded as a map from $X$ to $X$). Let us denote this group isomorphism by $\varphi_X$.

Thus, in particular, we obtain a group isomorphism $\varphi_{X_j} : S_{n,X_j} \to S_{X_j}$ for each $j \in [k]$. The direct product of these $k$ isomorphisms $\varphi_{X_j}$ is a group isomorphism

$$\varphi_{X_1} \times \varphi_{X_2} \times \cdots \times \varphi_{X_k} : S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k} \to S_{X_1} \times S_{X_2} \times \cdots \times S_{X_k},$$
$$(w_1, w_2, \ldots, w_k) \mapsto (\varphi_{X_1}(w_1), \varphi_{X_2}(w_2), \ldots, \varphi_{X_k}(w_k)).$$

Let us denote this group isomorphism $\varphi_{X_1} \times \varphi_{X_2} \times \cdots \times \varphi_{X_k}$ by $\varphi_\Pi$. In particular, $\varphi_\Pi$ is a bijection (since $\varphi_\Pi$ is a group isomorphism).

For each $j \in [k]$ and each $w \in K$, we define a map $w^{(j)} : [n] \to [n]$ as in our proof of Proposition 3.7.13 **(b)**. (Note that our set $K$ is precisely the set $K$ that was defined in the latter proof.)

For each $j \in [k]$ and each $w \in K$, we define a map $w^{(j)} : [n] \to [n]$ by setting

$$w^{(j)}(x) = \begin{cases} w(x), & \text{if } x \in X_j; \\ x, & \text{if } x \notin X_j \end{cases} \qquad \text{for each } x \in [n]. \tag{88}$$

Claim 2 from the proof of Proposition 3.7.13 **(b)** shows that

$$w^{(j)} \in S_{n,X_j} \qquad \text{for each } j \in [k] \text{ and } w \in K. \tag{89}$$

We now claim the following:

> *Claim 1:* Let $j \in [k]$ and $w \in K$. Then, the map $w^{\langle j \rangle}$ is well-defined and equals $\varphi_{X_j}\left(w^{(j)}\right)$.

*Proof of Claim 1.* We have $w \in K$. In other words, $w \in S_n$ is a permutation with the property that $w(X_i) \subseteq X_i$ for all $i \in [k]$ (by the definition of $K$). Applying the latter property to $i = j$, we obtain $w\left(X_j\right) \subseteq X_j$. Hence, for each $x \in X_j$, we have $w\left(\underbrace{x}_{\in X_j}\right) \in w\left(X_j\right) \subseteq X_j$. Thus, the map $w^{\langle j \rangle} : X_j \to X_j$ is well-defined (since the definition of this map says that $w^{\langle j \rangle}(x) = w(x)$ for each $x \in X_j$). It remains to show that it equals $\varphi_{X_j}\left(w^{(j)}\right)$.

Indeed, we first observe that $w^{(j)} \in S_{n,X_j}$ (by (89)). Hence, $\varphi_{X_j}\left(w^{(j)}\right)$ is well-defined and is a map from $X_j$ to $X_j$. The definition of $\varphi_{X_j}$ yields $\varphi_{X_j}\left(w^{(j)}\right) = w^{(j)}\,|_{X_j}^{X_j}$ (where $w^{(j)}\,|_{X_j}^{X_j}$ means the restriction of $w^{(j)}$ to $X_j$, regarded as a map from $X_j$ to $X_j$). Hence, for each $x \in X_j$, we have

$$\underbrace{\left(\varphi_{X_j}\left(w^{(j)}\right)\right)}_{= w^{(j)}|_{X_j}^{X_j}}(x) = \left(w^{(j)}\,|_{X_j}^{X_j}\right)(x) = w^{(j)}(x)$$

$$= \begin{cases} w(x), & \text{if } x \in X_j; \\ x, & \text{if } x \notin X_j \end{cases} \qquad \left(\text{by the definition of } w^{(j)}\right)$$

$$= w(x) \qquad (\text{since } x \in X_j)$$

$$= w^{\langle j \rangle}(x) \qquad \left(\text{since the definition of } w^{\langle j \rangle} \text{ yields } w^{\langle j \rangle}(x) = w(x)\right).$$

Thus, $\varphi_{X_j}\left(w^{(j)}\right) = w^{\langle j \rangle}$. In other words, $w^{\langle j \rangle}$ equals $\varphi_{X_j}\left(w^{(j)}\right)$. Thus, the proof of Claim 1 is complete. □

**(a)** Let $w \in K$ and $j \in [k]$. We must prove that the map $w^{\langle j \rangle}$ is well-defined and belongs to the symmetric group $S_{X_j}$.

Claim 1 shows that the map $w^{\langle j \rangle}$ is well-defined and equals $\varphi_{X_j}\left(w^{(j)}\right)$. Hence, $w^{\langle j \rangle} = \varphi_{X_j}\left(w^{(j)}\right) \in S_{X_j}$ (since $\varphi_{X_j}$ is a map from $S_{n,X_j}$ to $S_{X_j}$). In other words, $w^{\langle j \rangle}$ belongs to the symmetric group $S_{X_j}$. The proof of Lemma 3.12.20 **(a)** is thus complete.

**(b)** In the proof of Proposition 3.7.13 **(b)**, we defined a map

$$\begin{aligned}\Phi : K &\to S_{n,X_1} \times S_{n,X_2} \times \cdots \times S_{n,X_k}, \\ w &\mapsto \left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right).\end{aligned}$$

Consider this map $\Phi$. In the proof of Proposition 3.7.13 **(b)**, we have shown that this map $\Phi$ is a bijection.

Each $j \in [k]$ and $w \in K$ satisfy

$$\varphi_{X_j}\left(w^{(j)}\right) = w^{\langle j \rangle} \tag{90}$$

(since Claim 1 shows that $w^{\langle j \rangle}$ equals $\varphi_{X_j}\left(w^{(j)}\right)$).

The map $\varphi_\Pi \circ \Phi : K \to S_{X_1} \times S_{X_2} \times \cdots \times S_{X_k}$ is a bijection (since it is the composition of the two bijections $\varphi_\Pi$ and $\Phi$). But each $w \in K$ satisfies

$$\begin{aligned}(\varphi_\Pi \circ \Phi)(w) &= \varphi_\Pi(\Phi(w)) = \varphi_\Pi\left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right) \\ &\qquad \left(\begin{array}{c}\text{since the definition of } \Phi \\ \text{yields } \Phi(w) = \left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right)\end{array}\right) \\ &= \left(\varphi_{X_1} \times \varphi_{X_2} \times \cdots \times \varphi_{X_k}\right)\left(w^{(1)}, w^{(2)}, \ldots, w^{(k)}\right) \\ &\qquad \left(\text{since } \varphi_\Pi \text{ was defined to be } \varphi_{X_1} \times \varphi_{X_2} \times \cdots \times \varphi_{X_k}\right) \\ &= \left(\varphi_{X_1}\left(w^{(1)}\right),\ \varphi_{X_2}\left(w^{(2)}\right),\ \ldots,\ \varphi_{X_k}\left(w^{(k)}\right)\right) \\ &\qquad \left(\text{by the definition of } \varphi_{X_1} \times \varphi_{X_2} \times \cdots \times \varphi_{X_k}\right) \\ &= \left(w^{\langle 1 \rangle}, w^{\langle 2 \rangle}, \ldots, w^{\langle k \rangle}\right) \\ &\qquad \left(\text{since (90) says that } \varphi_{X_j}\left(w^{(j)}\right) = w^{\langle j \rangle} \text{ for each } j \in [k]\right).\end{aligned}$$

Hence, the map $\varphi_\Pi \circ \Phi$ can be rewritten as the map

$$\begin{aligned}K &\to S_{X_1} \times S_{X_2} \times \cdots \times S_{X_k}, \\ w &\mapsto \left(w^{\langle 1 \rangle}, w^{\langle 2 \rangle}, \ldots, w^{\langle k \rangle}\right).\end{aligned}$$

Thus, the map

$$\begin{aligned}K &\to S_{X_1} \times S_{X_2} \times \cdots \times S_{X_k}, \\ w &\mapsto \left(w^{\langle 1 \rangle}, w^{\langle 2 \rangle}, \ldots, w^{\langle k \rangle}\right)\end{aligned}$$

is a bijection (since $\varphi_\Pi \circ \Phi$ is a bijection). This proves Lemma 3.12.20 **(b)**.

**(c)** Let $w \in K$. For each subset $X$ of $[n]$, recall the default embedding $I_{X \to n} : S_X \to S_n$ defined in Definition 3.12.10. We claim the following:

*Claim 2:* Let $j \in [k]$. Then, $I_{X_j \to n}\left(w^{\langle j \rangle}\right) = w^{(j)}$.

*Proof of Claim 2.* The definition of $I_{X_j \to n}$ yields $I_{X_j \to n}\left(w^{\langle j \rangle}\right) = w^{\langle j \rangle} \uparrow^n$. (See Definition 3.12.10 for the meaning of the "$\uparrow$" symbol.) The definition of $w^{\langle j \rangle} \uparrow^n$ (see Definition 3.12.10) shows that

$$\left(w^{\langle j \rangle} \uparrow^n\right)(i) = \begin{cases} w^{\langle j \rangle}(i), & \text{if } i \in X_j; \\ i, & \text{if } i \notin X_j \end{cases} \tag{91}$$

for each $i \in [n]$. Thus, for each $x \in [n]$, we have

$$\begin{aligned} \left(w^{\langle j \rangle} \uparrow^n\right)(x) &= \begin{cases} w^{\langle j \rangle}(x), & \text{if } x \in X_j; \\ x, & \text{if } x \notin X_j \end{cases} \qquad \text{(by (91), applied to } i = x\text{)} \\ &= \begin{cases} w(x), & \text{if } x \in X_j; \\ x, & \text{if } x \notin X_j \end{cases} \qquad \left( \begin{array}{c} \text{since } w^{\langle j \rangle}(x) = w(x) \text{ if } x \in X_j \\ \text{(by the definition of } w^{\langle j \rangle}\text{)} \end{array} \right) \\ &= w^{(j)}(x) \end{aligned}$$

(since the definition of $w^{(j)}$ says that $w^{(j)}(x) = \begin{cases} w(x), & \text{if } x \in X_j; \\ x, & \text{if } x \notin X_j \end{cases}$). Therefore, $w^{\langle j \rangle} \uparrow^n = w^{(j)}$. Thus, $I_{X_j \to n}\left(w^{\langle j \rangle}\right) = w^{\langle j \rangle} \uparrow^n = w^{(j)}$. This proves Claim 2. □

*Claim 3:* Let $j \in [k]$. Then, $(-1)^{w^{(j)}} = (-1)^{w^{\langle j \rangle}}$.

*Proof of Claim 3.* Proposition 3.12.11 **(d)** (applied to $X = X_j$) shows that the default embedding $I_{X_j \to n}$ preserves signs; i.e., that we have

$$(-1)^{\sigma} = (-1)^{I_{X_j \to n}(\sigma)} \qquad \text{for any } \sigma \in S_{X_j}.$$

Applying the latter equality to $\sigma = w^{\langle j \rangle}$, we obtain

$$(-1)^{w^{\langle j \rangle}} = (-1)^{I_{X_j \to n}\left(w^{\langle j \rangle}\right)} = (-1)^{w^{(j)}}$$

(since Claim 2 yields $I_{X_j \to n}\left(w^{\langle j \rangle}\right) = w^{(j)}$). This proves Claim 3. □

Now, Claim 3 from the proof of Proposition 3.7.13 **(b)** shows that $w = w^{(1)} w^{(2)} \cdots w^{(k)}$. Hence,

$$\begin{aligned} (-1)^w &= (-1)^{w^{(1)} w^{(2)} \cdots w^{(k)}} \\ &= \underbrace{(-1)^{w^{(1)}}}_{\substack{=(-1)^{w^{\langle 1 \rangle}} \\ \text{(by Claim 3)}}} \underbrace{(-1)^{w^{(2)}}}_{\substack{=(-1)^{w^{\langle 2 \rangle}} \\ \text{(by Claim 3)}}} \cdots \underbrace{(-1)^{w^{(k)}}}_{\substack{=(-1)^{w^{\langle k \rangle}} \\ \text{(by Claim 3)}}} \qquad \text{(by the multiplicativity of the sign)} \\ &= (-1)^{w^{\langle 1 \rangle}} (-1)^{w^{\langle 2 \rangle}} \cdots (-1)^{w^{\langle k \rangle}}. \end{aligned}$$

This proves Lemma 3.12.20 **(c)**. □

**Exercise 3.12.2.** Let $n$ be a positive integer. Consider the map

$$\begin{aligned} \rho_n : S_n &\to S_{n-1}, \\ w &\mapsto I_{n-1\to n}^{-1}\left(t_{n,w(n)} w\right) \end{aligned}$$

(where $t_{n,n}$ is understood to mean id).

**(a)** $\boxed{1}$ Show that this is well-defined (i.e., that $t_{n,w(n)} w$ really lies in the image of the default embedding $I_{n-1\to n}$ for each $w \in S_n$).

**(b)** $\boxed{1}$ Describe what $\rho_n$ does to the disjoint cycle decomposition of a permutation $w$.

**(c)** $\boxed{1}$ Show that $\rho_n$ is **not** a group morphism whenever $n \geq 3$.

**(d)** $\boxed{1}$ Show that nevertheless, we have $\rho_n(uv) = \rho_n(u)\rho_n(v)$ and $\rho_n(vu) = \rho_n(v)\rho_n(u)$ whenever $u \in S_n$ and $v \in I_{n-1\to n}(S_{n-1})$.

### 3.12.3. Base change

So far, we have constructed morphisms between monoid algebras over the same base ring $\mathbf{k}$. But there are also ring morphisms between monoid algebras over different base rings. The simplest ones are constructed as follows:

**Proposition 3.12.21.** Let $f : \mathbf{k} \to \mathbf{l}$ be a ring morphism between two commutative rings $\mathbf{k}$ and $\mathbf{l}$. Let $M$ be any monoid. Then, the map

$$\begin{aligned} f_* : \mathbf{k}[M] &\to \mathbf{l}[M], \\ (\alpha_m)_{m\in M} &\mapsto (f(\alpha_m))_{m\in M} \end{aligned}$$

(sending each $\sum_{m\in M} \alpha_m m$ to $\sum_{m\in M} f(\alpha_m) m$, where $\alpha_m \in \mathbf{k}$ are scalars) is a ring morphism. It is furthermore a $\mathbf{k}$-algebra morphism if we equip $\mathbf{l}[M]$ with the $\mathbf{k}$-algebra structure that comes from restriction of scalars[21] via $f : \mathbf{k} \to \mathbf{l}$.

*Proof sketch.* This is just a matter of juggling definitions around and using linearity, but let me outline the process for the sake of completeness:

It suffices to prove the second claim (since a $\mathbf{k}$-algebra morphism is automatically a ring morphism).

[21] A reminder from abstract algebra: *Restriction of scalars* is the easiest way to turn an $\mathbf{l}$-module or $\mathbf{l}$-algebra into an $\mathbf{k}$-module or $\mathbf{k}$-algebra (respectively) using a ring morphism $f : \mathbf{k} \to \mathbf{l}$. Namely, if $V$ is an $\mathbf{l}$-module, then we make $V$ into a $\mathbf{k}$-module by setting

$$\kappa v := f(\kappa) v \qquad \text{for all } \kappa \in \mathbf{k} \text{ and } v \in V.$$

(This defines a scaling; an addition is already defined on $V$.) By the same rule, we can make $V$ into a $\mathbf{k}$-algebra if $V$ is an $\mathbf{l}$-algebra. This is what we do here (for $V = \mathbf{l}[M]$).

First, it is easy to see that the map $f_*$ is $\mathbf{k}$-linear (since the map $f : \mathbf{k} \to \mathbf{l}$ itself is $\mathbf{k}$-linear if we equip $\mathbf{l}$ with the $\mathbf{k}$-algebra structure coming from restriction of scalars). Moreover, it is easy to see that $f_*$ sends each standard basis vector $e_m$ of $\mathbf{k}[M]$ to the corresponding standard basis vector $e_m$ of $\mathbf{l}[M]$: That is, we have

$$f_*(e_m) = e_m \qquad \text{for each } m \in M. \tag{92}$$

*Proof of (92).* Let $m \in M$. Recall that $f$ is a ring homomorphism, and thus satisfies $f(1_{\mathbf{k}}) = 1_{\mathbf{l}}$ and $f(0_{\mathbf{k}}) = 0_{\mathbf{l}}$.

By its definition, the map $f_*$ acts on a family $(\alpha_m)_{m \in M} \in \mathbf{k}[M]$ by applying $f$ to each entry of this family. But the standard basis vector $e_m$ of $\mathbf{k}[M]$ is defined to be the family whose $m$-th entry is $1_{\mathbf{k}}$ and whose all other entries are $0_{\mathbf{k}}$. Thus, after we apply $f_*$ to this vector $e_m$, the $m$-th entry becomes $f(1_{\mathbf{k}}) = 1_{\mathbf{l}}$, while all the other entries become $f(0_{\mathbf{k}}) = 0_{\mathbf{l}}$. Hence, we obtain the family whose $m$-th entry is $1_{\mathbf{l}}$ and whose all other entries are $0_{\mathbf{l}}$. But this family is precisely the standard basis vector $e_m$ of $\mathbf{l}[M]$. So we have shown that $f_*$ sends the standard basis vector $e_m$ of $\mathbf{k}[M]$ to the corresponding standard basis vector $e_m$ of $\mathbf{l}[M]$. This proves (92). □

In particular, (92) yields $f_*(e_1) = e_1$ (where 1 is the neutral element of $M$). In other words, $f_*\left(1_{\mathbf{k}[M]}\right) = 1_{\mathbf{l}[M]}$ (since $1_{\mathbf{k}[M]} = e_1$ and $1_{\mathbf{l}[M]} = e_1$). In other words, the map $f_*$ respects the unity.

Next, we shall prove that $f_*$ respects multiplication. In other words, we shall prove that $f_*(\mathbf{ab}) = f_*(\mathbf{a}) \cdot f_*(\mathbf{b})$ for any $\mathbf{a}, \mathbf{b} \in \mathbf{k}[M]$. Since both sides of this equality are $\mathbf{k}$-linear in $\mathbf{a}$ and $\mathbf{b}$, we can reduce this to the case when both $\mathbf{a}$ and $\mathbf{b}$ are standard basis vectors (just like we did in the proof of Theorem 3.11.5 **(a)**). In other words, it suffices to prove that $f_*(e_m e_n) = f_*(e_m) \cdot f_*(e_n)$ for any $m, n \in M$. But this is easy: For any $m, n \in M$, we have $e_m e_n = e_{mn}$ and thus

$$\begin{aligned} f_*(e_m e_n) &= f_*(e_{mn}) = e_{mn} \qquad \text{(by (92))} \\ &= \underbrace{e_m}_{\substack{= f_*(e_m) \\ \text{(by (92))}}} \cdot \underbrace{e_n}_{\substack{= f_*(e_n) \\ \text{(by (92))}}} = f_*(e_m) \cdot f_*(e_n), \end{aligned}$$

qed. So we have shown that the map $f_*$ respects multiplication. Since $f_*$ is $\mathbf{k}$-linear and respects the unity, we conclude that $f_*$ is a $\mathbf{k}$-algebra morphism, and therefore a ring morphism. Proposition 3.12.21 is thus proved. □

**Definition 3.12.22.** The ring morphism $f_*$ constructed in Proposition 3.12.21 is called the *base change morphism* for $M$ from $\mathbf{k}$ to $\mathbf{l}$.

**Example 3.12.23.** Let $M$ be the monoid $\left\{x^0, x^1, x^2, \ldots\right\}$ with multiplication given by $x^i x^j = x^{i+j}$ and with neutral element $x^0$. As we have seen in Subsection 3.1.6, its monoid algebra $\mathbf{k}[M]$ is just the polynomial ring $\mathbf{k}[x]$.

Let $f : \mathbb{C} \to \mathbb{C}$ be complex conjugation (i.e., the map $\mathbb{C} \to \mathbb{C}$, $z \mapsto \overline{z}$); this is a ring morphism.

Now, the base change morphism $f_* : \mathbb{C}[x] \to \mathbb{C}[x]$ defined in Proposition 3.12.21 is the automorphism of the polynomial ring $\mathbb{C}[x]$ that conjugates every coefficient of the polynomial.

In particular, we can apply Definition 3.12.22 to $M = S_n$, obtaining a base change morphism $f : \mathbf{k}[S_n] \to \mathbf{l}[S_n]$.

For any commutative ring $\mathbf{k}$, there is a unique ring morphism $g : \mathbb{Z} \to \mathbf{k}$ (sending $1_{\mathbb{Z}}$ to $1_{\mathbf{k}}$, and thus sending each integer $m \in \mathbb{Z}$ to $m \cdot 1_{\mathbf{k}}$). Thus, by Proposition 3.12.21, this ring morphism yields a ring morphism

$$g_* : \mathbb{Z}[S_n] \to \mathbf{k}[S_n],$$

which sends every $\sum_{w \in S_n} \alpha_w w$ (with $\alpha_w \in \mathbb{Z}$) to $\sum_{w \in S_n} (\alpha_w \cdot 1_{\mathbf{k}})\, w$. This morphism $g_*$ (and, more generally, any base change morphism for $S_n$) sends every standard basis vector $w \in S_n$ of $\mathbb{Z}[S_n]$ to the corresponding standard basis vector $w$ of $\mathbf{k}[S_n]$. Thus, it sends $\nabla$ to $\nabla$, sends each jucys–murphy $\mathbf{m}_k$ to $\mathbf{m}_k$, sends every $\mathbf{t}_k$ to $\mathbf{t}_k$, and so on.

This has a nice application: If you want to prove an equality in $\mathbf{k}[S_n]$, such as $\mathbf{m}_2\mathbf{m}_5 = \mathbf{m}_5\mathbf{m}_2$, then it suffices to prove this equality in $\mathbb{Z}[S_n]$ (that is, it suffices to prove it for $\mathbf{k} = \mathbb{Z}$), because you can then apply the ring morphism $g_*$ to both sides and obtain the corresponding equality in $\mathbf{k}[S_n]$. This is a useful strategy for proving equalities (usually more complicated ones than $\mathbf{m}_2\mathbf{m}_5 = \mathbf{m}_5\mathbf{m}_2$), since $\mathbb{Z}[S_n]$ is a better-behaved ring than the generic $\mathbf{k}[S_n]$. This strategy is known as *base change*, and the slogan for it is "working over $\mathbb{Z}$ is enough". (Of course, this holds only if the equality you are proving can be stated in $\mathbb{Z}$, meaning that it contains no fractions or irrationals or other non-integer scalars. But it can be easily adapted to other situations. For instance, if your equality involves fractions and is stated for $\mathbb{Q}$-algebras $\mathbf{k}$, then you only need to prove it in $\mathbb{Q}[S_n]$. If your equality involves an arbitrary (unspecified) element $x \in \mathbf{k}$ (for instance, this applies to Corollary 3.8.23), then you only need to prove it for $x = X$ in the ring $(\mathbb{Z}[X])[S_n]$ (where $\mathbb{Z}[X]$ is the usual polynomial ring in one indeterminate $X$ over $\mathbb{Z}$), because then you can apply $g_*$ where $g : \mathbb{Z}[X] \to \mathbf{k}$ is the unique ring morphism that sends $X$ to $x$.)

**Remark 3.12.24.** A particularly simple kind of ring morphisms are the *canonical inclusions* (i.e., the obvious maps from a subring to a ring that send each element of the subring to itself). If $\mathbf{k}$ is a subring of a commutative ring $\mathbf{l}$, then the canonical inclusion $g : \mathbf{k} \to \mathbf{l}$ is a ring morphism, and thus yields a base change morphism $g_* : \mathbf{k}[S_n] \to \mathbf{l}[S_n]$. This morphism $g_*$ sends each $\sum_{w \in S_n} \alpha_w w$ to $\sum_{w \in S_n} \alpha_w w$, and thus we can treat it as being a canonical inclusion itself, so that $\mathbf{k}[S_n]$ becomes a subring of $\mathbf{l}[S_n]$. In particular,

$$\mathbb{Z}[S_n] \subseteq \mathbb{Q}[S_n] \subseteq \mathbb{R}[S_n] \subseteq \mathbb{C}[S_n]$$

(and all these "$\subseteq$" signs are "subring" signs).

## 3.13. Antipodal conjugacy

Let us briefly return to the combinatorics of permutations:

**Theorem 3.13.1.** Each permutation $w \in S_n$ is conjugate to its inverse $w^{-1}$ in the group $S_n$.

*Proof.* Let $w \in S_n$ be a permutation. Then, Proposition 3.8.7 **(b)** shows that the orbits of $w^{-1}$ are the orbits of $w$. In other words, the permutations $w$ and $w^{-1}$ have the same orbits, and therefore the same cycle type (since the cycle type is defined in terms of the orbits). But any two permutations having the same cycle type are conjugate (by Theorem 3.9.9). Thus, the permutations $w$ and $w^{-1}$ are conjugate. This proves Theorem 3.13.1. □

Note that Theorem 3.13.1 is specific to symmetric groups; it is false (e.g.) for all cyclic groups of order $> 2$, and for many other groups.

**Exercise 3.13.1.** [2] Prove the following stronger version of Theorem 3.13.1: Let $w \in S_n$. Then, there exists an involution $x \in S_n$ such that $w = xw^{-1}x^{-1}$ (or, equivalently, $w = xw^{-1}x$).

Theorem 3.13.1 has a linear analogue, which is surprisingly nontrivial:

**Theorem 3.13.2.** Assume that $\mathbf{k}$ is a field of characteristic 0 (such as $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$). Let $\mathbf{a} \in \mathbf{k}[S_n]$. Then, $\mathbf{a}$ is conjugate to $S(\mathbf{a})$ in $\mathbf{k}[S_n]$ (meaning that there exists some invertible element $\mathbf{u} \in \mathbf{k}[S_n]$ such that $\mathbf{a} = \mathbf{u}S(\mathbf{a})\mathbf{u}^{-1}$).

I will prove this theorem later (in Section 5.20). Note that the "characteristic 0" requirement is really needed here.

The following corollary of Theorem 3.13.1 will be used later on:

**Proposition 3.13.3.** We have $S(\mathbf{a}) = \mathbf{a}$ for all $\mathbf{a} \in Z(\mathbf{k}[S_n])$.

*Proof.* Theorem 3.9.5 (applied to $G = S_n$) shows that the center $Z(\mathbf{k}[S_n])$ of the group algebra $\mathbf{k}[S_n]$ is the $\mathbf{k}$-linear span of the conjugacy class sums. In other words, the conjugacy class sums span $Z(\mathbf{k}[S_n])$.

Let $\mathbf{a} \in Z(\mathbf{k}[S_n])$. We must prove the equality $S(\mathbf{a}) = \mathbf{a}$. Both sides of this equality are $\mathbf{k}$-linear in $\mathbf{a}$. Hence, in proving this equality, we can WLOG assume that $\mathbf{a}$ is a conjugacy class sum (since the conjugacy class sums span $Z(\mathbf{k}[S_n])$). Assume this. Thus, $\mathbf{a} = \mathbf{z}_C$ for some conjugacy class $C$ of $S_n$ (where $\mathbf{z}_C$ is as defined in Definition 3.9.4 **(c)**). Consider this $C$. Thus, $\mathbf{z}_C = \sum_{c \in C} c$ (by the definition of $\mathbf{z}_C$).

However, Theorem 3.13.1 shows that each $c \in C$ satisfies $c^{-1} \in C$ [22]. Thus, the map

$$\begin{aligned} C &\to C, \\ c &\mapsto c^{-1} \end{aligned}$$

is well-defined. This map is clearly inverse to itself (since $\left(c^{-1}\right)^{-1} = c$ for each $c \in C$), and thus is invertible, i.e., a bijection.

Applying the map $S$ to the equality $\mathbf{z}_C = \sum_{c \in C} c$, we obtain

$$\begin{aligned} S\left(\mathbf{z}_C\right) &= S\left(\sum_{c \in C} c\right) = \sum_{c \in C} \underbrace{S\left(c\right)}_{\substack{=c^{-1} \\ \text{(by the definition of } S\text{)}}} && \text{(since the map } S \text{ is } \mathbf{k}\text{-linear)} \\ &= \sum_{c \in C} c^{-1} = \sum_{c \in C} c && \left(\begin{array}{c} \text{here, we have substituted } c \text{ for } c^{-1} \text{ in} \\ \text{the sum, since the map } C \to C,\ c \mapsto c^{-1} \\ \text{is a bijection} \end{array}\right) \\ &= \mathbf{z}_C. \end{aligned}$$

In other words, $S\left(\mathbf{a}\right) = \mathbf{a}$ (since $\mathbf{a} = \mathbf{z}_C$). Proposition 3.13.3 is thus proved. □

Another proof of Proposition 3.13.3 can be obtained by combining Exercise 3.11.5 with (79).[23]

## 3.14. Computer algebra

### 3.14.1. SageMath

A symmetric group algebra $\mathbf{k}\left[S_n\right]$ is a rather tangible object, perfectly suited for computation and experiments. However, computing inside $\mathbf{k}\left[S_n\right]$ by hand quickly gets tiresome as $n$ gets large, since $\left|S_n\right| = n!$ grows very fast. Thus, it is advantageous to use computer algebra when working in $\mathbf{k}\left[S_n\right]$.

Alas, most CASs (computer algebra systems) do not know $\mathbf{k}\left[S_n\right]$. You can work around this by representing elements of $\mathbf{k}\left[S_n\right]$ as matrices (we will see some ways to do this in the next two chapters), or you can write your own code for programmable CASs like Mathematica or Maple. In this section, I shall instead show how to use SageMath: one of the few CASs that "natively" know $\mathbf{k}\left[S_n\right]$.

[22] *Proof.* Let $c \in C$. Then, $c \in C \subseteq S_n$. Hence, Theorem 3.13.1 (applied to $w = c$) shows that the permutation $c$ is conjugate to its inverse $c^{-1}$ in the group $S_n$. In other words, $c$ and $c^{-1}$ belong to the same conjugacy class of $S_n$. This conjugacy class must be $C$ (since $c \in C$). Thus, we conclude that $c^{-1}$ belongs to $C$ as well. In other words, $c^{-1} \in C$, qed.

[23] Also, Theorem 3.13.2 yields Proposition 3.13.3 when $\mathbf{k}$ is a field of characteristic 0.

*SageMath* (for short: *Sage*) is a CAS built atop the Python programming language; you can view it as a Python dialect with a huge mathematical library attached to it (particularly well-suited for algebra and combinatorics). If you are familiar with Python, you will find SageMath fairly intuitive (although it redefines some Python syntax).[24]

You can install SageMath on your computer (it runs natively on Linux and Mac; on Windows it runs via the WSL subsystem). But the easiest way to use SageMath is through the Sage cell server:

`https://sagecell.sagemath.org/`

This is an anonymous cloud server: You type in your code and it runs the computation for you (as long as it takes no longer than 30 seconds).

### 3.14.2. Permutations in SageMath

Let us see how SageMath is used. First, we play around with permutations in the symmetric groups $S_n$; the group algebras will come later.

I highly recommend pasting (or typing) all the examples into the Sage cell server ( `https://sagecell.sagemath.org/` ) as you read this section.

**Entering permutations.** Permutations in Sage are usually expressed in OLN. For example, the permutation oln $(512364)$ can be entered as follows:[25]

```
w = Permutation([5, 1, 2, 3, 6, 4])
```

Sage can then compute values of `w` in the obvious way: Typing in `w(1)` will yield the output `5`, whereas typing in `w(2)` will yield `1`, and so on.

There are other ways to construct permutations in SageMath. For example, permutations can be entered in cycle notation (i.e., via their DCD) as follows:

```
u = Permutation(((1,3),(2,4,6),(5,)))
```

This is the permutation with DCD $((1,3),\ (2,4,6),\ (5))$, that is, the permutation $\mathrm{cyc}_{1,3}\,\mathrm{cyc}_{2,4,6}\,\mathrm{cyc}_5 \in S_6$. (The extraneous-looking comma in "`(5,)`" is a programming quirk: Without the comma, Sage would understand "`(5)`" as meaning the **number** 5, rather than the 1-tuple $(5)$ that we want to enter.)

You can skip length-1 cycles when entering a permutation in DCD, as long as Sage understands correctly what the $n$ is. For example, the above permutation `u` can also be entered as `Permutation(((1,3),(2,4,6)))`; Sage will automatically understand that 5 is meant to be a fixed point.[26]

[24] Another CAS that knows group rings is *GAP* (see `https://www.gap-system.org/` ).

[25] Everything printed in `typewritten font` in this section is either an input for Sage or an output of Sage.

[26] Of course, if you want $\mathrm{cyc}_{1,2} \in S_3$, then you need to write `Permutation(((1,2),(3,)))`, since `Permutation(((1,2),))` will yield $\mathrm{cyc}_{1,2} \in S_2$. Although this is not that bad – see the handling of default extensions below.

In case of doubt, always consult the documentation. For instance, the documentation for permutations in Sage explains all the above and much more. (On the cell server, you can also type `Permutation??` for a quick overview of the class-level documentation, although this is often not sufficient.)

I also recommend that you acquaint yourself with some of Sage/Python's syntactic quirks and pitfalls (e.g., its load-bearing whitespace/indentation[27]; its use of square brackets for lists[28]; its use of `=` for assignment and `==` for equality[29]).

**Composing permutations.** Let us now multiply (i.e., compose) the two permutations $u = \operatorname{oln}(312)$ and $v = \operatorname{oln}(132)$ in Sage. The symbol for multiplication in Sage is `*` (you cannot just write `uv` or `u v` and hope that Sage will understand it as a product); so we try the following:

```
u = Permutation([3,1,2])
v = Permutation([1,3,2])
u*v
```

We get `[2, 1, 3]`, which (like any permutation that Sage outputs) should be read in one-line notation, and thus means $\operatorname{oln}(213)$. But wait – the product $uv$ is actually $\operatorname{oln}(321)$, not $\operatorname{oln}(213)$. What has gone wrong?

The answer is that we got bitten by Sage's nonstandard notations. Sage defines the product of permutations in the **opposite** way from us; i.e., it defines a product $uv$ of two permutations to be "first $u$ then $v$" (so that $(uv)(i) = v(u(i))$ for each $i$). Thus, what we call $uv$ is rather called $vu$ (or, more precisely, `v*u`) in Sage. Thus, if you are using our convention for multiplying permutations, then you should write all products in reverse when you feed them into Sage. This is an annoying source of mistakes and anxiety which it is, alas, too late to fix.[30]

(If you don't want to write products from right to left, you can also get **our** product $uv$ by `u.left_action_product(v)` .)

**Other features of permutations.** Sage can do a lot more with permutations. It can take their inverses:

```
u.inverse()
```

It can compute their signs:

[27] See, e.g., `https://note.nkmk.me/en/python-indentation-usage/` .

This has the unfortunate consequence that if you copypaste my sample code from these notes, you will probably get the indentation wrong and you will have to fix it manually.

[28] So `[4, 2, 5, 7]` in Sage/Python is what we would call $(4,2,5,7)$.

[29] Thus, saying `a = 15` in Sage/Python means "let $a$ be 15", whereas saying `a == 15` means "$a$ equals 15".

[30] Sage's convention is in fact inspired by the GAP CAS, which was written by group theorists (the "G" in "GAP" stands for "groups"), who often write functions to the right of the elements they act on: i.e., they write $x.f$ instead of $f(x)$. From this point of view, the convention makes sense, because a product $uv$ of two permutations should send any $i$ to $(i.u).v$. But Sage writes $f(x)$, not $x.f$, so this product sends any $i$ to $v(u(i))$, which looks awful.

```
u.sign()
```

It can even multiply two permutations in different symmetric groups:

```
Permutation([2,1]) * Permutation([3,2,1])
```

This may appear baroque, but there is nothing wild going on here; Sage is simply replacing the first permutation $\operatorname{oln}(21)$ with its default extension $I_{2\to 3}(\operatorname{oln}(21)) = \operatorname{oln}(213)$ in order to get it into the same group as the second permutation. In other words, Sage is automatically using the default embeddings $I_{m\to n} : S_m \to S_n$ as "coercions" (i.e., replacing each $\sigma \in S_m$ by $I_{m\to n}(\sigma)$ whenever necessary).

**Listing permutations.** Now let's list all the 24 permutations of $[4]$:

```
Permutations(4)
```

Sage dutifully answers "`Standard permutations of 4`", which is correct but not very useful. So we need to be explicit and ask for a list if we want a list:

```
list(Permutations(4))
```

Now Sage indeed produces a list of all permutations $w \in S_4$, in lexicographically increasing order (i.e., sorted by $w(1)$, then by $w(2)$, then by $w(3)$, then by $w(4)$).

We can quickly access specific entries of this list using Python's "bracket" notation for the entries of a list:

```
list(Permutations(4))[5]
```

This gives us the 6-th permutation in the list – yes, the 6-th, not the 5-th, because Sage (like Python) numbers things starting at 0 rather than 1. Thus, the first permutation in the list can be obtained by

```
list(Permutations(4))[0]
```

There is also a shortcut to the last entry of a list:

```
list(Permutations(4))[-1]
```

Generally, if `a` is a list, then `a[-1]` means its last entry in Python and in Sage.

By the way, the last permutation `list(Permutations(4))[-1]` is $\operatorname{oln}(4321)$, and this is not a coincidence: The last permutation in $S_n$ (in lexicographically increasing order) is always the permutation $\operatorname{oln}(n, n-1, \ldots, 2, 1)$. In Sage, you can also construct it as follows:

```
Permutation(list(range(n, 0, -1)))
```

How does this work? Python's `list(range(a, b, c))` creates an arithmetic sequence starting at $a$ and increasing by a step size of $c$ until reaching or surpassing $b$ (but $b$ itself is not included). Thus, `list(range(n, 0, -1))` gives the sequence $(n, n-1, \ldots, 2, 1)$. Finally, by calling `Permutation` on this sequence, we transform it into a permutation.

Likewise, the identity permutation $\mathrm{id} = \operatorname{oln}(1, 2, \ldots, n) \in S_n$ can be obtained by

```
Permutation(list(range(1, n+1, 1)))
```

Here you can leave out the `1` at the end (i.e., you can replace `range(1, n+1, 1)` by `range(1, n+1)`), since 1 is the default value for the step size of an arithmetic sequence in Sage.

**Permutations with a given property.** Let us now list all involutions in $S_4$, that is, all permutations $w \in S_4$ that satisfy $w^{-1} = w$. Sage has no inbuilt function for this (I believe), but we can just use the definition:

```
[w for w in Permutations(4) if w.inverse() == w]
```

This syntax is fairly self-explanatory, once you recall that square brackets in Sage stand for lists (so that we are asking for a list), and that the "double equality sign" `==` just means regular equality (whereas the "single equality sign" `=` means assignment). Thus, the expression above means "the list of all $w$ where $w$ ranges over the permutations of $[4]$ that satisfy $w^{-1} = w$". And Sage responds with the actual list of these $w$'s.

If we want to know the number of these $w$'s (rather than a list of them), then we can ask Sage for the length of this list:

```
len([w for w in Permutations(4) if w.inverse() == w])
```

Or, even better (certainly faster!), we can let Sage count them:

```
sum(1 for w in Permutations(4) if w.inverse() == w)
```

Yes, we have just asked Sage for the sum $\sum_{\substack{w \in S_4; \\ w^{-1} = w}} 1$. This is the most natural way to count these $w$'s, and unlike the previous way, it does not make Sage waste its time by listing them first.

Similar syntax can be used for other classes of permutations. For instance, let's say you want to count the permutations $w \in S_7$ that satisfy $w(2) > w(3) > \cdots > w(7)$. You can do this via

```
sum(1 for w in Permutations(7) if w(2) > w(3)
                    > w(4) > w(5) > w(6) > w(7))
```

A more conceptual way would be

```
sum(1 for w in Permutations(7) if
                  all(w(i) > w(i+1) for i in range(2, 7)))
```

In fact, `all(w(i) > w(i+1) for i in range(2, 7))` is Sage's (and Python's) syntax for "$w(i) > w(i+1)$ holds for all $i \in \{2, 3, \ldots, 6\}$", which is the mathematically rigorous way of saying $w(2) > w(3) > \cdots > w(7)$.

### 3.14.3. Symmetric group algebras in SageMath

**Invoking the symmetric group algebra.** Now let us work in a symmetric group algebra. We will work in the $\mathbb{Q}$-algebra $\mathbb{Q}[S_n]$, since it includes $\mathbb{Z}[S_n]$ as a subring and contains almost everything we will ever need.

Let us start with $\mathbb{Q}[S_3]$. Here is how to generate this algebra in Sage:

```
A = SymmetricGroupAlgebra(QQ, 3)
```

Let us now compute the somewhere-to-below shuffle

$$\mathbf{t}_1 = \mathrm{cyc}_1 + \mathrm{cyc}_{1,2} + \mathrm{cyc}_{1,2,3} = \mathrm{oln}\,(123) + \mathrm{oln}\,(213) + \mathrm{oln}\,(231)$$

from Definition 3.6.1. We try the obvious thing:

```
Permutation([1,2,3]) + Permutation([2,1,3]) + Permutation([2,3,1])
```

Sadly, this does not work. Indeed, it cannot work: The three permutations we are trying to add are just permutations, but we want to add the corresponding standard basis vectors, not the permutations themselves. By abuse of notation, we might write the former while meaning the latter, but this does not make them literally equal, and Sage cannot be fooled. Nor does Sage have the context needed to understand our abuse correctly: Are we asking for the sum $\mathrm{oln}\,(123) + \mathrm{oln}\,(213) + \mathrm{oln}\,(231)$ in $\mathbb{Z}\left[S_3\right]$ or in $\mathbb{Q}\left[S_3\right]$ or in $\mathbf{k}\left[S_3\right]$ for some completely different ring $\mathbf{k}$ ?

So we need to ask Sage to add the standard basis vectors, not the permutations. The surefire way to do this is by wrapping the permutations in `A(...)`'s (remember that `A` is $\mathbb{Q}\left[S_3\right]$). So we write

```
A(Permutation([1,2,3])) + A(Permutation([2,1,3]))
              + A(Permutation([2,3,1]))
```

(this should be a single line in your Sage input; if you really want to break it into two, then put a backslash $\backslash$ at the end of the first line). Or, even simpler, we can just write

```
A([1,2,3]) + A([2,1,3]) + A([2,3,1])
```

Sage understands `A([...])` to mean `A(Permutation([...]))` .

**Computing in the symmetric group algebra.** Let us call this sum `t1` (as it is our $\mathbf{t}_1$):

```
A = SymmetricGroupAlgebra(QQ, 3)
t1 = A([1,2,3]) + A([2,1,3]) + A([2,3,1])
```

Let us similarly define `t2` to be our $\mathbf{t}_2 = \mathrm{cyc}_2 + \mathrm{cyc}_{2,3}$:

```
t2 = A([1,2,3]) + A([1,3,2])
```

Now let us ask Sage whether we have $\mathbf{t}_1\mathbf{t}_2 = \mathbf{t}_2\mathbf{t}_1$ (that is, whether $\mathbf{t}_1$ and $\mathbf{t}_2$ commute):

```
t1*t2 == t2*t1
```

Sage answers "`False`", so we see once again that $\mathbf{t}_1$ and $\mathbf{t}_2$ do not commute. Keep in mind that Sage's convention for multiplying elements in $\mathbf{k}\left[S_n\right]$ is opposite from ours (just like for multiplying permutations), but fortunately this does not matter for checking whether two elements commute.

We can also check that the commutator $\mathbf{t}_1\mathbf{t}_2 - \mathbf{t}_2\mathbf{t}_1$ is nilpotent:

```
(t1*t2 - t2*t1) ** 2
```

(Again, `t1*t2 - t2*t1` is our $\mathbf{t}_2\mathbf{t}_1 - \mathbf{t}_1\mathbf{t}_2$, not our $\mathbf{t}_1\mathbf{t}_2 - \mathbf{t}_2\mathbf{t}_1$, but this does not matter as we square the element.)

Sage knows quite a few more things about symmetric group algebras. For example, we can compute $S\left(t_1\right)$ by

```
t1.antipode()
```

And we can compute the jucys–murphy $\mathbf{m}_3$ by

```
A.jucys_murphy(3)
```

Sage can also compute inverses of invertible elements of $\mathbf{k}[S_n]$ (at least for $\mathbf{k} = \mathbb{Q}$):

```
(t1 + t2).inverse()
```

**A bit of programming.** Let's now be more systematic. We are usually interested in families of elements, not single elements, and it gets tiresome to manually define them one by one. Sage is a programming language (a dialect of Python), so we can "teach it how to fish". For instance, let us code a function that computes the somewhere-to-below shuffles

$$\mathbf{t}_k = \sum_{\ell=k}^{n} \mathrm{cyc}_{k,k+1,\ldots,\ell} \in \mathbf{k}[S_n]$$

for arbitrary $n \in \mathbb{N}$ and $k \in [n]$. First, we observe that the cycle $\mathrm{cyc}_{k,k+1,\ldots,\ell}$ can be written in OLN as follows:

$$\mathrm{cyc}_{k,k+1,\ldots,\ell} = \mathrm{oln}\left( \underbrace{1,2,\ldots,k-1}_{\texttt{list(range(1, k))}},\ \underbrace{k+1,k+2,\ldots,\ell}_{\texttt{list(range(k+1, l+1))}},k,\ \underbrace{\ell+1,\ell+2,\ldots,n}_{\texttt{list(range(l+1, n+1))}} \right),$$

and thus can be obtained in Sage as follows:

```
Permutation(list(range(1, k)) + list(range(k+1, l+1))
                              + [k] + list(range(l+1, n+1)))
```

(The "+" operator in Python/Sage does not only add numbers; it also concatenates lists: For instance, $[u,v,w] + [x,y] = [u,v,w,x,y]$.)

To obtain $\mathbf{t}_k$, we need to sum these cycles $\mathrm{cyc}_{k,k+1,\ldots,\ell}$ for all $\ell \in \{k, k+1, \ldots, n\}$ (or, as Sage would say, for all `l in range(k, n+1)`). Thus, we end up with the following code:

```
def QS(n): # this is the group algebra Q[S_n]
        return SymmetricGroupAlgebra(QQ, n)
def t(n, k): # this is our t_k in Q[S_n]
        A = QS(n)
        return A.sum(A(list(range(1, k)) + list(range(k+1, l+1))
                                + [k] + list(range(l+1, n+1)))
                            for l in range(k, n+1))
```

You can check this by letting Sage compute `t(3, 1)` or `t(5, 2)` (for example).

Let us play around a bit more with these elements. For instance, Exercise 3.6.1 for $n = 5$ claims that $\mathbf{t}_5\mathbf{t}_4\mathbf{t}_3\mathbf{t}_2\mathbf{t}_1 = \nabla$ in $\mathbf{k}[S_5]$. Let us check this with Sage (for $\mathbf{k} = \mathbb{Q}$):

```
LHS = QS(5).prod(t(5, k) for k in range(1, 6))
```

```
RHS = QS(5).sum(QS(5)(w) for w in Permutations(5))
LHS == RHS
```

The `LHS` here is saying "the product of `t(5, k)` for $k \in \{1,2,3,4,5\}$ in the ring `QS(5)`", which in our language is $\mathbf{t}_5\mathbf{t}_4\mathbf{t}_3\mathbf{t}_2\mathbf{t}_1$. (Yes, Sage multiplies the `t(5, k)` in the order of increasing $k$, but because Sage's order of multiplication in $\mathbf{k}[S_n]$ is the opposite of ours, this results in $\mathbf{t}_5\mathbf{t}_4\mathbf{t}_3\mathbf{t}_2\mathbf{t}_1$ rather than $\mathbf{t}_1\mathbf{t}_2\mathbf{t}_3\mathbf{t}_4\mathbf{t}_5$.) The `RHS` is our $\nabla$. The last line (`LHS == RHS`) thus asks Sage whether $\mathbf{t}_5\mathbf{t}_4\mathbf{t}_3\mathbf{t}_2\mathbf{t}_1 = \nabla$. And Sage quickly answers `True`.

Incidentally, what happens if we multiply the $\mathbf{t}_k$'s in the opposite order? In other words, what is $\mathbf{t}_1\mathbf{t}_2\mathbf{t}_3\mathbf{t}_4\mathbf{t}_5$ ? Sage can easily answer this:

```
QS(5).prod(t(5, k) for k in range(5, 0, -1))
```

Alas, the result does not look like there is a simple pattern to its coefficients.

**Question 3.14.1.** Is there? Can we at least describe the permutations that appear in the result with nonzero coefficient? In other words, what permutations $w \in S_n$ can be written in the form

$$w = \operatorname{cyc}_{1,2,\ldots,i_1} \operatorname{cyc}_{2,3,\ldots,i_2} \operatorname{cyc}_{3,4,\ldots,i_3} \cdots \operatorname{cyc}_n$$

for some $i_1, i_2, \ldots, i_n \in [n]$ with $i_k \geq k$ for each $k \in [n]$ ? Apparently, the number of such permutations $w$ is given by OEIS Sequence A229046; can this be proved?

### 3.14.4. Exercises

In the following exercises, you are supposed to give both the answer and the Sage code used to obtain it.

**Exercise 3.14.1.** [1] Write Sage code to compute the $\mathbf{X}_{a,b}$ from Definition 3.10.20.

**Exercise 3.14.2.** [1] Write Sage code to compute the random-to-random shuffles $\mathbf{R}_k$ (Definition 3.10.14) or find the inbuilt function that computes them.

[**Hint:** They appear in Sage's `SymmetricGroupAlgebra` class under a different name.]

**Exercise 3.14.3.** [2] Playing with Exercise 3.13.1, you may have made the following conjecture: "If two permutations $u, v \in S_n$ are conjugate, then there exists an involution $w \in S_n$ such that $u = wvw^{-1}$ (or, equivalently, $u = wvw$)."

Use Sage to check whether this conjecture holds for all $n \leq 6$.

[**Hint:** This requires `for`-loops.]

**Exercise 3.14.4.** An *inversion* of a permutation $w \in S_n$ means a pair $(i,j) \in [n]^2$ satisfying $i < j$ and $w(i) > w(j)$. The # of inversions of a given permutation $w \in S_n$ will be denoted $\ell(w)$.

For each $t \in \mathbf{k}$, define the element

$$\mathbf{L}_t := \sum_{w \in S_n} t^{\ell(w)} w \in \mathbf{k}[S_n].$$

(In particular, $\mathbf{L}_1 = \nabla$ and $\mathbf{L}_{-1} = \nabla^-$ and $\mathbf{L}_0 = 1$.)

**(a)** $\boxed{1}$ Is it true that $\mathbf{L}_t \mathbf{L}_s = \mathbf{L}_s \mathbf{L}_t$ for all $t, s \in \mathbf{k}$ ?

**(b)** $\boxed{1}$ Is it true that $\operatorname{span}_{\mathbf{k}} \{\mathbf{L}_t \mid t \in \mathbf{k}\}$ is a $\mathbf{k}$-subalgebra of $\mathbf{k}[S_n]$ ?

[**Hint:** The cases $n \leq 3$ are not representative of the situation.]

## 3.15. A look at the structure of symmetric group algebras

Now let us determine the structure of the $\mathbf{k}$-algebra $\mathbf{k}[S_n]$ for small values of $n$ (specifically, 2 and 3). The word "structure" here refers to the question "is it isomorphic to an algebra we know?" and related questions. While at that, let us also try to understand the subalgebras $Z(\mathbf{k}[S_n])$ and $\mathrm{GZ}_n$ (defined in Corollary 3.5.4).

### 3.15.1. $\mathbf{k}[S_2]$

We begin with the case $n = 2$.

The group $S_2$ is abelian, so its group algebra $\mathbf{k}[S_2]$ is commutative. This algebra has basis $(1, s_1)$ (as a $\mathbf{k}$-module) with multiplication table

$$1 \cdot 1 = 1, \qquad 1 s_1 = s_1, \qquad s_1 1 = s_1, \qquad s_1 s_1 = 1.$$

Have you seen this $\mathbf{k}$-algebra before?

For comparison, consider the quotient ring $\mathbf{k}[x] / (x^2 - 1)$ of the polynomial ring $\mathbf{k}[x]$ modulo the principal ideal generated by $x^2 - 1$. This quotient ring is known as the ring of split-complex numbers[31] over $\mathbf{k}$, and has a basis $(\overline{1}, \overline{x})$ (where $\overline{p}$ means the residue class of $p \in \mathbf{k}[x]$ in the quotient ring) with multiplication table

$$\overline{1} \cdot \overline{1} = \overline{1}, \qquad \overline{1} \cdot \overline{x} = \overline{x}, \qquad \overline{x} \cdot \overline{1} = \overline{x}, \qquad \overline{x} \cdot \overline{x} = \overline{1}.$$

This is exactly the multiplication table we found for $\mathbf{k}[S_2]$ above, except that 1 and $s_1$ have been renamed as $\overline{1}$ and $\overline{x}$.

[31] The name is a reference to the ring of *complex numbers* over $\mathbf{k}$, which is defined as $\mathbf{k}[x] / (x^2 + 1)$. (Of course, for $\mathbf{k} = \mathbb{R}$, this is the standard ring $\mathbb{C}$ of complex numbers.) Thus, the split-complex numbers differ from the complex numbers only in that the polynomial $x^2 - 1$ is used instead of $x^2 + 1$.

Thus, we obtain a $\mathbf{k}$-algebra isomorphism

$$\begin{aligned}\mathbf{k}\left[x\right] / \left(x^2-1\right) &\to \mathbf{k}\left[S_2\right], \\ \overline{x} &\mapsto s_1.\end{aligned}$$

(We need not say that it sends $\overline{1} \mapsto 1$, because this is automatically true for $\mathbf{k}$-algebra morphisms.)

Thus, $\mathbf{k}\left[S_2\right]$ is isomorphic to $\mathbf{k}\left[x\right] / \left(x^2-1\right)$. We can say more if we know a bit about $\mathbf{k}$. For instance, for $\mathbf{k} = \mathbb{Q}$, we have

$$\begin{aligned}\mathbb{Q}\left[x\right] / \left(x^2-1\right) &= \mathbb{Q}\left[x\right] / \left(\left(x-1\right)\left(x+1\right)\right) \\ &\cong \underbrace{\left(\mathbb{Q}\left[x\right] / \left(x-1\right)\right)}_{\cong \mathbb{Q}} \times \underbrace{\left(\mathbb{Q}\left[x\right] / \left(x+1\right)\right)}_{\cong \mathbb{Q}} \\ &\qquad \left(\begin{array}{c}\text{by the Chinese Remainder Theorem} \\ \text{(for } \mathbb{Q}\text{-algebras), since the polynomials } x-1 \\ \text{and } x+1 \text{ are coprime}\end{array}\right) \\ &\cong \mathbb{Q} \times \mathbb{Q} \qquad \left(\begin{array}{c}\text{a direct product of } \mathbb{Q}\text{-algebras, with} \\ \text{entrywise addition and multiplication}\end{array}\right).\end{aligned}$$

This can also be shown more explicitly: There is a $\mathbb{Q}$-algebra isomorphism

$$\begin{aligned}\mathbb{Q}\left[x\right] / \left(x^2-1\right) &\to \mathbb{Q} \times \mathbb{Q}, \\ \overline{a+bx} &\mapsto \left(a+b,\ a-b\right) \qquad \text{for all } a, b \in \mathbb{Q}.\end{aligned}$$

More generally, for any commutative ring $\mathbf{k}$, there is a $\mathbf{k}$-algebra morphism

$$\begin{aligned}\mathbf{k}\left[x\right] / \left(x^2-1\right) &\to \mathbf{k} \times \mathbf{k}, \\ \overline{a+bx} &\mapsto \left(a+b,\ a-b\right) \qquad \text{for all } a, b \in \mathbf{k}.\end{aligned}$$

When 2 is invertible in $\mathbf{k}$, this is a $\mathbf{k}$-algebra **iso**morphism, so we conclude that

$$\mathbf{k}\left[S_2\right] \cong \mathbf{k}\left[x\right] / \left(x^2-1\right) \cong \mathbf{k} \times \mathbf{k} \qquad \text{in this case.}$$

If $\mathbf{k}$ is a field of characteristic 2, then the ring $\mathbf{k}\left[S_2\right]$ has a nonzero nilpotent element $\nabla$ (since Corollary 3.3.4 yields $\nabla^2 = 2! \cdot \nabla = 0$), so it cannot be isomorphic to $\mathbf{k} \times \mathbf{k}$. Instead, in this case, we have $\mathbf{k}\left[S_2\right] \cong \mathbf{k}\left[x\right] / \left(x^2\right)$ (the ring of dual numbers over $\mathbf{k}$), since $x^2 - 1 = x^2 - 1^2 = \left(x-1\right)^2$ can be turned into $x^2$ by substituting $x$ for $x-1$.

We need not analyze $Z\left(\mathbf{k}\left[S_2\right]\right)$ or $\mathrm{GZ}_2$ separately, since we have $Z\left(\mathbf{k}\left[S_2\right]\right) = \mathrm{GZ}_2 = \mathbf{k}\left[S_2\right]$.

### 3.15.2. $\mathbf{k}[S_3]$

Now, we move on to the case $n = 3$.

We have $|S_3| = 6$ and $S_3 = \{1,\ s_1,\ s_2,\ t_{1,3},\ z,\ z^{-1}\}$, where

$$z := \mathrm{cyc}_{1,2,3} \qquad \text{and thus} \qquad z^{-1} = \mathrm{cyc}_{1,3,2} = z^2.$$

So the $\mathbf{k}$-algebra $\mathbf{k}[S_3]$ has basis $\left(1,\ s_1,\ s_2,\ t_{1,3},\ z,\ z^{-1}\right)$ as a $\mathbf{k}$-module. How does it look like as a $\mathbf{k}$-algebra?

We begin with its center $Z(\mathbf{k}[S_3])$, since it is smaller. By our study of centers of group algebras (specifically, Theorem 3.9.5 and Corollary 3.9.10), we know that $Z(\mathbf{k}[S_3])$ has basis $(1,\ \mathbf{t},\ \mathbf{c})$, where

$$\mathbf{t} = s_1 + s_2 + t_{1,3} = (\text{sum of all transpositions}) \qquad \text{and}$$
$$\mathbf{c} = z + z^{-1} = \left(\text{sum of all cycles } \mathrm{cyc}_{i,j,k}\right).$$

The multiplication table for this basis is

$$\mathbf{t}^2 = 3 + 3\mathbf{c}, \qquad \mathbf{tc} = 2\mathbf{t}, \qquad \mathbf{c}^2 = 2 + \mathbf{c}$$

(the unitality and the commutativity of $Z(\mathbf{k}[S_3])$ take care of the rest). For $\mathbf{k} = \mathbb{Q}$, the first equality can be solved for $\mathbf{c}$, leading to $\mathbf{c} = \dfrac{\mathbf{t}^2 - 3}{3}$, and then the other two equalities become

$$\mathbf{t}(\mathbf{t} - 3)(\mathbf{t} + 3) = 0 \qquad \text{and} \qquad \mathbf{t}^2(\mathbf{t} - 3)(\mathbf{t} + 3) = 0$$

(after a bit of computation). Of course, the last equality is redundant, so we only need to deal with $\mathbf{t}(\mathbf{t} - 3)(\mathbf{t} + 3) = 0$. The three elements $1,\ \mathbf{t},\ \mathbf{t}^2$ are $\mathbb{Q}$-linearly independent, and the equality $\mathbf{t}(\mathbf{t} - 3)(\mathbf{t} + 3) = 0$ lets us express any higher power of $\mathbf{t}$ in terms of them. Thus, $\left(1,\ \mathbf{t},\ \mathbf{t}^2\right)$ is a basis of the $\mathbb{Q}$-vector space $Z(\mathbb{Q}[S_3])$, and we get

$$\begin{aligned} Z(\mathbb{Q}[S_3]) &\cong \mathbb{Q}[x] \,/\, (x(x-3)(x+3)) \\ &\qquad \left(\begin{array}{c} \text{via the } \mathbb{Q}\text{-algebra isomorphism} \\ \text{from } \mathbb{Q}[x] \,/\, (x(x-3)(x+3)) \text{ to } Z(\mathbb{Q}[S_3]) \\ \text{that sends } \overline{x} \text{ to } \mathbf{t} \end{array}\right) \\ &\cong \mathbb{Q} \times \mathbb{Q} \times \mathbb{Q} \qquad (\text{again by the Chinese Remainder Theorem}). \end{aligned}$$

More generally, we get a $\mathbf{k}$-algebra isomorphism

$$Z(\mathbf{k}[S_3]) \cong \mathbf{k} \times \mathbf{k} \times \mathbf{k}$$

whenever 2 and 3 are invertible in $\mathbf{k}$. When $3 = 0$ in $\mathbf{k}$, this is false, since $\mathbf{t}^2 = 0$ in this case. This is also false when $2 = 0$ in $\mathbf{k}$, since $(\mathbf{c} + \mathbf{t} + 1)^2 = 0$ in this case.

**Exercise 3.15.1.** Let $\mathbf{k}$ be the field with 3 elements (that is, $\mathbb{F}_3 = \mathbb{Z}/3$).

**(a)** $\boxed{2}$ Prove that the center $Z(\mathbf{k}[S_3])$ is neither isomorphic to $\mathbf{k} \times \mathbf{k} \times \mathbf{k}$, nor isomorphic to a quotient ring of $\mathbf{k}[x]$.

**(b)** $\boxed{2}$ Prove that $Z(\mathbf{k}[S_3])$ is isomorphic to the quotient ring $\mathbf{k}[x,y] / (x^2, xy, y^2)$.

[**Hint:** For part **(a)**, compute $(\kappa + \lambda \mathbf{t} + \mu \mathbf{c})^2$ for $\lambda, \mu, \kappa \in \mathbf{k}$, and show that this is always a $\mathbf{k}$-linear combination of 1 and $\lambda \mathbf{t} + \mu \mathbf{c}$.]

Next, we consider the Gelfand–Tsetlin subalgebra $\mathrm{GZ}_3$. We can easily see that $\mathrm{GZ}_3$ has basis $(1,\ \mathbf{t},\ \mathbf{c},\ s_1)$ (note that $\mathbf{m}_3 = \mathbf{t} - s_1$). The multiplication table consists of the formulas for $Z(\mathbf{k}[S_3])$ we found above, plus

$$s_1^2 = 1, \qquad \mathbf{t} s_1 = \mathbf{c} + 1, \qquad \mathbf{c} s_1 = \mathbf{t} - 1.$$

Can we identify this algebra? At least for $\mathbf{k} = \mathbb{Q}$, we can try to find an element with 4 linearly independent powers, since such an element will generate $\mathrm{GZ}_3$ and thus allow us to write $\mathrm{GZ}_3$ as a quotient of $\mathbf{k}[x]$.

We know such an element: the jucys–murphy $\mathbf{m}_3$, whose first 4 powers are linearly independent (at least when 2 is invertible in $\mathbf{k}$), and which satisfies

$$(\mathbf{m}_3 - 2)(\mathbf{m}_3 - 1)(\mathbf{m}_3 + 1)(\mathbf{m}_3 + 2) = 0.$$

Thus, when 2 is invertible in $\mathbf{k}$, we obtain

$$\mathrm{GZ}_3 \cong \mathbf{k}[x] / ((x-2)(x-1)(x+1)(x+2)).$$

When 2 and 3 are invertible in $\mathbf{k}$, this is furthermore isomorphic to $\mathbf{k} \times \mathbf{k} \times \mathbf{k} \times \mathbf{k}$ using the Chinese Remainder Theorem.

What about the whole group algebra $\mathbf{k}[S_3]$ ? Let us again start with the case $\mathbf{k} = \mathbb{Q}$. We begin by identifying central idempotents. A *central idempotent* in a $\mathbf{k}$-algebra $A$ means an idempotent in $Z(A)$. When $z$ is a central idempotent of $A$, we can decompose $A$ as a direct product $(zA) \times ((1-z)A)$ of two $\mathbf{k}$-algebras[32] (see [23wa, Exercise 2.10.4 and the following few paragraphs]). We already know a central idempotent of $\mathbb{Q}[S_3]$: namely, $\frac{1}{6}\nabla$ (since Corollary 3.3.4

[32]Here we are using the standard notation

$$zA = \{za \mid a \in A\}.$$

For an arbitrary $z \in Z(A)$, this is an ideal of $A$, and thus closed under multiplication. When $z$ is furthermore idempotent, this ideal $zA$ furthermore is a $\mathbf{k}$-algebra with unity $z$ (and multiplication inherited from $A$). Moreover, when $z$ is a central idempotent, $1-z$ is a central idempotent as well, and the two ideals $zA$ and $(1-z)A$ add up to the whole $A$ (that is, $A = (zA) \oplus ((1-z)A)$ as an internal direct sum).

shows that $\frac{1}{n!}\nabla$ is idempotent for any $n \in \mathbb{N}$). Thus, we can write $\mathbb{Q}[S_3]$ as a direct product

$$\mathbb{Q}[S_3] \cong \left(\left(\frac{1}{6}\nabla\right)\mathbb{Q}[S_3]\right) \times \left(\left(1-\frac{1}{6}\nabla\right)\mathbb{Q}[S_3]\right).$$

The first factor here is the $\mathbb{Q}$-algebra $\left(\frac{1}{6}\nabla\right)\mathbb{Q}[S_3]$, which has dimension 1 as a $\mathbb{Q}$-vector space (since any multiple of $\frac{1}{6}\nabla$ in $\mathbb{Q}[S_3]$ is a **scalar** multiple of $\nabla$, due to (17)), and thus is $\cong \mathbb{Q}$ as a $\mathbb{Q}$-algebra (why?). The second factor is the $\mathbb{Q}$-algebra $\left(1-\frac{1}{6}\nabla\right)\mathbb{Q}[S_3]$, which has dimension 5.

This second factor can be decomposed further, since it contains another central idempotent: namely, $\frac{1}{6}\nabla^-$. (The centrality and the idempotency of this element are shown just as easily as for $\frac{1}{6}\nabla$. To see that it belongs to $\left(1-\frac{1}{6}\nabla\right)\mathbb{Q}[S_3]$, we must check that $\nabla\nabla^- = 0$, but this follows easily from Theorem 3.7.11.) Thus, we obtain

$$\left(1-\frac{1}{6}\nabla\right)\mathbb{Q}[S_3] \cong \left(\left(\frac{1}{6}\nabla^-\right)\mathbb{Q}[S_3]\right) \times \left(\left(1-\frac{1}{6}\nabla-\frac{1}{6}\nabla^-\right)\mathbb{Q}[S_3]\right).$$

Again, the first factor is 1-dimensional and isomorphic to $\mathbb{Q}$. It remains to decompose the $\mathbb{Q}$-algebra $\left(1-\frac{1}{6}\nabla-\frac{1}{6}\nabla^-\right)\mathbb{Q}[S_3]$ here, which has dimension 4 as a $\mathbb{Q}$-vector space. At this point, we have run out of central idempotents (except the obvious ones – 0 and 1 – which are useless, as they yield trivial decompositions). What can we say about this 4-dimensional noncommutative algebra?

What follows is a deus ex machina; I don't know how to motivate it at this point (although soon I will). Consider the four elements

$$\begin{aligned}
\mathbf{x} &:= \frac{1}{3}\,\mathrm{id} + \frac{1}{6}\mathbf{m}_3 - \frac{1}{3}s_1 - \frac{1}{6}\left(z+z^{-1}\right), \\
\mathbf{y} &:= \frac{1}{3}\left(s_2 - z + z^{-1} - t_{1,3}\right), \\
\mathbf{z} &:= \frac{4}{3}S(\mathbf{y}) = \frac{-4}{3}T_{\mathrm{sign}}(\mathbf{y}), \\
\mathbf{w} &:= T_{\mathrm{sign}}(\mathbf{x})
\end{aligned}$$

of $\mathbb{Q}[S_3]$. Then, straightforward (if tiresome) computations show that $(\mathbf{x},\ \mathbf{y},\ \mathbf{z},\ \mathbf{w})$ is a basis of the very subspace $\left(1-\frac{1}{6}\nabla-\frac{1}{6}\nabla^-\right)\mathbb{Q}[S_3]$ that we are trying to

analyze. Moreover, this basis has a nice multiplication table:

$$\begin{array}{llll} \mathbf{xx}=\mathbf{x}, & \mathbf{xy}=\mathbf{y}, & \mathbf{xz}=0, & \mathbf{xw}=0, \\ \mathbf{yx}=0, & \mathbf{yy}=0, & \mathbf{yz}=\mathbf{x}, & \mathbf{yw}=\mathbf{y}, \\ \mathbf{zx}=\mathbf{z}, & \mathbf{zy}=\mathbf{w}, & \mathbf{zz}=0, & \mathbf{zw}=0, \\ \mathbf{wx}=0, & \mathbf{wy}=0, & \mathbf{wz}=\mathbf{z}, & \mathbf{ww}=\mathbf{w}. \end{array}$$

This is exactly the multiplication table for the elementary matrices

$$E_{1,1}=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad E_{1,2}=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad E_{2,1}=\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad E_{2,2}=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

in the matrix ring $\mathbb{Q}^{2\times 2}$, with $E_{1,1}$, $E_{1,2}$, $E_{2,1}$, $E_{2,2}$ renamed as $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, $\mathbf{w}$. Thus, there is a $\mathbb{Q}$-algebra isomorphism

$$\mathbb{Q}^{2\times 2} \to \left(1-\frac{1}{6}\nabla-\frac{1}{6}\nabla^{-}\right)\mathbb{Q}\left[S_3\right]$$

that sends $E_{1,1}$, $E_{1,2}$, $E_{2,1}$, $E_{2,2}$ renamed as $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, $\mathbf{w}$, respectively. Consequently,

$$\left(1-\frac{1}{6}\nabla-\frac{1}{6}\nabla^{-}\right)\mathbb{Q}\left[S_3\right] \cong \mathbb{Q}^{2\times 2}.$$

Altogether, we have obtained

$$\begin{aligned} \mathbb{Q}\left[S_3\right] &\cong \left(\left(\frac{1}{6}\nabla\right)\mathbb{Q}\left[S_3\right]\right) \times \underbrace{\left(\left(1-\frac{1}{6}\nabla\right)\mathbb{Q}\left[S_3\right]\right)}_{\cong \left(\left(\frac{1}{6}\nabla^{-}\right)\mathbb{Q}[S_3]\right)\times\left(\left(1-\frac{1}{6}\nabla-\frac{1}{6}\nabla^{-}\right)\mathbb{Q}[S_3]\right)} \\ &\cong \underbrace{\left(\left(\frac{1}{6}\nabla\right)\mathbb{Q}\left[S_3\right]\right)}_{\cong\mathbb{Q}} \times \underbrace{\left(\left(\frac{1}{6}\nabla^{-}\right)\mathbb{Q}\left[S_3\right]\right)}_{\cong\mathbb{Q}} \times \underbrace{\left(\left(1-\frac{1}{6}\nabla-\frac{1}{6}\nabla^{-}\right)\mathbb{Q}\left[S_3\right]\right)}_{\cong\mathbb{Q}^{2\times 2}} \\ &\cong \mathbb{Q}\times\mathbb{Q}\times\mathbb{Q}^2. \end{aligned}$$

Explicitly, we have a $\mathbb{Q}$-algebra isomorphism

$$\begin{aligned} \mathbb{Q}\times\mathbb{Q}\times\mathbb{Q}^{2\times 2} &\to \mathbb{Q}\left[S_3\right], \\ (1,0,0) &\mapsto \frac{1}{6}\nabla, \\ (0,1,0) &\mapsto \frac{1}{6}\nabla^{-}, \\ (0,0,E_{1,1}) &\mapsto \mathbf{x}, \\ (0,0,E_{1,2}) &\mapsto \mathbf{y}, \\ (0,0,E_{2,1}) &\mapsto \mathbf{z}, \\ (0,0,E_{2,2}) &\mapsto \mathbf{w}. \end{aligned}$$

More generally, the same construction defines a $\mathbf{k}$-algebra isomorphism

$$\mathbf{k} \times \mathbf{k} \times \mathbf{k}^{2\times 2} \to \mathbf{k}\left[S_3\right]$$

for any commutative ring $\mathbf{k}$ as long as 2 and 3 are invertible in $\mathbf{k}$.

Now how on earth did I find the elements $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, $\mathbf{w}$ that came to our rescue here? The next chapter will give us an idea.

### 3.15.3. About the general case

We cannot go much farther than $n = 3$ with such ad-hoc methods. But some patterns are already visible in these cases. Not all these patterns persist. For example, we observed above that $\mathrm{GZ}_3 = \mathrm{AlgGen}\left\{\mathbf{m}_3\right\}$ for $\mathbf{k} = \mathbb{Q}$, but not all $n$ satisfy $\mathrm{GZ}_n = \mathrm{AlgGen}\left\{\mathbf{m}_n\right\}$. Likewise, $Z\left(\mathbb{Q}\left[S_n\right]\right) = \mathrm{AlgGen}\left\{\mathbf{t}\right\}$ (where $\mathbf{t}$ is the sum of all transpositions in $S_n$) holds for $n = 3$ but not for $n = 5$.

But what is true is the following:

**Theorem 3.15.1.** Assume that $n!$ is invertible in $\mathbf{k}$. Then:

**(a)** The $\mathbf{k}$-algebra $\mathbf{k}\left[S_n\right]$ is isomorphic to a direct product of matrix rings over $\mathbf{k}$ (including $\mathbf{k}^{1\times 1} \cong \mathbf{k}$). That is,

$$\mathbf{k}\left[S_n\right] \cong \prod_{\lambda\in\Lambda} \mathbf{k}^{f_\lambda \times f_\lambda}$$

for some finite set $\Lambda$ and some positive integers $f_\lambda$ for all $\lambda \in \Lambda$. Moreover, the set $\Lambda$ is the set of all partitions of $n$, whereas the integer $f_\lambda$ is the # of so-called standard tableaux of shape $\lambda$ (a combinatorial object we will define later).

**(b)** The center $Z\left(\mathbf{k}\left[S_n\right]\right)$ is isomorphic to the direct product $\mathbf{k}^{|\Lambda|} = \prod_{\lambda\in\Lambda} \mathbf{k}$. In other words, it is isomorphic to $\mathbf{k}^{p(n)}$, where $p\left(n\right)$ is the # of partitions of $n$.

**(c)** The Gelfand–Tsetlin subalgebra $\mathrm{GZ}_n$ is isomorphic to the direct product $\mathbf{k}^{t(n)}$, where $t\left(n\right)$ is the # of all standard tableaux of all size-$n$ partition shapes, or, equivalently, the # of all involutions in $S_n$.

This is one of the main results of this course, and is both deep and surprising; we will prove it later, using the so-called seminormal basis of $\mathbf{k}\left[S_n\right]$.

**Remark 3.15.2.** When $\mathbf{k}$ is a field of characteristic 0, the numbers $p\left(n\right)$ and $t\left(n\right)$ in Theorem 3.15.1 are the dimensions of the $\mathbf{k}$-vector spaces $Z\left(\mathbf{k}\left[S_n\right]\right)$ and $\mathrm{GZ}_n$. Here is a little table of these numbers in Theorem 3.15.1:

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $p\left(n\right)$ | 1 | 1 | 2 | 3 | 5 | 7 | 11 | 15 | 22 | 30 | 42 | 56 |
| $t\left(n\right)$ | 1 | 1 | 2 | 4 | 10 | 26 | 76 | 232 | 764 | 2620 | 9496 | 35696 |

.

Note that both numbers $p(n)$ and $t(n)$ are much smaller than $n! = |S_n| = \dim(\mathbf{k}[S_n])$ for $n \geq 3$.

In the OEIS (Online Encyclopedia of Integer Sequences), the sequences $(p(n))_{n\geq 0}$ and $(t(n))_{n\geq 0}$ appear as A000041 and A000085, respectively.

The case when $n!$ is **not** invertible in $\mathbf{k}$ (this is known as the *modular case*) is much less well-behaved; in this case, $\mathbf{k}[S_n]$ is not a direct product of known rings. In particular, if $\mathbf{k}$ is a field of positive characteristic $\leq n$, then $\nabla \in \mathbf{k}[S_n]$ is nilpotent (with $\nabla^2 = 0$), so any description of $\mathbf{k}[S_n]$ will have to deal with both idempotents and nilpotents, which dooms it to be more complicated. No good descriptions of $\mathbf{k}[S_n]$ in the general case are known, although some partial results have been revealing themselves slowly over the last 80 years (e.g., see [Kuenze15] for some work).

## 3.16. More exercises

The following is an assortment of further exercises on $\mathbf{k}[S_n]$.

**Exercise 3.16.1.** For any $a, b \in [n]$, we define the element

$$\nabla_{b,a} := \sum_{\substack{w\in S_n; \\ w(a)=b}} w \in \mathbf{k}[S_n].$$

(This is a particular case of the rook sums $\nabla_{B,A}$ we will define later on.)

**(a)** $\boxed{3}$ Prove that every $a, b, c, d \in [n]$ satisfy

$$\nabla_{d,c}\nabla_{b,a} = \begin{cases} (n-2)!\,(\nabla - \nabla_{d,a}), & \text{if } b \neq c; \\ (n-1)!\nabla_{d,a}, & \text{if } b = c. \end{cases}$$

**(b)** $\boxed{2}$ Let $a, b \in [n]$, and let $X$ be a subset of $[n]$ having size $|X| > 2$. Prove that

$$\nabla_X^- \nabla_{b,a} = \nabla_{b,a}\nabla_X^- = 0.$$

The next exercise introduces another family of commuting elements of $\mathbf{k}[S_n]$:

**Exercise 3.16.2.** $\boxed{3}$ Let $z_1, z_2, \ldots, z_n \in \mathbf{k}$ be some scalars such that each difference $z_i - z_j$ with $i \neq j$ is invertible in $\mathbf{k}$. (For instance, if $\mathbf{k}$ is a field, then $z_1, z_2, \ldots, z_n$ can be any $n$ distinct elements of $\mathbf{k}$.) Set

$$\mathbf{H}_i := \sum_{\substack{j\in[n]; \\ j\neq i}} \frac{1}{z_j - z_i} t_{i,j} \in \mathbf{k}[S_n] \qquad \text{for each } i \in [n].$$

Prove that the elements $\mathbf{H}_1$, $\mathbf{H}_2$, $\ldots$, $\mathbf{H}_n$ commute. (These elements are known as the *Knizhnik–Zamolodchikov sums* in $\mathbf{k}[S_n]$.)

The next exercise continues the hunt for polynomials that annihilate elements of $\mathbf{k}[S_n]$:

**Exercise 3.16.3.** **(a)** [2] Let $P \in \mathbb{Z}[x]$ be the polynomial $(x - n! + 1)(x - 1)(x + 1)$. Let $w_0$ be the permutation $\operatorname{oln}(n, n-1, \ldots, 2, 1) \in S_n$. Prove that $P(\nabla - w_0) = 0$.

**(b)** [1] For each $\mathbf{a} \in \mathbf{k}[S_n]$ and each $m \in \mathbb{N}$, prove that

$$(\nabla + \mathbf{a})^m = \mathbf{a}^m + \lambda\nabla \qquad \text{for some scalar } \lambda \in \mathbf{k}.$$

**(c)** [1] Let $w \in S_n$, and let $m$ be the order of $w$ in $S_n$ (that is, the smallest positive integer satisfying $w^m = 1$). Let $Q \in \mathbb{Z}[x]$ be the polynomial $(x - n! + 1)\left(x^m - (-1)^m\right)$. Prove that $Q(\nabla - w) = 0$.

Next, we present a celebrated result of Olshanskii as a (difficult) exercise. First, we recall a notion from ring theory:

**Definition 3.16.1.** Let $B$ be a subset of a ring $A$. Then, the *centralizer* of $B$ in $A$ denotes the subset

$$\{a \in A \mid ab = ba \text{ for all } b \in B\} \text{ of } A$$

(consisting of all elements of $A$ that commute with all elements of $B$). This is itself a subring of $A$, and is denoted by $Z_A(B)$. (See [23wa, §2.3.5] for some examples.)

**Exercise 3.16.4.** Assume that $n \geq 1$. As in Remark 3.12.16, let us regard $\mathbf{k}[S_{n-1}]$ as a subring of $\mathbf{k}[S_n]$ by identifying each $\mathbf{a} \in \mathbf{k}[S_{n-1}]$ with its image $(I_{n-1\to n})_*(\mathbf{a})$ under the default embedding of group algebras.

**(a)** [1] Prove that the jucys–murphy $\mathbf{m}_n \in \mathbf{k}[S_n]$ commutes with all elements of $\mathbf{k}[S_{n-1}]$. In other words, prove that $\mathbf{m}_n \in Z_{\mathbf{k}[S_n]}(\mathbf{k}[S_{n-1}])$.

**(b)** [7] Prove that

$$Z_{\mathbf{k}[S_n]}(\mathbf{k}[S_{n-1}]) = \operatorname{AlgGen}\left(Z(\mathbf{k}[S_{n-1}]) \cup \{\mathbf{m}_n\}\right).$$

In other words, prove that an element $\mathbf{a} \in \mathbf{k}[S_n]$ commutes with all elements of $\mathbf{k}[S_{n-1}]$ if and only if $\mathbf{a}$ lies in the $\mathbf{k}$-subalgebra of $\mathbf{k}[S_n]$ generated by the elements of the center of $\mathbf{k}[S_{n-1}]$ and the $n$-th jucys–murphy $\mathbf{m}_n$.

The following exercise is about the group algebras of cyclic groups, not symmetric groups. It offers some contrast to the analysis of $\mathbf{k}[S_n]$ we have seen above.

**Exercise 3.16.5.** Let $n$ be a positive integer. Consider the cyclic group $C_n = \{g^0, g^1, \ldots, g^{n-1}\}$ of size $n$. Let $\zeta := e^{2\pi i/n} \in \mathbb{C}$ (so that the complex numbers $\zeta^0, \zeta^1, \ldots, \zeta^{n-1}$ are the vertices of a regular $n$-gon inscribed in the unit circle).

For each $i \in \mathbb{Z}$, define the element

$$\Delta_i := \sum_{k=0}^{n-1} \zeta^{ik} g^k \in \mathbb{C}\left[C_n\right].$$

(Note that this depends only on the residue class of $i$ modulo $n$, since $\zeta^n = 1$.)

**(a)** $\boxed{1}$ Prove that $\Delta_i^2 = n\Delta_i$ for each $i \in \mathbb{Z}$.

**(b)** $\boxed{2}$ Prove that $\Delta_i \Delta_j = 0$ for any distinct $i, j \in \{0, 1, \ldots, n-1\}$.

**(c)** $\boxed{2}$ Consider the direct product $\mathbb{C}^n$ of the $\mathbb{C}$-algebras $\underbrace{\mathbb{C}, \mathbb{C}, \ldots, \mathbb{C}}_{n \text{ times}}$. Prove that the map

$$\mathbb{C}^n \to \mathbb{C}\left[C_n\right],$$
$$(a_1, a_2, \ldots, a_n) \mapsto \frac{1}{n} \sum_{i=0}^{n-1} a_i \Delta_i$$

is a $\mathbb{C}$-algebra isomorphism. (This isomorphism is – up to scalar multiple – known as the *discrete Fourier transform*.)

**(d)** $\boxed{2}$ Prove that the $\mathbb{R}$-algebra $\mathbb{R}\left[C_3\right]$ is **not** isomorphic to $\mathbb{R}^3$. (It is not enough to show that $\zeta \notin \mathbb{R}$ for $n = 3$; you should also show that no other isomorphism can exist.)

**(e)** $\boxed{2}$ Prove that $\mathbb{R}\left[C_4\right] \cong \mathbb{R} \times \mathbb{R} \times \mathbb{C}$ as $\mathbb{R}$-algebras.

Note that this exercise shows that the structure of $\mathbf{k}\left[C_n\right]$ can change from $\mathbf{k} = \mathbb{R}$ to $\mathbf{k} = \mathbb{C}$. But this does not happen for $\mathbf{k}\left[S_n\right]$, since Theorem 3.15.1 **(a)** describes the structure independently of $\mathbf{k}$ when $\mathbf{k}$ is a $\mathbb{Q}$-algebra.

# 4. Actions and representations

In this chapter, we will introduce group actions and representations of groups and algebras. This might feel like a sideshow at first, but I hope to show that it is actually one of the best ways to study groups and algebras. Viewing groups or algebras through the prism of their actions is how many of the above results were first discovered, and in some cases (e.g., for Theorem 3.10.10 **(b)**) it is the only way I know to prove them. It will also explain to some extent how the elements $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}$ of $\mathbb{Q}\left[S_3\right]$ were found in Subsection 3.15.2.

## 4.1. Group actions and $G$-sets

**Convention 4.1.1.** For the entirety of Section 4.1, we let $G$ be a group (not necessarily finite).

### 4.1.1. Motivation

Here is a slightly provocative slogan: Group elements want to be permutations. I.e., a group wants to act.

This might sound strange. After all, we nowadays define groups as abstract structures consisting of disembodied objects that multiply via a multiplication table. But the elements of the (historically) first groups ever defined (by Galois in the early 19th century) were permutations of sets. That is, the first groups ever defined were subgroups of symmetric groups.

Even groups that are not defined in this way can be forced into this mold:

**Example 4.1.2.** Recall the additive group $(\mathbb{Z}, +, 0)$. Its elements are integers, not permutations, but you can replace them by permutations that get multiplied (i.e., composed) just like the corresponding integers are added. To be specific: For each $n \in \mathbb{Z}$, we consider the map

$$\begin{aligned} \tau_n : \mathbb{Z} &\to \mathbb{Z}, \\ k &\mapsto n + k, \end{aligned}$$

which we call the *translation* by $n$ (since it is geometrically a translation along the number line). These translations $\tau_n$ are permutations of $\mathbb{Z}$ (that is, they belong to the symmetric group $S_{\mathbb{Z}}$), and they multiply like the $n$'s add: That is, they satisfy $\tau_n \tau_m = \tau_{n+m}$ for all $n, m \in \mathbb{Z}$, as well as $\tau_0 = \mathrm{id}$. Hence, there is a group morphism

$$\begin{aligned} \mathbb{Z} &\to S_{\mathbb{Z}}, \\ n &\mapsto \tau_n. \end{aligned}$$

Moreover, this morphism is injective (meaning that $\tau_n \neq \tau_m$ for all $n \neq m$), so its image is a subgroup of $S_{\mathbb{Z}}$ isomorphic to $\mathbb{Z}$. Thus, $\mathbb{Z}$ is isomorphic to a group of actual permutations (namely, the image we just found).

We can do this trick for any group, not just for $\mathbb{Z}$. Thus we obtain:

**Theorem 4.1.3** (Cayley's theorem)**.** Let $G$ be any group. For each $g \in G$, we define the permutation

$$\begin{aligned} \tau_g : G &\to G, \\ k &\mapsto gk \end{aligned}$$

of $G$. Then, the map

$$G \to S_G,$$
$$g \mapsto \tau_g$$

is an injective group morphism. Thus, its image is a subgroup of $S_G$ isomorphic to $G$.

And therefore $G$ is isomorphic to an actual group of permutations. This can be useful, since permutations have some features that elements of an arbitrary group might not have.[33]

Yet this is just the tip of the iceberg. We can turn the elements of a group $G$ not just into the permutations $\tau_g$, but also into other permutations of other sets, and we will soon see that this too can be rather useful.

### 4.1.2. Definition

Let us give this a name:

**Definition 4.1.4. (a)** Let $G$ be a group, and let $X$ be a set. A *left $G$-action on $X$* (aka a *left action of $G$ on $X$*) means a map $\rightharpoonup : G \times X \to X$ that satisfies the following two axioms:

1. *Associativity:* We have $g \rightharpoonup (h \rightharpoonup x) = (gh) \rightharpoonup x$ for all $g, h \in G$ and $x \in X$. Here and in the following, we use infix notation for the map $\rightharpoonup$, meaning that we write $g \rightharpoonup x$ for the image of $(g, x) \in G \times X$ under this map.
2. *Unitality:* We have $1_G \rightharpoonup x = x$ for all $x \in X$. (Here, $1_G$ means the neutral element of $G$.)

We often omit the word "left" in "left $G$-action" (or "left action"). We also often write $g \cdot x$ or $gx$ for $g \rightharpoonup x$, as long as this creates no ambiguity (i.e., as long as it cannot be mistaken for another kind of product, as it could be in Example 4.1.8 below). Thus, the axioms

$$g \rightharpoonup (h \rightharpoonup x) = (gh) \rightharpoonup x \qquad \text{and} \qquad 1_G \rightharpoonup x = x$$

become

$$g\,(hx) = (gh)\,x \qquad \text{and} \qquad 1_G x = x,$$

which look exactly like the associativity and the unitality axioms for a group (or monoid), but of course are different in that we are not composing two

[33] A well-known application of Theorem 4.1.3 is the fact that any finite group of order $4n+2$ (with $n \in \mathbb{N}$) must have a subgroup of index 2 (that is, of order $2n+1$). See `https://math.stackexchange.com/questions/225987/` for a proof of this fact.

elements of the same set but rather one element of $G$ with one element of $X$.

**(b)** A *left $G$-set* means a set $X$ equipped with a left $G$-action on $X$. We furthermore say that "*$G$ acts on $X$ from the left*" to mean that $X$ is a left $G$-set. The words "from the left" are frequently omitted here.

Analogously, we can define right $G$-actions and right $G$-sets. Here, the map $\rightharpoonup : G \times X \to X$ is replaced by a map

$$\leftharpoonup : X \times G \to X,$$

and the axioms are replaced by

$$(x \leftharpoonup h) \leftharpoonup g = x \leftharpoonup (hg) \qquad \text{and} \qquad x \leftharpoonup 1_G = x.$$

We prefer to use left $G$-actions, but sometimes right $G$-actions are more natural or convenient. The two concepts are easily translated into one another anyway (see Exercise 4.1.3 below).

**Convention 4.1.5.** The words "action", "$G$-set" and "act" shall be understood by default to mean "left action", "left $G$-set" and "act from the left", unless we qualify them by the word "right".

Definition 4.1.4 can be generalized by replacing "group" by "monoid". No other changes are necessary, and the notion of a left action of a monoid is indeed useful, although not as mainstream and well-behaved as that of a left action of a group.

### 4.1.3. Examples

There are lots of $G$-actions in nature. Let us list just a few examples. We begin with three $G$-actions on the set $G$ itself (i.e., the set $X$ in Definition 4.1.4 is $G$ itself):

**Example 4.1.6.** Let $G$ be any group. The *left regular $G$-action* is the left $G$-action on $G$ given by

$$g \rightharpoonup h = gh \qquad \text{for all } g \in G \text{ and } h \in G.$$

In other words, this action is just the multiplication map of $G$.

**Example 4.1.7.** Let $G$ be any group. The *inverse right regular $G$-action* is the left $G$-action on $G$ given by

$$g \rightharpoonup h = hg^{-1} \qquad \text{for all } g \in G \text{ and } h \in G.$$

Note that the formula $g \rightharpoonup h = hg$ would not define a left $G$-action, since it would fail the associativity axiom.

**Example 4.1.8.** Let $G$ be any group. The *conjugation $G$-action* is the left $G$-action on $G$ given by

$$g \rightharpoonup h = ghg^{-1} \qquad \text{for all } g \in G \text{ and } h \in G.$$

**Remark 4.1.9.** In Definition 4.1.4, it was suggested to abbreviate $g \rightharpoonup x$ as $gx$ when $g$ is an element of the group $G$ and $x$ is an element of a left $G$-set $X$. When the $G$-set $X$ is $G$ itself, this abbreviation is dangerous, since $gx$ could also mean the product of $g$ and $x$ in the group $G$. In Example 4.1.6, this would be unproblematic, since $g \rightharpoonup x$ agrees with the product of $g$ and $x$ by definition. However, doing the same in Example 4.1.7 or Example 4.1.8 would be a recipe for disaster.

Our next example is a $G$-action that can be defined on **any** set $X$, although not a very interesting one:

**Example 4.1.10.** Let $G$ be any group, and let $X$ be any set. The *trivial $G$-action on $X$* is the left $G$-action on $X$ given by

$$g \rightharpoonup x = x \qquad \text{for all } g \in G \text{ and } x \in X.$$

The next example is one of the most important in group theory. It answers the rather natural question "if a quotient $G/H$ of a group $G$ is not a group, then what is it?":

**Example 4.1.11.** Let $G$ be any group. Let $H$ be a subgroup of $G$. Let

$$G/H = \{\text{left cosets of } H \text{ in } G\} = \{uH \mid u \in G\}.$$

This set $G/H$ is (in general) not a group (unless $H$ is a normal subgroup of $G$), but it becomes a left $G$-set if we define the left $G$-action on $G/H$ by

$$g \rightharpoonup (uH) = (gu)\, H \qquad \text{for all } g \in G \text{ and } u \in G.$$

Here, of course, you can abbreviate $g \rightharpoonup x$ as $gx$, so this just takes the form

$$g\,(uH) = (gu)\, H \qquad \text{for all } g \in G \text{ and } u \in G.$$

So much for actions of arbitrary groups $G$. But there are many more examples for specific groups:

**Example 4.1.12.** Let $A$ be any set. The symmetric group $S_A$ acts on the set $A$ by

$$g \rightharpoonup x = g(x) \qquad \text{for all } g \in S_A \text{ and } x \in A.$$

This is a left $S_A$-action on $A$, and is called the *natural action* of $S_A$.

**Example 4.1.13.** Let $A$ be any set. Consider the set $\mathcal{P}(A)$ of all subsets of $A$. (This is the power set of $A$.) The symmetric group $S_A$ acts on $\mathcal{P}(A)$ by

$$g \rightharpoonup U = g(U) \qquad \text{for all } g \in S_A \text{ and } U \in \mathcal{P}(A).$$

(Here, as usual, $g(U)$ means the set $\{g(u) \mid u \in U\}$.)

For example, for $A = [5]$ and $g = \mathrm{cyc}_{1,2,3} \in S_A$, we have

$$g \rightharpoonup \{2,3,5\} = g(\{2,3,5\}) = \{g(2), g(3), g(5)\} = \{3,1,5\} = \{1,3,5\}.$$

Note that if $g \in S_A$ is a permutation of $A$, then $|g(U)| = |U|$ for any subset $U$ of $A$. Thus, we can restrict the action in Example 4.1.13 to the subsets of a given size $k$:

**Example 4.1.14.** Let $A$ be any set. Let $k \in \mathbb{N}$. Consider the set $\mathcal{P}_k(A)$ of all $k$-element subsets of $A$. The symmetric group $S_A$ acts on $\mathcal{P}_k(A)$ by

$$g \rightharpoonup U = g(U) \qquad \text{for all } g \in S_A \text{ and } U \in \mathcal{P}_k(A).$$

Here is a slightly more exotic action of a symmetric group:

**Example 4.1.15.** Let $A$ be a finite set. Then, the symmetric group $S_A$ acts on the two-element set $\{1,-1\}$ by

$$g \rightharpoonup x = (-1)^g x \qquad \text{for all } g \in S_A \text{ and } x \in \{1,-1\}.$$

In other words, $g \rightharpoonup x = x$ if $g$ is even, and $g \rightharpoonup x = -x$ if $g$ is odd.

*Proof of Example 4.1.15 (sketched).* Consider the map $\rightharpoonup : S_A \times \{1,-1\} \to \{1,-1\}$ defined by

$$g \rightharpoonup x = (-1)^g x \qquad \text{for all } g \in S_A \text{ and } x \in \{1,-1\}.$$

We must prove that this map is a left $S_A$-action on $\{1,-1\}$. In other words, we must prove that it satisfies the two axioms in Definition 4.1.4 **(a)**.

We begin with the first axiom. This axiom claims that $g \rightharpoonup (h \rightharpoonup x) = (gh) \rightharpoonup x$ for all $g, h \in S_A$ and $x \in \{1,-1\}$. To prove this, we fix $g, h \in S_A$ and $x \in \{1,-1\}$. Then, one of the basic properties of signs of permutations shows that $(-1)^{gh} = (-1)^g \cdot (-1)^h$. Now, the definition of $\rightharpoonup$ yields

$$g \rightharpoonup (h \rightharpoonup x) = (-1)^g \underbrace{(h \rightharpoonup x)}_{\substack{=(-1)^h x \\ \text{(by the definition of } \rightharpoonup\text{)}}} = (-1)^g (-1)^h x$$

and

$$(gh) \rightharpoonup x = \underbrace{(-1)^{gh}}_{=(-1)^g \cdot (-1)^h} x = (-1)^g \cdot (-1)^h x = (-1)^g (-1)^h x.$$

Comparing these two equalities, we obtain $g \rightharpoonup (h \rightharpoonup x) = (gh) \rightharpoonup x$. Thus, we have shown that $\rightharpoonup$ satisfies the first axiom in Definition 4.1.4 **(a)**.

The proof of the second axiom is similar but even more trivial (we need the fact that $(-1)^{\mathrm{id}} = 1$). Thus, both axioms are satisfied, so that $\rightharpoonup$ is really a left $S_A$-action. This proves Example 4.1.15. □

**Example 4.1.16.** Let $A$ be any set. Let $n \in \mathbb{N}$. Then, the symmetric group $S_A$ acts on the set

$$A^n = \{\text{all } n\text{-tuples of elements of } A\}$$

by

$$g \rightharpoonup (a_1, a_2, \ldots, a_n) = (g(a_1), g(a_2), \ldots, g(a_n)) \quad \text{for all } g \in S_A \text{ and } (a_1, a_2, \ldots, a_n) \in A^n.$$

This is called *entrywise action* of $S_A$, or the *action on entries*.

For example, for $A = \{a, b, c, d, e\}$ and $n = 3$, we have

$$\mathrm{cyc}_{a,b,c} \rightharpoonup (d, c, a) = (d, a, b).$$

**Example 4.1.17.** Let $A$ be any set. Let $n \in \mathbb{N}$. Then, the symmetric group $S_n$ acts on the set

$$A^n = \{\text{all } n\text{-tuples of elements of } A\}$$

by

$$g \rightharpoonup (a_1, a_2, \ldots, a_n) = \left(a_{g^{-1}(1)}, a_{g^{-1}(2)}, \ldots, a_{g^{-1}(n)}\right) \quad \text{for all } g \in S_n \text{ and } (a_1, a_2, \ldots, a_n) \in A^n.$$

This is called the *place permutation action* of $S_n$, or the *action on places*.

For example, for $A = \{a, b, c, d, e\}$ and $n = 5$, we have

$$\mathrm{cyc}_{1,2,3} \rightharpoonup (a, b, c, d, e) = (c, a, b, d, e) \qquad \text{and}$$
$$t_{2,5} \rightharpoonup (a, b, c, d, e) = (a, e, c, d, b).$$

*Proof of Example 4.1.17.* Let us prove the claim of Example 4.1.17, to illustrate how such claims can be shown. The proof is simple and straightforward, but the road is somewhat slippery due to the many subscripts involved.

We must prove that the map $\rightharpoonup$ introduced in Example 4.1.17 really is a left $S_n$-action on $A^n$. In other words, we must show that it satisfies the two axioms in Definition 4.1.4

**(a)**. The second axiom (unitality) is satisfied for pretty obvious reasons[34], so we focus on verifying the first (associativity). Thus, we need to prove that

$$g \rightharpoonup (h \rightharpoonup x) = (gh) \rightharpoonup x \qquad \text{for all } g, h \in S_n \text{ and } x \in A^n.$$

For this purpose, we fix $g, h \in S_n$ and $x \in A^n$. Write $x$ as $x = (x_1, x_2, \ldots, x_n)$. Thus,

$$h \rightharpoonup x = h \rightharpoonup (x_1, x_2, \ldots, x_n) = \left(x_{h^{-1}(1)}, x_{h^{-1}(2)}, \ldots, x_{h^{-1}(n)}\right)$$

(by the definition of $\rightharpoonup$). Let us denote this $n$-tuple $h \rightharpoonup x$ by $(y_1, y_2, \ldots, y_n)$. Thus,

$$(y_1, y_2, \ldots, y_n) = h \rightharpoonup x = \left(x_{h^{-1}(1)}, x_{h^{-1}(2)}, \ldots, x_{h^{-1}(n)}\right).$$

In other words,

$$y_i = x_{h^{-1}(i)} \qquad \text{for each } i \in [n]. \tag{93}$$

Now, from $h \rightharpoonup x = (y_1, y_2, \ldots, y_n)$, we obtain

$$\begin{aligned} g \rightharpoonup (h \rightharpoonup x) &= g \rightharpoonup (y_1, y_2, \ldots, y_n) \\ &= \left(y_{g^{-1}(1)}, y_{g^{-1}(2)}, \ldots, y_{g^{-1}(n)}\right) \end{aligned} \tag{94}$$

(by the definition of $\rightharpoonup$). However, each $k \in [n]$ satisfies

$$\begin{aligned} y_{g^{-1}(k)} &= x_{h^{-1}(g^{-1}(k))} \qquad \left(\text{by (93), applied to } i = g^{-1}(k)\right) \\ &= x_{(gh)^{-1}(k)} \qquad \left(\text{since } h^{-1}\left(g^{-1}(k)\right) = \underbrace{\left(h^{-1}g^{-1}\right)}_{=(gh)^{-1}}(k) = (gh)^{-1}(k)\right). \end{aligned}$$

Hence,

$$\left(y_{g^{-1}(1)}, y_{g^{-1}(2)}, \ldots, y_{g^{-1}(n)}\right) = \left(x_{(gh)^{-1}(1)}, x_{(gh)^{-1}(2)}, \ldots, x_{(gh)^{-1}(n)}\right).$$

Thus, (94) can be rewritten as

$$g \rightharpoonup (h \rightharpoonup x) = \left(x_{(gh)^{-1}(1)}, x_{(gh)^{-1}(2)}, \ldots, x_{(gh)^{-1}(n)}\right).$$

Comparing this with

$$\begin{aligned} (gh) \rightharpoonup x &= (gh) \rightharpoonup (x_1, x_2, \ldots, x_n) \qquad (\text{since } x = (x_1, x_2, \ldots, x_n)) \\ &= \left(x_{(gh)^{-1}(1)}, x_{(gh)^{-1}(2)}, \ldots, x_{(gh)^{-1}(n)}\right) \qquad (\text{by the definition of } \rightharpoonup), \end{aligned}$$

we obtain $g \rightharpoonup (h \rightharpoonup x) = (gh) \rightharpoonup x$. Thus, we have shown that the map $\rightharpoonup$ satisfies the first axiom (associativity) in Definition 4.1.4 **(a)**. As we said, this completes the proof of Example 4.1.17. □

[34] Indeed, $\mathrm{id} \rightharpoonup (a_1, a_2, \ldots, a_n) = \left(a_{\mathrm{id}^{-1}(1)}, a_{\mathrm{id}^{-1}(2)}, \ldots, a_{\mathrm{id}^{-1}(n)}\right) = (a_1, a_2, \ldots, a_n)$ for all $(a_1, a_2, \ldots, a_n) \in A^n$.

Note that Example 4.1.16 and Example 4.1.17 define actions of two different symmetric groups ($S_A$ and $S_n$) on the same set $A^n$. When $A = [n]$, the two groups agree, but the actions do not.

**Example 4.1.18.** Let $n, m \in \mathbb{N}$. Consider the set $\mathbf{k}^{n\times m}$ of $n \times m$-matrices with entries in $\mathbf{k}$. We use the notation $\left(a_{i,j}\right)_{i\in[n],\ j\in[m]}$ for the $n \times m$-matrix whose $(i,j)$-th entry is $a_{i,j}$ for all $i \in [n]$ and $j \in [m]$.

Then, the Cartesian product $S_n \times S_m$ of the symmetric groups $S_n$ and $S_m$ acts on $\mathbf{k}^{n\times m}$ by permuting rows and columns:

$$(g,h) \rightharpoonup \left(a_{i,j}\right)_{i\in[n],\ j\in[m]} = \left(a_{g^{-1}(i),h^{-1}(j)}\right)_{i\in[n],\ j\in[m]}$$
$$\text{for all } g \in S_n \text{ and } h \in S_m \text{ and } \left(a_{i,j}\right)_{i\in[n],\ j\in[m]} \in \mathbf{k}^{n\times m}.$$

In other words, for any $(g,h) \in S_n \times S_m$ and any matrix $A \in \mathbf{k}^{n\times m}$, the matrix $(g,h) \rightharpoonup A$ is obtained from $A$ by permuting the rows using $g$ and permuting the columns using $h$.

For instance, for $n = 3$ and $m = 3$, we have

$$\left(s_1,\ \mathrm{cyc}_{1,2,3}\right) \rightharpoonup \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = \begin{pmatrix} a_{2,3} & a_{2,1} & a_{2,2} \\ a_{1,3} & a_{1,1} & a_{1,2} \\ a_{3,3} & a_{3,1} & a_{3,2} \end{pmatrix}.$$

Other groups have interesting actions, too:

**Example 4.1.19.** Let $n \in \mathbb{N}$. The group

$$\mathrm{GL}_n(\mathbb{R}) = \{\text{invertible } n \times n\text{-matrices with real entries}\}$$

acts on the set $\mathbb{R}^{n\times 1} = \{\text{column vectors of size } n \text{ over } \mathbb{R}\}$ by

$$A \rightharpoonup v = Av \qquad \text{for all } A \in \mathrm{GL}_n(\mathbb{R}) \text{ and } v \in \mathbb{R}^{n\times 1}.$$

The same applies to any ring instead of $\mathbb{R}$.

And there are many more (see, e.g., [Conrad, "Group actions", §2]).

### 4.1.4. $G$-sets vs. morphisms $G \to S_X$

You can think of $G$-sets as a non-additive analogue of $\mathbf{k}$-modules ("non-additive" because neither $G$ nor the $G$-set are equipped with an addition). The left $G$-action $\rightharpoonup : G \times X \to X$ for a $G$-set is thus an analogue of the scaling map $\mathbf{k} \times V \to V$, $(\lambda, v) \mapsto \lambda v$ that scales vectors in a $\mathbf{k}$-module $V$. The group elements $g \in G$ here correspond to the scalars $\lambda \in \mathbf{k}$.

Just like for $\mathbf{k}$-modules, we can fix a "scalar" (really a group element) $g \in G$ and consider the "scaling" by $g$ as a map from $X$ to $X$:

**Definition 4.1.20.** Let $G$ be a group, and let $X$ be a left $G$-set. Let $g \in G$. Then, the map

$$\begin{aligned} \tau_g : X &\to X, \\ x &\mapsto g \rightharpoonup x \end{aligned}$$

is called the *action* of $g$ on $X$.

This map $\tau_g$ generalizes the translations $\tau_n : \mathbb{Z} \to \mathbb{Z}$ from Example 4.1.2 (which are obtained as a special case for $G = \mathbb{Z}$ and $X = \mathbb{Z}$ and $g = n$, where $G$ acts on $X$ by the left regular $G$-action defined in Definition 4.1.6), and the translations $\tau_g : G \to G$ from Theorem 4.1.3 (which, again, arise by letting $X$ be the left regular $G$-action). This fulfills the "dream" of $g$ to act on something.

Let $G$ be a group, and let $X$ be a left $G$-set. The actions of the elements of $G$ on $X$ behave very much like the elements themselves: For example, the action of $1_G$ is

$$\tau_{1_G} = \mathrm{id}_X \tag{95}$$

(since $\tau_{1_G}(x) = 1 \rightharpoonup x = x$ for each $x \in X$). Moreover, for any two elements $g, h \in G$, the action of the product $gh$ is the composition of the actions of $g$ and $h$: that is, we have

$$\tau_{gh} = \tau_g \tau_h \tag{96}$$

(this is just restating the axiom $(gh) \rightharpoonup x = g \rightharpoonup (h \rightharpoonup x)$ in Definition 4.1.4 **(a)**). This easily yields that the action $\tau_g$ of any element $g \in G$ is invertible, with inverse $\tau_{g^{-1}}$ (since (96) yields $\tau_{gg^{-1}} = \tau_g \tau_{g^{-1}}$ and thus $\tau_g \tau_{g^{-1}} = \tau_{gg^{-1}} = \tau_{1_G} = \mathrm{id}_X$ (by (95)), and similarly $\tau_{g^{-1}} \tau_g = \mathrm{id}$). Hence, this action $\tau_g$ is a permutation of $X$, and therefore belongs to the symmetric group $S_X$. We thus obtain a map

$$\begin{aligned} \rho : G &\to S_X, \\ g &\mapsto \tau_g \end{aligned}$$

(which sends each $g \in G$ to its action on $X$). This map $\rho$ is furthermore a group morphism (by (96) and (95)).

**Definition 4.1.21.** We shall call this morphism $\rho$ the *curried form* of the left $G$-action $\rightharpoonup$ on $X$.

We note that we can reconstruct the left $G$-action $\rightharpoonup$ on $X$ from this morphism $\rho$, since all $g \in G$ and $x \in X$ satisfy

$$g \rightharpoonup x = \underbrace{\tau_g}_{=\rho(g)}(x) = (\rho(g))(x).$$

Thus, the morphism $\rho$ encodes the left $G$-action $\rightharpoonup$.

Note that $\rho$ is not always injective. For example, if $X$ has a trivial $G$-action (as defined in Example 4.1.10), then $\rho(g) = \mathrm{id}_X$ for each $g \in G$.

Thus, we have transformed each left $G$-set $X$ into a morphism $\rho : G \to S_X$.

Let us now forget that we fixed $X$, and work in the opposite direction: If we are given a set $X$ (without a $G$-action provided) and a group morphism $\rho : G \to S_X$, then we can define a left $G$-action on $X$ by

$$g \rightharpoonup x = (\rho(g))(x) \qquad \text{for all } g \in G \text{ and } x \in X.$$

**Definition 4.1.22.** We shall call this $G$-action the *uncurried form* of $\rho$.

Altogether, we thus obtain the following theorem:

**Theorem 4.1.23.** Let $G$ be a group. Let $X$ be a set. Then, a left $G$-action on $X$ is "the same as" a group morphism from $G$ to $S_X$. To be more precise:

**(a)** If $\rightharpoonup$ is a left $G$-action on $X$, then the curried form of $\rightharpoonup$ is a group morphism from $G$ to $S_X$.

**(b)** If $\rho$ is a group morphism from $G$ to $S_X$, then the uncurried form of $\rho$ is a left $G$-action on $X$.

**(c)** Currying and uncurrying are mutually inverse: I.e., the uncurried form of the curried form of a left $G$-action $\rightharpoonup$ is $\rightharpoonup$ itself, and conversely, the curried form of the uncurried form of a group morphism $\rho : G \to S_X$ is $\rho$ itself.

*Proof sketch.* Straightforward. (Part **(a)** was essentially proved above.) □

**Remark 4.1.24.** The words "curried" and "uncurried" refer to the general notions of *currying* and *uncurrying* in computer science. Namely, *currying* is the operation that transforms an arbitrary map $f : A \times B \to C$ with two inputs into a map

$$\widetilde{f} : A \to \{\text{maps from } B \text{ to } C\},$$
$$a \mapsto (\text{the map from } B \text{ to } C \text{ that sends each } b \text{ to } f(a,b)).$$

(In essence, $\widetilde{f}$ is a version of $f$ that takes the two inputs of $f$ "one at a time".) Conversely, *uncurrying* is the operation that transforms a map $g : A \to \{\text{maps from } B \text{ to } C\}$ (whose outputs themselves are maps) into a two-input map

$$g' : A \times B \to C,$$
$$(a,b) \mapsto (g(a))(b).$$

It is easy to see that currying and uncurrying are mutually inverse operations. The curried form of a left $G$-action $\rightharpoonup$ is exactly the result of currying $\rightharpoonup$, except that its target is not $\{\text{maps from } X \text{ to } X\}$ but rather the symmetric group $S_X$. Likewise, the uncurried form of a group morphism $\rho : G \to S_X$ is the result of uncurrying $\rho$, where we treat $\rho$ as a map from $G$ to $\{\text{maps from } X \text{ to } X\}$.

Theorem 4.1.23 can be extended to the case when $G$ is just a monoid (rather than a group), provided that we replace the symmetric group $S_X$ by the "symmetric monoid" $\{\text{all maps from } X \text{ to } X\}$.

**Example 4.1.25.** Consider the left regular $G$-action (as defined in Example 4.1.6). Its curried form is just the morphism $G \to S_G$ from Cayley's theorem (Theorem 4.1.3).

**Example 4.1.26.** Let $A$ be any set. Consider the natural action of $S_A$ on $A$ (as defined in Example 4.1.12). Its curried form is the identity morphism $S_A \to S_A$.

**Example 4.1.27.** Let $A$ be a finite set. Consider the left $S_A$-action on $\{1, -1\}$ given by

$$g \rightharpoonup x = (-1)^g x \qquad \text{for all } g \in S_A \text{ and } x \in \{1, -1\}.$$

The curried form of this left $S_A$-action $\rightharpoonup$ is the map

$$\begin{aligned} S_A &\to S_{\{1,-1\}}, \\ g &\mapsto \left(\text{multiplication by } (-1)^g\right). \end{aligned}$$

This is "more or less" the sign homomorphism

$$\begin{aligned} S_A &\to \{1, -1\}, \\ g &\mapsto (-1)^g, \end{aligned}$$

except that instead of the actual numbers 1 and $-1$ it produces the multiplications by these numbers (but this is the same thing up to isomorphism).

### 4.1.5. $G$-equivariant maps aka $G$-set morphisms

Whenever you define an algebraic object, the next logical step is to define morphisms between these objects. For $G$-sets, it is pretty clear how to do this:

**Definition 4.1.28.** Let $X$ and $Y$ be two left $G$-sets. Then, a *$G$-set morphism* (aka a *morphism of $G$-sets*, aka a *$G$-equivariant map*) means a map $f : X \to Y$ with the property that all $g \in G$ and $x \in X$ satisfy

$$f(g \rightharpoonup x) = g \rightharpoonup f(x). \tag{97}$$

This requirement is similar to the axiom $f(\lambda x) = \lambda f(x)$ in the definition of a $\mathbf{k}$-module morphism (aka $\mathbf{k}$-linear map).

**Example 4.1.29.** Let $A$ be any set. Let $n \in \mathbb{N}$. Consider $A^n$ as a left $S_n$-set via the place permutation action of $S_n$ on $A^n$ (defined in Example 4.1.17). Also consider $A$ as a left $S_n$-set via the trivial $S_n$-action on $A$ (as defined in Example 4.1.10, i.e., defined by $g \rightharpoonup x = x$ for all $g \in S_n$ and $x \in A$). Then:

**(a)** The map

$$f_1 : A \to A^n,$$
$$a \mapsto \left( \underbrace{a, a, \ldots, a}_{n \text{ times}} \right)$$

is an $S_n$-set morphism. Indeed, all $g \in S_n$ and $x \in A$ satisfy

$$f_1 \left( \underbrace{g \rightharpoonup x}_{=x} \right) = f_1(x) = \left( \underbrace{x, x, \ldots, x}_{n \text{ times}} \right) \qquad \text{and}$$
$$g \rightharpoonup f_1(x) = g \rightharpoonup \left( \underbrace{x, x, \ldots, x}_{n \text{ times}} \right) = \left( \underbrace{x, x, \ldots, x}_{n \text{ times}} \right)$$

(by the definition of $\rightharpoonup$) and therefore (by comparing these two equalities) $f_1(g \rightharpoonup x) = g \rightharpoonup f_1(x)$, so that the requirement (97) is satisfied for $f = f_1$.

**(b)** Assume that $A$ is an additive abelian monoid (i.e., a set equipped with a binary operation $+$ that is associative and commutative and has a neutral element). For instance, $A$ can be $\mathbb{N}$ or any ring like $\mathbb{Z}$ or $\mathbb{Q}$.

Then, the map

$$f_2 : A^n \to A,$$
$$(a_1, a_2, \ldots, a_n) \mapsto a_1 + a_2 + \cdots + a_n$$

is also an $S_n$-set morphism. Indeed, the requirement (97) (applied to $x = (a_1, a_2, \ldots, a_n)$) is here just saying that

$$a_{g^{-1}(1)} + a_{g^{-1}(2)} + \cdots + a_{g^{-1}(n)} = a_1 + a_2 + \cdots + a_n$$

for all $g \in S_n$ and $(a_1, a_2, \ldots, a_n) \in A^n$; but this is a known property of abelian monoids ("generalized commutativity").

**(c)** The map

$$f_3 : A^n \to A,$$
$$(a_1, a_2, \ldots, a_n) \mapsto a_1$$

is **not** an $S_n$-set morphism (unless $n \leq 1$ or $|A| \leq 1$). Indeed, if we pick any $x = (x_1, x_2, \ldots, x_n) \in A^n$ with $x_1 \neq x_2$, then we have $f_3(s_1 \rightharpoonup x) = f_3(x_2, x_1, x_3, x_4, \ldots, x_n) = x_2$ but $s_1 \rightharpoonup f_3(x) = f_3(x) = x_1 \neq x_2$.

**(d)** The map

$$f_4 : A^n \to A^n,$$
$$(a_1, a_2, \ldots, a_n) \mapsto (a_n, a_{n-1}, \ldots, a_1)$$

(which reverses each $n$-tuple given to it as an input) is not an $S_n$-set morphism (unless $n \leq 2$ or $|A| \leq 1$). Indeed, if we pick any $x = (x_1, x_2, \ldots, x_n) \in A^n$ with $x_1 \neq x_2$, then we have

$$f_4(s_1 \rightharpoonup x) = f_4(x_2, x_1, x_3, x_4, \ldots, x_n) = (x_n, x_{n-1}, \ldots, x_3, x_1, x_2) \quad \text{but}$$
$$s_1 \rightharpoonup f_4(x) = s_1 \rightharpoonup (x_n, x_{n-1}, \ldots, x_1) = (x_{n-1}, x_n, x_{n-2}, x_{n-3}, \ldots, x_1),$$

which are different.

**Example 4.1.30.** Let $X$ and $Y$ be two left $G$-sets with trivial $G$-actions (i.e., we assume that $g \rightharpoonup x = x$ for all $g \in G$ and $x \in X$, and that $g \rightharpoonup y = y$ for all $g \in G$ and $y \in Y$). Then, **any** map $f : X \to Y$ is $G$-equivariant.

*Proof.* Let $f : X \to Y$ be any map. For each $g \in G$ and each $x \in X$, we then have $g \rightharpoonup f(x) = f(x)$ (since the $G$-action on $Y$ is trivial) and $g \rightharpoonup x = x$ (since the $G$-action on $X$ is trivial) and thus

$$f(g \rightharpoonup x) = f(x) = g \rightharpoonup f(x).$$

But this means precisely that $f$ is $G$-equivariant. This proves Example 4.1.30. □

As usual, the nicest morphisms are isomorphisms:

**Definition 4.1.31.** Let $X$ and $Y$ be two left $G$-sets. Then, a *$G$-set isomorphism* (aka *isomorphism of $G$-sets*) means a $G$-set morphism $f : X \to Y$ whose inverse exists and is again a $G$-set morphism.

**Proposition 4.1.32.** The "is again a $G$-set morphism" part in Definition 4.1.31 is redundant (i.e., follows from the other requirements). In other words, if a $G$-set morphism $f : X \to Y$ is invertible, then its inverse is again a $G$-set morphism.

*Proof.* This is proved in the same straightforward way as the analogous claims about groups, rings, $\mathbf{k}$-modules and many other algebraic structures. □

**Definition 4.1.33.** Let $X$ and $Y$ be two left $G$-sets. Then, we say that $X$ and $Y$ are *isomorphic* (this is written $X \cong Y$) if there exists a $G$-set isomorphism from $X$ to $Y$.

Here are some examples of $G$-set isomorphisms:

**Example 4.1.34.** We have $G/G \cong \{1\}$ as left $G$-sets (where $G/G$ is a left $G$-set as defined in Example 4.1.11, whereas $\{1\}$ has a trivial $G$-action as defined in Example 4.1.10). Indeed, there is a $G$-set isomorphism

$$\begin{aligned} G/G &\to \{1\}, \\ gG &\mapsto 1. \end{aligned}$$

**Example 4.1.35.** We have $G/\{1_G\} \cong G$ as left $G$-sets (where $G/\{1_G\}$ is a left $G$-set as defined in Example 4.1.11, whereas $G$ has the left regular $G$-action as defined in Example 4.1.6). Indeed, there is a $G$-set isomorphism

$$\begin{aligned} G/\{1_G\} &\to G, \\ g\{1_G\} &\mapsto g. \end{aligned}$$

**Example 4.1.36.** Let $A$ be any set. Consider the left $S_A$-set $\mathcal{P}(A)$ defined in Example 4.1.13. Then, the map

$$\begin{aligned} \mathcal{P}(A) &\to \mathcal{P}(A), \\ U &\mapsto A \setminus U \end{aligned}$$

(which sends each subset of $A$ to its complement) is an $S_A$-set isomorphism.

**Example 4.1.37.** Let $A$ be any set. Consider the left $S_A$-set $\mathcal{P}_k(A)$ defined in Example 4.1.13 for each $k \in \mathbb{N}$. Also, consider the left $S_A$-set $A$ defined in Example 4.1.12. Then, we have $\mathcal{P}_1(A) \cong A$ as $S_A$-sets. To be more specific: the map

$$\begin{aligned} A &\to \mathcal{P}_1(A), \\ a &\mapsto \{a\} \end{aligned}$$

is an $S_A$-set isomorphism.

**Example 4.1.38.** Consider the left $S_n$-action on $\mathcal{P}([n])$ defined in Example 4.1.13 (for $A = [n]$), and consider the left $S_n$-action on $\{0,1\}^n$ is defined as in Example 4.1.17. Then, we have $\mathcal{P}([n]) \cong \{0,1\}^n$ as left $S_n$-sets. Indeed, there is a $G$-set isomorphism

$$\begin{aligned} \{0,1\}^n &\to \mathcal{P}([n]), \\ (a_1, a_2, \ldots, a_n) &\mapsto \{i \in [n] \mid a_i = 1\}. \end{aligned}$$

**Exercise 4.1.1.** [1] Prove the latter claim about the map $\{0,1\}^n \to \mathcal{P}([n])$. You can take for granted the well-known fact that this map is well-defined and bijective.

### 4.1.6. Combining $G$-sets

We can think of $G$-sets as a set-theoretical analogue of vector spaces (with the map $\rightharpoonup : G \times X \to X$ being analogous to the scaling map $\mathbf{k} \times V \to V$ on a $\mathbf{k}$-vector space). Two vector spaces can be combined in two important ways: tensor products and direct products. The analogues of these two combinations for $G$-sets are Cartesian products and disjoint unions.

We begin with the Cartesian products, since they are the easiest to define (even though the disjoint unions are easier to visualize):

**Definition 4.1.39. (a)** If $X$ and $Y$ are two left $G$-sets, then the Cartesian product $X \times Y$ also becomes a left $G$-set by setting

$$g \rightharpoonup (x,y) = (g \rightharpoonup x,\ g \rightharpoonup y) \qquad \text{for all } g \in G \text{ and } x \in X \text{ and } y \in Y.$$

**(b)** More generally, for any left $G$-sets $X_1, X_2, \ldots, X_n$, we make the Cartesian product $X_1 \times X_2 \times \cdots \times X_n$ into a left $G$-set by setting

$$g \rightharpoonup (x_1, x_2, \ldots, x_n) = (g \rightharpoonup x_1,\ g \rightharpoonup x_2,\ \ldots,\ g \rightharpoonup x_n)$$
$$\text{for all } g \in G \text{ and } (x_1, x_2, \ldots, x_n) \in X_1 \times X_2 \times \cdots \times X_n.$$

This $G$-set is called the *Cartesian product* (or *direct product*) of these left $G$-sets $X_1, X_2, \ldots, X_n$. The $G$-action on this $G$-set is called the *entrywise action* (since $g$ acts on each entry of a $n$-tuple independently).

**(c)** For any left $G$-set $X$ and any $n \in \mathbb{N}$, we let $X^n$ denote the Cartesian product $\underbrace{X \times X \times \cdots \times X}_{n \text{ times}}$. This is called the *$n$-th Cartesian power* of $X$.

**Example 4.1.40.** Let $A$ be any set, and let $n \in \mathbb{N}$. As we saw in Example 4.1.12, the set $A$ becomes a left $S_A$-set using the natural action (given by $g \rightharpoonup a = g(a)$ for all $g \in S_A$ and $a \in A$). The $n$-th Cartesian power $A^n$ of this $S_A$-set $A$ (as defined in Definition 4.1.39 **(c)**) is precisely the $S_A$-set $A^n$ introduced in Example 4.1.16 (i.e., the one where a permutation $g \in S_A$ is applied to each entry of the $n$-tuple independently).

Now let us move on to the disjoint unions.

Recall that the *disjoint union* $X \sqcup Y$ of two (not necessarily disjoint) sets $X$ and $Y$ is defined by

$$X \sqcup Y := (X \times \{1\}) \cup (Y \times \{2\}).$$

Thus, it contains a "clone" $(x,1)$ of each element $x \in X$ as well as a "clone" $(y,2)$ of each element $y \in Y$. The "clones" are pairs, whose second entries 1 and 2 have been chosen to disambiguate the "clones" coming from $X$ from the "clones" coming from $Y$ (so that, even if the sets $X$ and $Y$ have an element $z$ in common, the two "clones" $(z,1)$ and $(z,2)$ will be distinguishable in $X \sqcup Y$). When the sets $X$ and $Y$ are themselves disjoint, we can (by abuse of notation) identify the clones $(x,1)$ and $(y,2)$ with the elements $x$ and $y$ themselves, which allows us to pretend that $X \sqcup Y$ is the literal union $X \cup Y$ in this case.

More generally, the *disjoint union* $X_1 \sqcup X_2 \sqcup \cdots \sqcup X_n$ of $n$ (not necessarily disjoint) sets $X_1, X_2, \ldots, X_n$ is defined by

$$\begin{aligned} X_1 \sqcup X_2 \sqcup \cdots \sqcup X_n &:= (X_1 \times \{1\}) \cup (X_2 \times \{2\}) \cup \cdots \cup (X_n \times \{n\}) \\ &= \{(x,i) \mid i \in [n] \text{ and } x \in X_i\}. \end{aligned}$$

It contains a "clone" $(x,i)$ of each element $x \in X_i$ for each $i \in [n]$, and again the second entries of these "clones" allow us to distinguish "clones" coming from different sets.

We can equip disjoint unions of $G$-sets with a canonical $G$-set structure:

**Definition 4.1.41.** **(a)** If $X$ and $Y$ are two left $G$-sets, then the disjoint union

$$X \sqcup Y = (X \times \{1\}) \cup (Y \times \{2\})$$

also becomes a left $G$-set by setting

$$g \rightharpoonup (x,1) = \left( \underbrace{g \rightharpoonup x}_{G\text{-action on } X}, \ 1 \right) \qquad \text{for all } g \in G \text{ and } x \in X, \qquad \text{and}$$

$$g \rightharpoonup (y,2) = \left( \underbrace{g \rightharpoonup y}_{G\text{-action on } Y}, \ 2 \right) \qquad \text{for all } g \in G \text{ and } y \in Y.$$

**(b)** More generally, for any left $G$-sets $X_1, X_2, \ldots, X_n$, we make the disjoint union $X_1 \sqcup X_2 \sqcup \cdots \sqcup X_n$ into a left $G$-set by setting

$$g \rightharpoonup (x,i) = \left( \underbrace{g \rightharpoonup x}_{G\text{-action on } X_i}, \ i \right) \qquad \text{for all } g \in G,\ i \in [n] \text{ and } x \in X_i.$$

This $G$-set is called the *disjoint union* (or *direct sum*) of these left $G$-sets $X_1, X_2, \ldots, X_n$.

**(c)** For any left $G$-set $X$ and any $n \in \mathbb{N}$, we let $X^{\sqcup n}$ denote the disjoint union $\underbrace{X \sqcup X \sqcup \cdots \sqcup X}_{n \text{ times}}$. This is called the *$n$-th multiple* of $X$.

We note a theorem that we will not need, but is worth knowing:

**Theorem 4.1.42.** Any finite left $G$-set is isomorphic to a disjoint union of $G/H$'s for $H$ being subgroups of $G$.

*Proof sketch.* Combine §1 and Theorem 3.6 of [Conrad, "Transitive group actions"]. Specifically, the former shows that any finite left $G$-set is isomorphic to a disjoint union of transitive $G$-sets, whereas the latter reveals that each transitive $G$-set is isomorphic to $G/H$ for some subgroup $H$ of $G$. □

This is also true for infinite $G$-sets, if we define infinite disjoint unions in the obvious way.

### 4.1.7. $G$-subsets and quotients

Just like $\mathbf{k}$-modules have $\mathbf{k}$-submodules (and likewise for groups, rings, etc.), $G$-sets have subobjects called "$G$-subsets":

**Definition 4.1.43.** Let $X$ be a left $G$-set.

**(a)** A *$G$-subset* (aka *sub-$G$-set*) of $X$ means a subset $Y$ of $X$ such that

$$\text{all } g \in G \text{ and } y \in Y \text{ satisfy } g \rightharpoonup y \in Y.$$

**(b)** If $Y$ is a $G$-subset of $X$, then $Y$ becomes a left $G$-set in its own right, by simply restricting the left $G$-action on $X$ to $Y$. (That is, the left $G$-action on $Y$ is just the restriction of the map $\rightharpoonup : G \times X \to X$ to $G \times Y$.) Clearly, the canonical inclusion map $i : Y \to X$ (sending each $y \in Y$ to $y$) is a $G$-set morphism.

**Example 4.1.44.** Let $A$ be any set. Consider $A^3$ as a left $S_3$-set via the place permutation action of $S_3$ on $A^3$ (defined in Example 4.1.17, for $n = 3$). Then, the subset

$$\left\{(a,b,c) \in A^3 \mid a = b = c\right\} = \{(a,a,a) \mid a \in A\}$$

of $A^3$ is an $S_3$-subset of $A^3$. Moreover, if $A$ is an additive group, then the subset

$$\left\{(a,b,c) \in A^3 \mid a+b+c = 0\right\}$$

of $A^3$ is an $S_3$-subset of $A^3$ as well (since the sum $a+b+c$ does not change when we permute the entries of the triple $(a,b,c)$). But the subset

$$\left\{(a,b,c) \in A^3 \mid a \neq b \text{ and } b \neq c\right\}$$

is not an $S_3$-subset of $A^3$ (unless $|A| \leq 1$), since (for any two distinct elements $u$ and $v$ of $A$) the triple $(u,v,u)$ belongs to it but the triple $s_1 \rightharpoonup (u,v,u) = (v,u,u)$ does not. The reader can easily generalize these examples to $A^n$ instead of $A^3$.

Just like $\mathbf{k}$-modules have quotient modules, $G$-sets have quotient $G$-sets. However, the quotient is not a quotient by a $G$-subset, but rather a quotient by a "$G$-invariant equivalence relation". To explain this, let me first recall how quotients by equivalence relations are defined for usual sets (not $G$-sets): If $\sim$ is an equivalence relation on a given set $X$, then the *quotient set* of $X$ by this relation $\sim$ is defined to be the set of all $\sim$-equivalence classes on $X$. This quotient set is denoted by $X/\sim$. For each $x \in X$, we denote the $\sim$-equivalence class that contains $x$ as $[x]_\sim$. (Thus, $X/\sim = \{[x]_\sim \mid x \in X\}$.) For example, if $n$ is an integer, then the residue classes of integers modulo $n$ are the equivalence classes of the relation "congruent modulo $n$", and thus their set $\mathbb{Z}/n$ is the quotient set $\mathbb{Z}/\overset{n}{\equiv}$, where $\overset{n}{\equiv}$ is this relation.

Now, we can define quotient $G$-sets:

**Definition 4.1.45.** Let $\sim$ be an equivalence relation on a left $G$-set $X$.

**(a)** We say that this relation $\sim$ is *$G$-invariant* if it has the property that for every $x, y \in X$ and every $g \in G$, the implication

$$(x \sim y) \implies (g \rightharpoonup x \sim g \rightharpoonup y) \qquad \text{holds.}$$

In words: The relation $\sim$ is $G$-invariant if and only if any two $\sim$-equivalent elements of $X$ remain $\sim$-equivalent to each other under the action of any $g \in G$. (This does **not** mean that $x \sim g \rightharpoonup x$ for all $x \in X$.)

**(b)** If $\sim$ is a $G$-invariant equivalence relation on $X$, then the quotient set

$$X/\sim \ = \{\text{all } \sim\text{-equivalence classes on } X\}$$

becomes a left $G$-set, with the $G$-action defined by

$$g \rightharpoonup U = g(U) \qquad \text{for any } \sim\text{-equivalence class } U.$$

Equivalently, we can rewrite this definition of $\rightharpoonup$ as

$$g \rightharpoonup [x]_\sim = [g \rightharpoonup x]_\sim \qquad \text{for any } \sim\text{-equivalence class } [x]_\sim.$$

The map

$$\begin{aligned} X &\to X/\sim, \\ x &\mapsto [x]_\sim \end{aligned}$$

(which sends each element of $X$ to its $\sim$-equivalence class) is thus a $G$-set morphism.

**Exercise 4.1.2.** [1] Prove all claims implicitly made in Definition 4.1.45 (i.e., that $X/\sim$ really becomes a left $G$-set, and that the map is a $G$-set morphism).

**Example 4.1.46.** Let $A$ be any set. Let $n \in \mathbb{N}$. Consider $A^n$ as a left $S_A$-set via the entrywise action (defined in Example 4.1.16). We can define an equivalence relation $\sim$ on $A^n$ by setting

$$(a_1, a_2, \ldots, a_n) \sim \left(a_{\sigma(1)}, a_{\sigma(2)}, \ldots, a_{\sigma(n)}\right) \quad \text{for all } (a_1, a_2, \ldots, a_n) \in A^n \text{ and } \sigma \in S_n.$$

In other words, we declare two $n$-tuples in $A^n$ to be equivalent if they can be obtained from each other by permuting the entries. (For example, $(2,4,2,5) \sim (4,5,2,2)$ but not $\sim (1,4,1,5)$.)

It is easy to see that this relation $\sim$ is $S_A$-invariant, so that the quotient set $A^n/\sim$ becomes an $S_A$-set. The elements of this quotient set $A^n/\sim$ are $\sim$-equivalence classes, i.e., $n$-tuples "up to permutation". They are known as *unordered $n$-tuples* of elements of $A$. Thus, we have obtained an $S_A$-action on the set of unordered $n$-tuples of elements of $A$.

### 4.1.8. Restriction

Any $\mathbb{C}$-vector space is automatically an $\mathbb{R}$-vector space, since $\mathbb{R}$ is a subfield of $\mathbb{C}$. More generally, if $\mathbf{l}$ is a subring of $\mathbf{k}$, then any $\mathbf{k}$-module automatically becomes an $\mathbf{l}$-module. Something analogous holds for $G$-sets and subgroups:

**Definition 4.1.47.** Let $H$ be a subgroup of the group $G$. Then, every left $G$-set $X$ automatically becomes a left $H$-set by setting

$$h \rightharpoonup x = h \rightharpoonup x \qquad \text{for each } h \in H \text{ and } x \in X$$

(where the $\rightharpoonup$ on the right hand side is the left $G$-action that we are given, whereas the $\rightharpoonup$ on the left hand side is the left $H$-action that we are defining). This left $H$-set is called the *restriction* of the original $G$-set to the subgroup $H$.

This notion of restriction can be generalized:

**Definition 4.1.48.** Let $G$ and $H$ be any two groups. Let $\alpha : H \to G$ be a group morphism. Then, every left $G$-set $X$ automatically becomes a left $H$-set by setting

$$h \rightharpoonup x = \alpha(h) \rightharpoonup x \qquad \text{for each } h \in H \text{ and } x \in X.$$

This left $H$-set is called the *restriction* of the original $G$-set along the morphism $\alpha$.

When $H$ is a subgroup of $G$, the canonical inclusion map $i : H \to G,\ h \mapsto h$ is a group morphism, and then the restriction of a left $G$-set along this morphism $i$ is precisely its restriction to the subgroup $H$ (as defined in Definition 4.1.47). Thus, Definition 4.1.48 generalizes Definition 4.1.47.

**Example 4.1.49.** Let $X$ be any set. The trivial $G$-action on $X$ (that is, the left $G$-action defined by $g \rightharpoonup x = x$, as in Example 4.1.10) is the restriction of the trivial $\{1\}$-action on $X$ along the morphism

$$\begin{aligned} \pi : G &\to \{1\}, \\ g &\mapsto 1. \end{aligned}$$

**Exercise 4.1.3.** $\boxed{2}$ Let $G$ be a group. Let $X$ be a set. Prove the following:

**(a)** If $\rightharpoonup : G \times X \to X$ is a left $G$-action on $X$, then the map $\leftharpoonup : X \times G \to X$ defined by

$$x \leftharpoonup g = g^{-1} \rightharpoonup x \qquad \text{for all } g \in G \text{ and } x \in X$$

is a right $G$-action on $X$. Thus, any left $G$-action on $X$ can be converted into a right $G$-action on $X$.

**(b)** If $\leftharpoonup : X \times G \to X$ is a right $G$-action on $X$, then the map $\rightharpoonup : G \times X \to X$ defined by

$$g \rightharpoonup x = x \leftharpoonup g^{-1} \qquad \text{for all } g \in G \text{ and } x \in X$$

is a left $G$-action on $X$. Thus, any right $G$-action on $X$ can be converted into a left $G$-action on $X$.

**(c)** These two conversions are mutually inverse: I.e.,

- if we convert a left $G$-action into a right $G$-action and then convert the resulting right $G$-action back into a left $G$-action, then we recover the original left $G$-action,
- and likewise for right $G$-actions.

### 4.1.9. Understanding $G$-actions for cyclic groups $G$

Recall that a group is called *cyclic* if it can be generated by a single element. From basic abstract algebra, we know what the cyclic groups are (up to isomorphism):

- For each positive integer $n$, there is a cyclic group $C_n$ of size $n$. This group is generated by an element $g$ that satisfies $g^n = 1$; thus it consists of $n$ distinct elements $g^0, g^1, \ldots, g^{n-1}$ with $g^n = g^0 = 1$. Of course, this group is nothing but the additive abelian group $(\mathbb{Z}/n, +, 0)$, rewritten multiplicatively (so we write $g^i$ for the residue class $\bar{i} \in \mathbb{Z}/n$).

- There is also an infinite cyclic group $C_\infty$, which is generated by an element $g$ that satisfies no relations. This is just the additive abelian group $(\mathbb{Z}, +, 0)$, rewritten multiplicatively (so we write $g^i$ for the integer $i$).

Any cyclic group is isomorphic to one of the $C_n$'s or $C_\infty$. The cyclic groups have the reputation of being the simplest groups to study, and so we shall now try to better understand their actions.

We ask ourselves: What "is" a $G$-set "really"? In other words, what is the easiest way to view a $G$-action? For general $G$, we cannot improve on the definition of a $G$-action (or on the equivalent description in Theorem 4.1.23). However, when $G$ is a cyclic group, we can reframe this definition in a much more compact form.

We begin with the most trivial example: that of the cyclic group $C_1$, also known as the trivial group $\{1\}$, also isomorphic to the symmetric groups $S_0$ and $S_1$.

**Example 4.1.50.** Consider the trivial group $\{1\}$.

What is a left $\{1\}$-set? In other words, what does a left $\{1\}$-action on a set $X$ mean? It means a map $\rightharpoonup : \{1\} \times X \to X$ that satisfies the two axioms from Definition 4.1.4 **(a)**. But the second axiom is saying that $1 \rightharpoonup x = x$ for all $x \in X$ (which uniquely determines the map $\rightharpoonup$), whereas the first axiom is just saying that $1 \rightharpoonup (1 \rightharpoonup x) = (1 \cdot 1) \rightharpoonup x$ (since the only $g \in \{1\}$ is 1), which follows automatically from the second axiom. Thus, a left $\{1\}$-action $\rightharpoonup$ on $X$ is uniquely determined by the rule

$$1 \rightharpoonup x = x \qquad \text{for all } x \in X,$$

and moreover, the map $\rightharpoonup$ defined by this rule really is a left $\{1\}$-action on $X$. Hence, any set $X$ has a unique left $\{1\}$-action on it. Consequently, a left $\{1\}$-set is just a set equipped with this unique left $\{1\}$-action. Roughly speaking, this means that a left $\{1\}$-set is just a set (the $\{1\}$-action "comes for free").

This shows that $G$-sets are a generalization of sets. Nice, but not very exciting. But this is what we get by looking at the trivial group. Let us now move onward to slightly more interesting groups.

**Example 4.1.51.** Consider the cyclic group $C_2 = \{1, g\}$ (also known as the symmetric group $S_2 = \{1, s_1\}$ if we identify the generator $g$ with $s_1$).

What is a left $C_2$-set? In other words, what does a left $C_2$-action on a set $X$ mean? To define such an action, we must specify $1 \rightharpoonup x$ and $g \rightharpoonup x$ for each $x \in X$. Again, $1 \rightharpoonup x$ is uniquely determined by the second axiom in Definition 4.1.4 **(a)**, which requires $1 \rightharpoonup x$ to be $x$. As for $g \rightharpoonup x$, we know by the first axiom that

$$g \rightharpoonup (g \rightharpoonup x) = \underbrace{(gg)}_{=g^2=1} \rightharpoonup x = 1 \rightharpoonup x = x.$$

Thus, the map

$$\begin{aligned} \tau_g : X &\to X, \\ x &\mapsto g \rightharpoonup x \end{aligned}$$

must be an involution. Conversely, any involution of $X$ can be used as $\tau_g$ and thus defines a $C_2$-action on $X$ (why?). So a $C_2$-action on $X$ is "the same as" an involution of $X$. And thus a left $C_2$-set is just a set with an involution on it.

This can also be seen using the curried form (i.e., using Theorem 4.1.23): In its curried form, a $C_2$-action on $X$ is a group morphism $\rho : C_2 \to S_X$. Such a group morphism is uniquely determined by its value $\rho(g)$ (since $\rho(1)$ must be $\mathrm{id} \in S_X$), and this value must be an involution since $(\rho(g))^2 = \rho(g^2) = \rho(1) = \mathrm{id}$. Conversely, for any involution $w \in S_X$, we can define a group morphism $\rho : C_2 \to S_X$ by setting $\rho(1) = \mathrm{id}$ and $\rho(g) = w$, and this defines a left $C_2$-action on $X$.

So a left $C_2$-set is just a set $X$ with an involution, i.e., with a permutation $w \in S_X$ satisfying $w^2 = \mathrm{id}$. This generalizes to arbitrary cyclic groups $C_n$:

**Theorem 4.1.52.** Let $n$ be a positive integer. Let $C_n$ be the cyclic group of size $n$, with generator $g$ (so that $C_n = \{g^0, g^1, \ldots, g^{n-1}\}$ and $g^n = g^0 = 1$). Let $X$ be a set. Then, a left $C_n$-action on $X$ is "the same as" a permutation $w \in S_X$ satisfying $w^n = \mathrm{id}$. To be more precise:

**(a)** If $\rightharpoonup$ is a left $C_n$-action on $X$, then the map

$$\begin{aligned} \tau_g : X &\to X, \\ x &\mapsto g \rightharpoonup x \end{aligned}$$

is a permutation $w \in S_X$ satisfying $w^n = \mathrm{id}$.

**(b)** Conversely, if $w \in S_X$ is a permutation satisfying $w^n = \mathrm{id}$, then we can define a left $C_n$-action on $X$ by setting

$$g^i \rightharpoonup x = w^i(x) \qquad \text{for all } i \in \mathbb{Z} \text{ and } x \in X.$$

**(c)** These two conversions (from a left $C_n$-action to a permutation $w$, and vice versa) are mutually inverse.

*Proof sketch.* Much of this is straightforward, so I only show some milestones.

**(a)** Let $\rightharpoonup$ be a left $C_n$-action on $X$. It suffices to show that $\tau_g^n = \mathrm{id}$.

Consider the curried form $\rho : C_n \to S_X$ of the left $C_n$-action $\rightharpoonup$. Then, $\rho$ is a group morphism. Hence, $\rho(g^n) = (\rho(g))^n$. In view of $\rho(g) = \tau_g$ (by the definition of $\rho$) and $g^n = 1$, we can rewrite this as $\rho(1) = \tau_g^n$. However, $\rho(1) = \mathrm{id}$ (since $\rho$ is a group morphism). Comparing these two equalities, we obtain $\tau_g^n = \mathrm{id}$. Thus, Theorem 4.1.52 **(a)** is proved.

**(b)** Let $w \in S_X$ be a permutation satisfying $w^n = \mathrm{id}$. We want to define a left $C_n$-action on $X$ by setting

$$g^i \rightharpoonup x = w^i(x) \qquad \text{for all } i \in \mathbb{Z} \text{ and } x \in X.$$

First, we need to show that this map $\rightharpoonup$ is well-defined, i.e., that the value $w^i(x)$ that we are assigning to $g^i \rightharpoonup x$ depends only on $g^i$ and on $x$ (but not on the specific choice of $i$). In other words, we must prove that if $i$ and $j$ are two integers satisfying $g^i = g^j$, then $w^i(x) = w^j(x)$ for each $x \in X$. But this is not hard: If two integers $i$ and $j$ satisfy $g^i = g^j$, then $i \equiv j \bmod n$ (since $g \in C_n$ has order $n$), and thus $i = j + nk$ for some $k \in \mathbb{Z}$; but this entails $w^i = w^{j+nk} = w^j \left( \underbrace{w^n}_{=\mathrm{id}} \right)^k = w^j \mathrm{id}^k = w^j$, so that $w^i(x) = w^j(x)$ for each $x \in X$.

Thus, we can define a map $\rightharpoonup : C_n \times X \to X$ by setting

$$g^i \rightharpoonup x = w^i(x) \qquad \text{for all } i \in \mathbb{Z} \text{ and } x \in X.$$

It remains to show that this map $\rightharpoonup$ is an action of $C_n$ on $X$. This requires verifying the two axioms in Definition 4.1.4 **(a)**. But this is straightforward (e.g., the first axiom boils down to $w^i\left(w^j(x)\right) = w^{i+j}(x)$ for all $i, j \in \mathbb{Z}$ and $x \in X$). This proves Theorem 4.1.52 **(b)**.

**(c)** We must prove the following claims:

> *Claim 1:* If we convert a left $C_n$-action $\rightharpoonup$ into a permutation $w \in S_X$, and then convert the latter permutation $w$ back into a left $C_n$-action $\rightharpoonup'$, then this new left $C_n$-action $\rightharpoonup'$ is the original $C_n$-action $\rightharpoonup$.
>
> *Claim 2:* If we convert a permutation $w \in S_X$ satisfying $w^n = \mathrm{id}$ into a left $C_n$-action $\rightharpoonup$, and then convert the latter $C_n$-action $\rightharpoonup$ back into a permutation $w' \in S_X$, then this $w'$ is the original $w$.

*Proof of Claim 1.* Let $\rightharpoonup$ be a left $C_n$-action on $X$. Convert this left $C_n$-action into a permutation $w \in S_X$, and then convert the latter permutation $w$ back into a left $C_n$-action $\rightharpoonup'$. We must prove that this $\rightharpoonup'$ is the original $C_n$-action $\rightharpoonup$.

Consider the curried form $\rho : C_n \to S_X$ of the original left $C_n$-action $\rightharpoonup$. Then, $\rho$ is a group morphism. The definition of $\rho$ yields $\rho(g) = \tau_g$ (where $\tau_g$ is defined as in Theorem 4.1.52 **(a)**).

By the definition of $w$, we have $w = \tau_g$. Comparing this with $\rho(g) = \tau_g$, we find $w = \rho(g)$. Thus, for each $i \in \mathbb{Z}$, we have

$$w^i = (\rho(g))^i = \rho\left(g^i\right) \tag{98}$$

(since $\rho$ is a group morphism).

By the definition of $\rightharpoonup'$, we have

$$g^i \rightharpoonup' x = w^i(x) \qquad \text{for all } i \in \mathbb{Z} \text{ and } x \in X.$$

Hence, for all $i \in \mathbb{Z}$ and $x \in X$, we have

$$\begin{aligned} g^i \mathrel{\rightharpoonup'} x &= w^i(x) \\ &= \left(\rho\left(g^i\right)\right)(x) \qquad \left(\text{since (98) yields } w^i = \rho\left(g^i\right)\right) \\ &= g^i \mathrel{\rightharpoonup} x \qquad \left(\begin{array}{c} \text{since } \rho\left(g^i\right) \text{ is defined to be the map } \tau_{g^i} : X \to X \\ \text{that sends each } y \in X \text{ to } g^i \mathrel{\rightharpoonup} y \end{array}\right). \end{aligned}$$

In other words, for all $h \in C_n$ and $x \in X$, we have $h \mathrel{\rightharpoonup'} x = h \mathrel{\rightharpoonup} x$ (since each $h \in C_n$ can be written as $g^i$ for some $i \in \mathbb{Z}$). In other words, the map $\mathrel{\rightharpoonup'}$ equals the map $\mathrel{\rightharpoonup}$. This proves Claim 1. □

*Proof of Claim 2.* Let $w \in S_X$ be a permutation satisfying $w^n = \mathrm{id}$. Convert this permutation $w$ into a left $C_n$-action $\mathrel{\rightharpoonup}$, and then convert the latter $C_n$-action $\mathrel{\rightharpoonup}$ back into a permutation $w' \in S_X$. We must prove that this $w'$ is the original $w$.

The definition of $w'$ yields $w' = \tau_g$ (where $\tau_g$ is defined as in Theorem 4.1.52 **(a)**). Thus, for each $x \in X$, we have

$$\begin{aligned} w'(x) = \tau_g(x) &= g \mathrel{\rightharpoonup} x \qquad \left(\text{by the definition of } \tau_g\right) \\ &= g^1 \mathrel{\rightharpoonup} x \qquad \left(\text{since } g = g^1\right) \\ &= w^1(x) \qquad \left(\text{by the definition of } \mathrel{\rightharpoonup}\right). \end{aligned}$$

Hence, $w' = w^1 = w$. This proves Claim 2. □

Thus, the proof of Theorem 4.1.52 **(c)** is complete. □

Something similar (but simpler) holds for the infinite cyclic group $C_\infty$:

**Theorem 4.1.53.** Let $C_\infty$ be the infinite cyclic group, with generator $g$. Let $X$ be a set. Then, a left $C_\infty$-action on $X$ is "the same as" a permutation $w \in S_X$. To be more precise:

**(a)** If $\mathrel{\rightharpoonup}$ is a left $C_\infty$-action on $X$, then the map

$$\begin{aligned} \tau_g : X &\to X, \\ x &\mapsto g \mathrel{\rightharpoonup} x \end{aligned}$$

is a permutation $w \in S_X$.

**(b)** Conversely, if $w \in S_X$ is a permutation, then we can define a left $C_\infty$-action on $X$ by setting

$$g^i \mathrel{\rightharpoonup} x = w^i(x) \qquad \text{for all } i \in \mathbb{Z} \text{ and } x \in X.$$

**(c)** These two conversions (from a left $C_\infty$-action to a permutation $w$, and vice versa) are mutually inverse.

*Proof.* Similar to Theorem 4.1.52. □

Thus, a left $C_\infty$-set is nothing but a set with a permutation.

**Remark 4.1.54.** There is a similar result about monoid actions: Let $M$ be the monoid $(\mathbb{N}, +, 0)$. Then, a left $M$-set is just a set with a function from this set to itself.

Alas, $S_n$-actions for $n \geq 3$ cannot be described as just "sets with a map on them", since the group $S_n$ cannot be expressed through a single generator.[35] This makes them all the more interesting.

### 4.1.10. Faithfulness

Here is another useful property of certain actions:

**Definition 4.1.55. (a)** A left $G$-set $X$ is said to be *faithful* if it has the following property: If $g, h \in G$ are two elements that satisfy

$$(g \rightharpoonup x = h \rightharpoonup x \text{ for all } x \in X),$$

then $g = h$.

In other words, a left $G$-set $X$ is *faithful* if and only if the curried form $\rho : G \to S_X$ of its $G$-action is injective.

**(b)** Instead of saying that a left $G$-set $X$ is faithful, we also say that the action of $G$ on $X$ is faithful.

So a left $G$-set is faithful if different elements of $G$ can be told apart based on how they act on it. Thus, a faithful $G$-set allows us to replace the elements $g$ of $G$ by the corresponding permutations $\tau_g : X \to X$ without accidentally equating two different elements.

The following are some examples of faithfulness and its lack:

[35] They can be described in a similar, but more complicated way. For instance, a known fact (the "Moore presentation" of $S_n$) says that the symmetric group $S_n$ can be described by the generators $s_1, s_2, \ldots, s_{n-1}$ and the relations

$$\begin{aligned} s_i^2 &= 1 \qquad \text{for all } i \in [n-1]; \\ s_i s_j &= s_j s_i \qquad \text{for all } i, j \in [n-1] \text{ satisfying } |i-j| > 1; \\ s_i s_{i+1} s_i &= s_{i+1} s_i s_{i+1} \qquad \text{for all } i \in [n-2] \end{aligned}$$

(see [Willia03, Theorem 1.2.4] for a proof). Using this fact, we can then describe a left $S_n$-set as being "just" a set $X$ equipped with $n-1$ permutations $w_1, w_2, \ldots, w_{n-1} \in S_X$ that satisfy

$$\begin{aligned} w_i^2 &= 1 \qquad \text{for all } i \in [n-1]; \\ w_i w_j &= w_j w_i \qquad \text{for all } i, j \in [n-1] \text{ satisfying } |i-j| > 1; \\ w_i w_{i+1} w_i &= w_{i+1} w_i w_{i+1} \qquad \text{for all } i \in [n-2]. \end{aligned}$$

(These permutations $w_i$ are just the actions of the corresponding $s_i \in S_n$ on $X$.) Whether this description is any easier (or more intuitive) than the original definition of a left $S_n$-set, however, is another question.

- The left regular $G$-action on $G$ (defined in Example 4.1.6) is faithful. This is the "injective" part of Cayley's Theorem (Theorem 4.1.3).
- The inverse right regular $G$-action on $G$ (defined in Example 4.1.7) is also faithful.
- The conjugation action of $G$ on $G$ (defined in Example 4.1.8), given by

$$g \rightharpoonup x = gxg^{-1} \qquad \text{for all } g \in G \text{ and } x \in G,$$

  is often not faithful. For example, if $g$ is any element of the center of $G$, then $g \rightharpoonup x = \underbrace{gx}_{=xg} g^{-1} = x \underbrace{gg^{-1}}_{=1} = x$ for each $x \in G$; thus, all elements of the center of $G$ act in the same way (as the identity map). Hence, the conjugation action of $G$ is not faithful unless the center of $G$ is trivial. The converse is also true (see Exercise 4.1.5 for a proof).
- The trivial action of $G$ on a set $X$ is only faithful if $G$ is trivial.

**Exercise 4.1.4.** [1] Let $X$ be a left $G$-set. Prove that $X$ is faithful if and only if the only $g \in G$ satisfying ($g \rightharpoonup x = x$ for all $x \in X$) is $1_G$.

**Exercise 4.1.5.** [1] Prove that the conjugation action of $G$ on $G$ is faithful if and only if $Z(G) = \{1\}$.

**Exercise 4.1.6.** [1] Let $A$ be any set. Let $k \in \mathbb{N}$. Consider the left $S_A$-set $\mathcal{P}_k(A)$ from Example 4.1.14. What conditions must $k$ satisfy in order for this $S_A$-set to be faithful?

**Exercise 4.1.7.** [1] Let $n$ and $m$ be positive integers, and assume that the ring $\mathbf{k}$ is not trivial. Prove that the action of $S_n \times S_m$ on $\mathbf{k}^{n \times m}$ is faithful.

### 4.1.11. Further reading

Group actions have lots of applications; in particular they come helpful in proving group-theoretical results such as the Sylow theorems. See [Conrad, "Group actions", "Transitive group actions", "The Sylow theorems"], [Goodma16, Chapter 5], [McClea23, Chapter 2], [Carter09, Chapter 9] and other texts for this and much more about group actions.

Group actions can also be used to solve enumerative problems, like counting *necklaces* (equivalence classes of $n$-tuples under cyclic rotations). For an introduction to this, see [Davis23] or [Loehr11, Chapter 9]. A book-length treatment can be found in [Kerber99].

## 4.2. Linear group actions and representations of $G$

Let us continue the study of $G$-sets, but now combining it with the well-known concept of a $\mathbf{k}$-module. (You can think of $\mathbf{k}$ as being your favorite field – such as $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{Z}/p = \mathbb{F}_p$ for some prime $p$. Many authors restrict themselves to the case of $\mathbf{k} = \mathbb{C}$ and find an interesting theory already at that level. When $\mathbf{k}$ is a field, the $\mathbf{k}$-modules are also known as $\mathbf{k}$-vector spaces, so you have linear algebra at your disposal.)

**Convention 4.2.1.** For the entirety of Section 4.2, we let $G$ be a group (not necessarily finite).

### 4.2.1. Linear actions, aka representations

**Definition 4.2.2.** Let $G$ be a group.

**(a)** Let $V$ be a $\mathbf{k}$-module. A left action $\rightharpoonup : G \times V \to V$ of $G$ on $V$ is said to be $\mathbf{k}$-*linear* (or, for short, *linear*) if for each $g \in G$, the map

$$\begin{aligned} \tau_g : V &\to V, \\ x &\mapsto g \rightharpoonup x \end{aligned}$$

is $\mathbf{k}$-linear. Explicitly, this condition is saying that

$$\begin{aligned} &g \rightharpoonup (x+y) = (g \rightharpoonup x) + (g \rightharpoonup y) && \text{for all } g \in G \text{ and } x, y \in V; \\ &g \rightharpoonup (\lambda x) = \lambda\,(g \rightharpoonup x) && \text{for all } g \in G \text{ and } \lambda \in \mathbf{k} \text{ and } x \in V; \\ &g \rightharpoonup 0_V = 0_V && \text{for all } g \in G. \end{aligned}$$

**(b)** A *left representation of $G$ over $\mathbf{k}$* (aka a *left $G$-representation over $\mathbf{k}$*) means a $\mathbf{k}$-module $V$ equipped with a $\mathbf{k}$-linear left $G$-action on $V$.

We usually omit the words "left" and "over $\mathbf{k}$", unless the context suggests something different. Thus, we just say "representation of $G$" instead of "left representation of $G$ over $\mathbf{k}$". Some authors also refer to representations of $G$ as "*$G$-modules*", but we will avoid this notation in favor of a better one that we will soon introduce.

Let us revisit the above examples of $G$-actions, to see which of them are $\mathbf{k}$-linear. Obviously, we only need to consider the ones where the $G$-set is a $\mathbf{k}$-module.

**Example 4.2.3.** If $X$ is a $\mathbf{k}$-module, then the trivial $G$-action on $X$ (as defined in Example 4.1.10) is linear (since the map $\mathrm{id}_X : X \to X$ is $\mathbf{k}$-linear).

**Example 4.2.4.** If $A$ is a nontrivial $\mathbf{k}$-module, then the natural action of $S_A$ on $A$ (as defined in Example 4.1.12) is not linear, since a permutation of $A$ is usually not $\mathbf{k}$-linear.

**Example 4.2.5.** Let $A$ be a $\mathbf{k}$-module, and $n \in \mathbb{N}$. Then, the place permutation action of $S_n$ on $A^n$ (as defined in Example 4.1.17) is $\mathbf{k}$-linear. To check this, we need to verify the three equalities in Definition 4.2.2 **(a)**; let us check the first of them: We must prove that

$$g \rightharpoonup (x+y) = (g \rightharpoonup x) + (g \rightharpoonup y) \qquad \text{for all } g \in S_n \text{ and } x, y \in A^n.$$

To prove this, we write $x$ and $y$ as $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$, and observe that

$$\begin{aligned}
g \rightharpoonup (x+y) &= g \rightharpoonup \underbrace{((x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n))}_{=(x_1+y_1,\ x_2+y_2,\ \ldots,\ x_n+y_n)} \\
&= g \rightharpoonup (x_1 + y_1,\ x_2 + y_2,\ \ldots,\ x_n + y_n) \\
&= \left( x_{g^{-1}(1)} + y_{g^{-1}(1)},\ x_{g^{-1}(2)} + y_{g^{-1}(2)},\ \ldots,\ x_{g^{-1}(n)} + y_{g^{-1}(n)} \right) \\
&\qquad \text{(by the definition of } \rightharpoonup\text{)} \\
&= \underbrace{\left( x_{g^{-1}(1)},\ x_{g^{-1}(2)},\ \ldots,\ x_{g^{-1}(n)} \right)}_{\substack{=g \rightharpoonup (x_1, x_2, \ldots, x_n) \\ \text{(by the definition of } \rightharpoonup\text{)}}} + \underbrace{\left( y_{g^{-1}(1)},\ y_{g^{-1}(2)},\ \ldots,\ y_{g^{-1}(n)} \right)}_{\substack{=g \rightharpoonup (y_1, y_2, \ldots, y_n) \\ \text{(by the definition of } \rightharpoonup\text{)}}} \\
&= \left( g \rightharpoonup \underbrace{(x_1, x_2, \ldots, x_n)}_{=x} \right) + \left( g \rightharpoonup \underbrace{(y_1, y_2, \ldots, y_n)}_{=y} \right) \\
&= (g \rightharpoonup x) + (g \rightharpoonup y).
\end{aligned}$$

The other two equalities are proved similarly.

Thus, $A^n$ is a representation of $S_n$ over $\mathbf{k}$.

**Example 4.2.6.** The action of $S_n \times S_m$ on $\mathbf{k}^{n \times m}$ defined in Example 4.1.18 is $\mathbf{k}$-linear. This is straightforward to check.

**Example 4.2.7.** Consider the trivial group $\{1\}$. For any $\mathbf{k}$-module $X$, there is a unique left $\{1\}$-action on $X$ (as we learned in Example 4.1.50), and this $\{1\}$-action is $\mathbf{k}$-linear (since the action $\tau_1$ of 1 is $\mathrm{id}_X$, which is a $\mathbf{k}$-linear map). Thus, a representation of $\{1\}$ is just a $\mathbf{k}$-module.

**Example 4.2.8.** As we saw in Example 4.1.51, an action of the size-2 cyclic group $C_2$ on a set $X$ is just an involution of $X$. Assuming that $X$ is a $\mathbf{k}$-module, this action is $\mathbf{k}$-linear if and only if the involution is $\mathbf{k}$-linear. For example, the involution

$$\begin{aligned}
&\mathbf{k} \to \mathbf{k}, \\
&x \mapsto -x
\end{aligned}$$

is $\mathbf{k}$-linear (and thus gives a $\mathbf{k}$-linear action of $C_2$), but the involution

$$\begin{aligned} \mathbf{k} &\to \mathbf{k}, \\ x &\mapsto 1 - x \end{aligned}$$

is not.

Thus, a representation of $C_2$ is just a $\mathbf{k}$-module equipped with a $\mathbf{k}$-linear involution.

**Example 4.2.9.** As we saw in Theorem 4.1.52, an action of the size-$n$ cyclic group $C_n$ on a set $X$ is just a permutation $w$ of $X$ satisfying $w^n = \mathrm{id}$. Assuming that $X$ is a $\mathbf{k}$-module, this action is $\mathbf{k}$-linear if and only if the map $w$ is $\mathbf{k}$-linear. For example, for $\mathbf{k} = \mathbb{R}$ and $X = \mathbb{R}^2$, we can take $w$ to be the left multiplication by the matrix

$$\begin{pmatrix} \cos\dfrac{2\pi}{n} & -\sin\dfrac{2\pi}{n} \\ \sin\dfrac{2\pi}{n} & \cos\dfrac{2\pi}{n} \end{pmatrix} \in \mathbb{R}^{2\times 2},$$

aka a rotation by $2\pi/n$ around the origin (viewing $X = \mathbb{R}^2$ as the Cartesian plane). If we take $w$ to be this rotation, then $\mathbb{R}^2$ thus becomes a representation of $C_n$ over $\mathbb{R}$, and the action of $C_n$ is given explicitly by

$$\begin{aligned} g^k \rightharpoonup v = (&v \text{ rotated by } 2\pi k/n \text{ around the origin}) \\ &\text{for each } k \in \mathbb{Z} \text{ and } v \in \mathbb{R}^2. \end{aligned}$$

**Example 4.2.10.** The natural action of $\mathrm{GL}_n(\mathbb{R})$ on $\mathbb{R}^{n\times 1}$ (that is, the action defined in Example 4.1.19) is $\mathbb{R}$-linear. More generally, the natural action of the group

$$\mathrm{GL}_n(\mathbf{k}) := \{\text{all invertible } n \times n\text{-matrices over } \mathbf{k}\}$$

on $\mathbf{k}^{n\times 1}$ is $\mathbf{k}$-linear. This is called the *natural representation* of $\mathrm{GL}_n(\mathbf{k})$.

We can generalize the group $\mathrm{GL}_n(\mathbf{k})$ as follows:

**Definition 4.2.11.** Let $V$ be any $\mathbf{k}$-module. Then, we define

$$\begin{aligned} \mathrm{GL}(V) &:= \{\text{all invertible } \mathbf{k}\text{-linear maps from } V \text{ to } V\} \\ &= \{\text{all automorphisms of the } \mathbf{k}\text{-module } V\} \\ &= \{\text{all } \mathbf{k}\text{-linear permutations of } V\}. \end{aligned}$$

This is a group under composition, and is called the *general linear group* of $V$. It is a subgroup of the symmetric group $S_V$.

In particular, if $V = \mathbf{k}^{n\times 1}$ (the $\mathbf{k}$-module of column vectors of size $n$), then $\mathrm{GL}(V) \cong \mathrm{GL}_n(\mathbf{k})$ by the canonical isomorphism that sends each linear map $f : V \to V$ to the $n \times n$-matrix that represents this map in the standard basis of $\mathbf{k}^{n\times 1}$. Now, Example 4.2.10 generalizes as follows:

**Example 4.2.12.** For any $\mathbf{k}$-module $V$, we have a left $\mathrm{GL}(V)$-action on $V$ given by

$$f \rightharpoonup v = f(v) \qquad \text{for all } f \in \mathrm{GL}(V) \text{ and } v \in V.$$

This action is $\mathbf{k}$-linear, and is called the *natural action* of $\mathrm{GL}(V)$. The $\mathbf{k}$-module $V$, equipped with this action, is called the *natural representation* of $\mathrm{GL}(V)$.

### 4.2.2. The curried form

As we have seen in Definition 4.1.21, we can curry an arbitrary action of $G$ on any set $X$ to obtain a group morphism $\rho : G \to S_X$. Similarly, we can curry a linear action of $G$ on a $\mathbf{k}$-module $X$ to obtain a group morphism $\rho : G \to \mathrm{GL}(X)$. In fact, this is the same morphism $\rho$ as in Definition 4.1.21, just with its target $S_X$ replaced by $\mathrm{GL}(X)$:

**Theorem 4.2.13.** Let $X$ be a left $G$-set that is also a $\mathbf{k}$-module. Then, the left action $\rightharpoonup$ of $G$ on $X$ is $\mathbf{k}$-linear if and only if its curried form $\rho : G \to S_X$ satisfies $\rho(G) \subseteq \mathrm{GL}(X)$.

*Proof sketch.* This is more or less a restatement of Definition 4.2.2 **(a)**. To wit, consider the map

$$\begin{aligned} \tau_g : X &\to X, \\ x &\mapsto g \rightharpoonup x \end{aligned}$$

for each $x \in X$. Then, the definition of the curried form $\rho$ shows that

$$\rho(g) = \tau_g \qquad \text{for each } g \in G. \tag{99}$$

But we have the following chain of equivalences:

$$\begin{aligned}
&(\text{the left action } \rightharpoonup \text{ of } G \text{ on } X \text{ is } \mathbf{k}\text{-linear}) \\
&\iff \left(\text{the map } \tau_g \text{ is } \mathbf{k}\text{-linear for each } g \in G\right) \\
&\qquad\qquad (\text{by Definition 4.2.2 } \textbf{(a)}) \\
&\iff (\text{the map } \rho(g) \text{ is } \mathbf{k}\text{-linear for each } g \in G) \\
&\qquad\qquad \left(\text{since (99) yields } \tau_g = \rho(g) \text{ for each } g \in G\right) \\
&\iff (\text{we have } \rho(g) \in \mathrm{GL}(X) \text{ for each } g \in G) \\
&\qquad\qquad \left( \begin{array}{c} \text{because for each } g \in G, \text{ we have } \rho(g) \in S_X \\ \text{(since } \rho \text{ is a map from } G \text{ to } S_X\text{), which shows that} \\ \text{the map } \rho(g) \text{ is invertible, and thus } \rho(g) \text{ is } \mathbf{k}\text{-linear} \\ \text{if and only if } \rho(g) \in \mathrm{GL}(X) \end{array} \right) \\
&\iff (\rho(G) \subseteq \mathrm{GL}(X)).
\end{aligned}$$

This proves Theorem 4.2.13. □

Theorem 4.2.13 yields the following **k**-linear analogue of Theorem 4.1.23:

**Theorem 4.2.14.** Let $G$ be a group. Let $X$ be a $\mathbf{k}$-module. Then, a $\mathbf{k}$-linear left $G$-action on $X$ is "the same as" a group morphism from $G$ to $\mathrm{GL}(X)$. To be more precise:

**(a)** If $\rightharpoonup$ is a $\mathbf{k}$-linear left $G$-action on $X$, then the curried form of $\rightharpoonup$ is a group morphism from $G$ to $\mathrm{GL}(X)$.

**(b)** If $\rho$ is a group morphism from $G$ to $\mathrm{GL}(X)$, then the uncurried form of $\rho$ is a $\mathbf{k}$-linear left $G$-action on $X$.

**(c)** Currying and uncurrying are mutually inverse.

*Proof sketch.* **(a)** Let $\rightharpoonup$ be a $\mathbf{k}$-linear left $G$-action on $X$. Let $\rho$ be the curried form of $\rightharpoonup$. Then, $\rho$ is a group morphism from $G$ to $S_X$ (by the definition of a curried form), but also satisfies $\rho(G) \subseteq \mathrm{GL}(X)$ (by Theorem 4.2.13). Hence, $\rho$ is a group morphism from $G$ to $\mathrm{GL}(X)$. This proves Theorem 4.2.14 **(a)**.

**(b)** Let $\rho$ be a group morphism from $G$ to $\mathrm{GL}(X)$. We must prove that the uncurried form of $\rho$ is a $\mathbf{k}$-linear left $G$-action on $X$.

Let $\rightharpoonup$ be this uncurried form. Then, $\rightharpoonup$ is a $G$-action on $X$, and the curried form of $\rightharpoonup$ is $\rho$ (by Theorem 4.1.23 **(c)**). Hence, Theorem 4.2.13 yields that the left action $\rightharpoonup$ of $G$ on $X$ is $\mathbf{k}$-linear (since $\rho(G) \subseteq \mathrm{GL}(X)$). In other words, the uncurried form of $\rho$ is a $\mathbf{k}$-linear left $G$-action on $X$ (since this uncurried form is exactly $\rightharpoonup$). This proves Theorem 4.2.14 **(b)**.

**(c)** This is just a particular case of Theorem 4.1.23 **(c)**. □

**Corollary 4.2.15.** A representation of $G$ is "the same as" a $\mathbf{k}$-module $X$ equipped with a group morphism $\rho : G \to \mathrm{GL}(X)$.

This description is used as the definition of a representation of $G$ in most textbooks (e.g., [Prasad15, Definition 1.1.1], [Serre77, §1.1], [McClea23, Definition 4.1], [CSScTo10, §1.1.1], [Elman22, §109], [Roseng15, §1.5]).

### 4.2.3. Outlook on representation theory

The above definitions are the starting point for the study of representations of groups – a wide field of mathematical research known as *representation theory of groups*. Even the most classical setting, in which the group $G$ is finite and the ring $\mathbf{k}$ is a field of characteristic 0, is full of surprises and nontrivial results; by now, this part of the theory is well-exposed in several textbooks ([EGHetc11], [Serre77], [Prasad15], [CurRei62], [AlFais25], etc.). The case of finite groups $G$ and positive-characteristic fields $\mathbf{k}$ is much wilder (in the sense that there are much fewer general results and much more variety of behavior), and is a topic of ongoing research known as "*modular representation theory*" (see [Serre77, Part III], [Lassue23], [Morel19, Chapter III], [CurRei62, Chapter XII], [Schnei13] or [Webb16] for introductions, and – e.g. – [Zimmer14, Chapter 2] for a deeper dive). Infinite groups $G$ have also been studied extensively, often based on their topological and geometrical structures (see [Sternb95, Chapters 4–5], [Proces07], [Morel19, Chapter V] and many other places[36]).

Our focus in this course is different: We are concerned almost entirely with the case when $G$ is the symmetric group $S_n$, but we try to avoid any unnecessary assumptions on $\mathbf{k}$; we will even allow $\mathbf{k}$ to be an arbitrary commutative ring much of the time. We will classify the irreducible[37] representations of

[36] For an infinite group $G$, there is usually not much that can be said about all $G$-representations without any restrictions; they are "too wild" (e.g., there are uncountably many 1-dimensional $\mathbb{Q}$-linear representations of the multiplicative group $\mathbb{Q}^\times = (\mathbb{Q} \setminus \{0\}, \cdot, 1)$ over $\mathbb{Q}$; and this is far from the worst that can happen). Thus, much of the actual science is devoted to specific classes of groups (topological groups, Lie groups, pro-finite groups, etc.) and specific classes of representations (polynomial, rational, holomorphic, etc.). Among all these settings, two are particularly related to ours:

1. The polynomial representations of the general linear group $\mathrm{GL}_n\,\mathbf{k}$ have a particularly nice theory that mirrors and is closely related to the representation theory of symmetric groups; see [Fulton97, Chapter 8] or [Proces07, Chapter 9] for fairly elementary treatments.
2. There is also a representation theory of the infinite symmetric group (which, in this context, means the group of all permutations $\sigma$ of $\{1, 2, 3, \ldots\}$ that fix all but finitely many elements); see [Meliot17, Chapter 11] or [BorOls16].

However, we will not veer into any of this in the present text.

[37] Roughly speaking, a representation is said to be "irreducible" if it has no smaller representations "inside it". We will make this precise soon.

$S_n$ over characteristic-0 fields $\mathbf{k}$, but we will also (try to) analyze some non-irreducible ones in some detail, in particular decomposing them explicitly into subrepresentations. Thus, compared to most of the literature, we restrict ourselves to a much smaller set of groups $G$ (viz., the symmetric groups), but we allow ourselves to require less of the field $\mathbf{k}$ and to ask some deeper questions that make no sense for arbitrary $G$.

This said, there already exist several texts that are (fully or in great part) devoted to the representation theory of the symmetric groups specifically: [JamKer81], [Sagan01], [Feray16], [Kerber99], [James78], [Wildon18], [Howe22], [BleSch05], [Zelevi81b], [Knutso73], [Meliot17], [Prasad15] and many others. We will see a lot of overlap with many of these texts, but we will often make different choices in what results we cover and how we prove them.

### 4.2.4. Representations of $S_n$: first examples

We have already seen a few examples of representations of $S_n$. Let us recall them again:

**Example 4.2.16.** For any $\mathbf{k}$-module $V$, the trivial $S_n$-action on $V$ is $\mathbf{k}$-linear, and thus gives a representation of $S_n$. This is called a *trivial representation* of $S_n$. Recall that its action is given by

$$g \rightharpoonup v = v \qquad \text{for all } g \in S_n \text{ and } v \in V.$$

When $V = \mathbf{k}$, we refer to this trivial representation as "**the** trivial representation" of $S_n$ (with a definite article), and we denote it by $\mathbf{k}_{\text{triv}}$.

**Example 4.2.17.** The symmetric group $S_n$ acts on the $\mathbf{k}$-module $\mathbf{k}^n$ by

$$g \rightharpoonup (a_1, a_2, \ldots, a_n) = \left(a_{g^{-1}(1)}, a_{g^{-1}(2)}, \ldots, a_{g^{-1}(n)}\right)$$
$$\text{for all } g \in S_n \text{ and } (a_1, a_2, \ldots, a_n) \in A^n.$$

This is the place permutation action on $\mathbf{k}^n$ we introduced in Example 4.1.17 (for $A = \mathbf{k}$). As we know from Example 4.2.5, this left $S_n$-action is $\mathbf{k}$-linear, and therefore $\mathbf{k}^n$ becomes a representation of $S_n$. This representation is called the *natural representation* of $S_n$.

The trivial representation has a slightly less trivial sibling, which is also defined for any $\mathbf{k}$-module $V$:

**Example 4.2.18.** For any $\mathbf{k}$-module $V$, there is a $\mathbf{k}$-linear left $S_n$-action on $V$ given by

$$g \rightharpoonup v = (-1)^g v \qquad \text{for all } g \in S_n \text{ and } v \in V.$$

This gives a representation of $S_n$, which is called a *sign representation* of $S_n$.

When $V = \mathbf{k}$, we refer to this sign representation as "**the** sign representation" of $S_n$ (with a definite article), and we denote it by $\mathbf{k}_{\mathrm{sign}}$.

*Proof of Example 4.2.18 (sketched).* Similarly to Example 4.1.15, we can show that the formula $g \rightharpoonup v = (-1)^g v$ really defines a left $S_n$-action on $V$. That this $S_n$-action is $\mathbf{k}$-linear is obvious (since the action of a $g \in S_n$ on $V$ is just scaling by the number $(-1)^g$). $\square$

Inspired by these examples, we can easily come up with another (combining the sign representation with the natural one):

**Example 4.2.19.** The *sign-twisted natural representation* of $S_n$ is the $\mathbf{k}$-module $\mathbf{k}^n$ equipped with the left $S_n$-action given by

$$g \rightharpoonup (a_1, a_2, \ldots, a_n) = (-1)^g \left(a_{g^{-1}(1)}, a_{g^{-1}(2)}, \ldots, a_{g^{-1}(n)}\right)$$
$$\text{for all } g \in S_n \text{ and } (a_1, a_2, \ldots, a_n) \in A^n.$$

Here are some further representations of $S_n$:

**Example 4.2.20. (a)** Consider the polynomial ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$ in $n$ indeterminates $x_1, x_2, \ldots, x_n$ over $\mathbf{k}$. We define a left $S_n$-action on this polynomial ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$ by setting

$$g \rightharpoonup f = f\left(x_{g(1)}, x_{g(2)}, \ldots, x_{g(n)}\right)$$
$$\text{for all } g \in S_n \text{ and } f(x_1, x_2, \ldots, x_n) \in \mathbf{k}[x_1, x_2, \ldots, x_n].$$

(Thus, a permutation $g \in S_n$ acts on a polynomial $f$ by permuting the indeterminates.) This is indeed a left $S_n$-action (see [21s, Proposition 7.1.4] for a proof), and is easily seen to be $\mathbf{k}$-linear, so that the polynomial ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$ becomes a representation of $S_n$.

**(b)** Let $d \in \mathbb{N}$. Define $\mathbf{k}[x_1, x_2, \ldots, x_n]_d$ to be the set of all homogeneous polynomials of degree $d$ in $\mathbf{k}[x_1, x_2, \ldots, x_n]$. That is,

$$\mathbf{k}[x_1, x_2, \ldots, x_n]_d = \mathrm{span}_{\mathbf{k}}\left\{x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} \mid a_1 + a_2 + \cdots + a_n = d\right\}.$$

Thus, $\mathbf{k}[x_1, x_2, \ldots, x_n]_d$ is a $\mathbf{k}$-submodule of $\mathbf{k}[x_1, x_2, \ldots, x_n]$. We define a left $S_n$-action on this $\mathbf{k}$-submodule $\mathbf{k}[x_1, x_2, \ldots, x_n]_d$ in the same way as we defined it on the whole $\mathbf{k}[x_1, x_2, \ldots, x_n]$ in part **(a)**. Thus, $\mathbf{k}[x_1, x_2, \ldots, x_n]_d$ becomes a representation of $S_n$ as well.

What other representations of $S_n$ are there? We will soon see.

### 4.2.5. Morphisms of representations

Let us now define morphisms of representations. As with any algebraic structure, this is done in the obvious way: The morphisms are just the maps that preserve all the structures. In detail:

**Definition 4.2.21.** Let $G$ be a group. Let $V$ and $W$ be two representations of $G$ over $\mathbf{k}$. Let $f : V \to W$ be any map. Then, we say that $f$ is a *morphism of representations* (aka a *$G$-representation morphism*) if $f$ is both $\mathbf{k}$-linear (i.e., a morphism of $\mathbf{k}$-modules) and $G$-equivariant (i.e., a morphism of $G$-sets).

More classical synonyms for "morphism of representations" are "*intertwining map*" or "*intertwiner*", although I am not sure how standard the meanings of these words are (and I believe that authors are avoiding them nowadays).

**Example 4.2.22.** Consider the natural representation $\mathbf{k}^n$ of the symmetric group $S_n$ (defined in Example 4.2.17). Consider furthermore the trivial representation $\mathbf{k}$ of the symmetric group $S_n$ (defined in Example 4.2.16). Then:

**(a)** The map

$$\begin{aligned} \mathbf{k} &\to \mathbf{k}^n, \\ a &\mapsto \left( \underbrace{a, a, \ldots, a}_{n \text{ times}} \right) \end{aligned}$$

is a morphism of representations (since it is clearly $\mathbf{k}$-linear, and since we know from Example 4.1.29 **(a)** that it is $S_n$-equivariant).

**(b)** The map

$$\begin{aligned} \mathbf{k}^n &\to \mathbf{k}, \\ (a_1, a_2, \ldots, a_n) &\mapsto a_1 + a_2 + \cdots + a_n \end{aligned}$$

is a morphism of representations (since it is clearly $\mathbf{k}$-linear, and since we know from Example 4.1.29 **(b)** that it is $S_n$-equivariant).

**(c)** The map

$$\begin{aligned} \mathbf{k}^n &\to \mathbf{k}, \\ (a_1, a_2, \ldots, a_n) &\mapsto a_1 a_2 \cdots a_n \end{aligned}$$

is $S_n$-equivariant, but not a morphism of representations, since it is not $\mathbf{k}$-linear (unless $n = 1$ or $\mathbf{k}$ is trivial).

With morphisms come isomorphisms, which again are defined in the usual way:

**Definition 4.2.23.** An *isomorphism of representations* is an invertible morphism of representations whose inverse is also a morphism of representations.

**Proposition 4.2.24.** The "is also a morphism of representations" part in Definition 4.2.23 is redundant (i.e., follows from the other requirements). In other words, if a morphism of representations is invertible, then its inverse is again a morphism of representations.

*Proof.* Straightforward and standard, just like Proposition 4.1.32. □

**Definition 4.2.25.** Let $X$ and $Y$ be two representations of a group $G$. Then, we say that $X$ and $Y$ are *isomorphic* (this is written $X \cong Y$) if there exists an isomorphism of representations from $X$ to $Y$.

### 4.2.6. Subrepresentations

Subrepresentations are defined in the same way as any kinds of sub-objects in algebra:

**Definition 4.2.26.** Let $V$ be a representation of a group $G$ over $\mathbf{k}$. A *subrepresentation* of $V$ means a $\mathbf{k}$-submodule of $V$ that is simultaneously a $G$-subset of $V$.

Any such subrepresentation automatically is a representation of $G$ itself (by restricting the addition, the scaling and the $G$-action from $V$ to it).

Recall the natural representation $\mathbf{k}^n$ of the symmetric group $S_n$ (defined in Example 4.2.17). Let us study its subrepresentations. For convenience's sake, let us define a notation: Let $(e_1, e_2, \ldots, e_n)$ be the standard basis of $\mathbf{k}^n$ (so that $e_i \in \mathbf{k}^n$ is the $n$-tuple $(0, 0, \ldots, 0, 1, 0, 0, \ldots, 0)$ whose $i$-th entry is 1 and whose all other entries are 0). The action of the symmetric group $S_n$ then satisfies

$$g \rightharpoonup e_i = e_{g(i)} \qquad \text{for all } i \in [n]\,. \tag{100}$$

Now, what subrepresentations does the natural representation $\mathbf{k}^n$ of $S_n$ have? Here are four:

- One subrepresentation of $\mathbf{k}^n$ is the *diagonal subrepresentation*

$$\begin{aligned} D\left(\mathbf{k}^n\right) &:= \left\{(a_1, a_2, \ldots, a_n) \in \mathbf{k}^n \mid a_1 = a_2 = \cdots = a_n\right\} \\ &= \left\{(a, a, \ldots, a) \mid a \in \mathbf{k}\right\} \end{aligned}$$

  (we understand $(a, a, \ldots, a)$ to be an $n$-tuple here). This is a free $\mathbf{k}$-module of rank[38] 1, with basis

$$((1, 1, \ldots, 1)) = (e_1 + e_2 + \cdots + e_n)\,.$$

[38] The *rank* of a free $\mathbf{k}$-module is the number of elements in its basis. When $\mathbf{k}$ is a field, this is the classical concept of the dimension of a $\mathbf{k}$-vector space. I don't know why the word "dimension" is not commonly used for arbitrary $\mathbf{k}$, but the custom is to call it "rank".

In geometric terms, $D(\mathbf{k}^n)$ is a line through the origin.[39] As an $S_n$-representation, $D(\mathbf{k}^n)$ is isomorphic to the trivial representation $\mathbf{k}_{\text{triv}}$ of $S_n$ (defined in Example 4.2.16), since the map

$$\begin{aligned} \mathbf{k}_{\text{triv}} &\to D(\mathbf{k}^n), \\ a &\mapsto (a, a, \ldots, a) \end{aligned}$$

is an isomorphism of representations (why?).

- Another subrepresentation of $\mathbf{k}^n$ is the *reflection subrepresentation* (aka *zero-sum subrepresentation*)

$$R(\mathbf{k}^n) := \{(a_1, a_2, \ldots, a_n) \in \mathbf{k}^n \mid a_1 + a_2 + \cdots + a_n = 0\}.$$

It consists of all $n$-tuples in $\mathbf{k}^n$ whose entries sum to 0. This is a free $\mathbf{k}$-module of rank $n-1$ (unless $n = 0$), with basis

$$(e_1 - e_n,\ e_2 - e_n,\ \ldots,\ e_{n-1} - e_n)$$

(prove this, if you have not seen this before!). Geometrically, $R(\mathbf{k}^n)$ is the hyperplane through the origin perpendicular to the line $D(\mathbf{k}^n)$ (with respect to the dot product).

- Another subrepresentation of $\mathbf{k}^n$ is the zero submodule $\{0\}$. (In fact, for any representation $V$ of any group $G$, its zero submodule $\{0_V\}$ is a subrepresentation of $V$.)

- Another subrepresentation of $\mathbf{k}^n$ is $\mathbf{k}^n$ itself. (In fact, any representation $V$ has itself as a subrepresentation.)

Can we find any others?

For $\mathbf{k} = \mathbb{Z}$, we can find a lot. For instance, for any integers $u, v, w$, the subset

$$\begin{aligned} R_{u,v,w}(\mathbb{Z}^n) := \Big\{ & (a_1, a_2, \ldots, a_n) \in \mathbb{Z}^n \mid a_1 \equiv a_2 \equiv \cdots \equiv a_n \equiv 0 \bmod u \\ & \text{and } a_1 \equiv a_2 \equiv \cdots \equiv a_n \bmod v \\ & \text{and } a_1 + a_2 + \cdots + a_n \equiv 0 \bmod w \Big\} \end{aligned}$$

of $\mathbb{Z}^n$ is a subrepresentation of the natural $S_n$-representation $\mathbb{Z}^n$. (Keep in mind that "$a \equiv b \bmod 0$" means "$a = b$", so that we can encode any equalities as congruences.) Are there even more? I don't know.

When $\mathbf{k}$ is a field, however, the answer is much simpler:

[39] These "geometric terms" are, of course, not to be taken too literally; after all, $\mathbf{k}$ is just a commutative ring here.

**Theorem 4.2.27.** Assume that $\mathbf{k}$ is a field. Then, the only subrepresentations of the natural $S_n$-representation $\mathbf{k}^n$ are $D(\mathbf{k}^n)$, $R(\mathbf{k}^n)$, $\{0\}$ and $\mathbf{k}^n$.

*Proof idea.* Let $W$ be a subrepresentation of $\mathbf{k}^n$. If $W \subseteq D(\mathbf{k}^n)$, then $W$ is either $D(\mathbf{k}^n)$ or $\{0\}$ (why?). If not, then $W$ contains some vector

$$w = (w_1, w_2, \ldots, w_n)$$

with two distinct consecutive entries (why?). Pick such a $w = (w_1, w_2, \ldots, w_n)$, and pick some $i \in [n-1]$ such that $w_i \neq w_{i+1}$. Then, since $W$ is a subrepresentation, we must have $s_i \rightharpoonup w \in W$, and thus $w - s_i \rightharpoonup w \in W$. But

$$\begin{aligned} w - s_i \rightharpoonup w &= (w_1, w_2, \ldots, w_n) - (w_1, w_2, \ldots, w_{i-1}, w_{i+1}, w_i, w_{i+2}, \ldots, w_n) \\ &= (0, 0, \ldots, 0, w_i - w_{i+1}, w_{i+1} - w_i, 0, 0, \ldots, 0) \\ &= (w_i - w_{i+1})(e_i - e_{i+1}) \end{aligned}$$

(where $e_1, e_2, \ldots, e_n$ are the standard basis vectors again). Thus,

$$(w_i - w_{i+1})(e_i - e_{i+1}) = w - s_i \rightharpoonup w \in W.$$

Hence, $e_i - e_{i+1} \in W$ (since $w_i - w_{i+1}$ is nonzero and therefore invertible in the field $\mathbf{k}$). From this, we can easily derive that $e_i - e_n \in W$ (why?). This, in turn, easily entails that $e_k - e_n \in W$ for each $k \in [n-1]$ (why?). From this, we obtain $R(\mathbf{k}^n) \subseteq W$ (why?). But this entails that $W$ is either $R(\mathbf{k}^n)$ or $\mathbf{k}^n$ (why?). Thus, Theorem 4.2.27 is proved. □

**Exercise 4.2.1.** [3] Fill in the "why?"s in the above proof.

As mentioned above, when $\mathbf{k} = \mathbb{R}$, we can visualize $\mathbf{k}^n = \mathbb{R}^n$ geometrically as $n$-dimensional space, and then the subrepresentations $D(\mathbf{k}^n)$ and $R(\mathbf{k}^n)$ become (respectively) a line and a hyperplane orthogonal to it. This geometric picture is a good guide for the case $\mathbf{k} = \mathbb{R}$, but it can mislead when $\mathbf{k}$ is (e.g.) a finite field. Indeed, you might expect the orthogonality to ensure that $R(\mathbf{k}^n)$ and $D(\mathbf{k}^n)$ are complementary subspaces (or $\mathbf{k}$-submodules, to be more precise); but this reasoning only works when $\mathbf{k}$ is an ordered field like $\mathbb{R}$ or $\mathbb{Q}$. Thus, the question of how $R(\mathbf{k}^n)$ and $D(\mathbf{k}^n)$ interact in general needs to be answered algebraically. Here is a partial answer:

**Proposition 4.2.28.** **(a)** If $n$ is invertible in $\mathbf{k}$, then

$$\mathbf{k}^n = R(\mathbf{k}^n) \oplus D(\mathbf{k}^n) \qquad (\text{as } \mathbf{k}\text{-modules}).$$

That is, $R(\mathbf{k}^n)$ and $D(\mathbf{k}^n)$ are complementary $\mathbf{k}$-submodules of $\mathbf{k}^n$ in this case.

**(b)** If $n = 0$ in $\mathbf{k}$, then $D(\mathbf{k}^n) \subseteq R(\mathbf{k}^n)$.

**Exercise 4.2.2.** [2] Prove this proposition.

Note that Proposition 4.2.28 does not cover all possibilities, since there can be commutative rings $\mathbf{k}$ in which $n$ is neither invertible nor 0 (for example, 2 is neither invertible nor zero in $\mathbb{Z}/4$ or $\mathbb{Z}/6$). But any field $\mathbf{k}$ falls under either part **(a)** or part **(b)** of Proposition 4.2.28, since any element of a field is either invertible or zero.

### 4.2.7. Quotients

We know what quotients of $\mathbf{k}$-modules are, and we know (from Definition 4.1.45 **(b)**) what quotients of $G$-sets are. Thus we can easily define quotients of representations:

**Definition 4.2.29.** Let $G$ be a group. Let $V$ be a representation of $G$ over $\mathbf{k}$. Let $W$ be a subrepresentation of $V$. Consider the quotient $\mathbf{k}$-module

$$\begin{aligned} V/W &= \{\text{cosets of } W \text{ in } V\} \\ &= \{\text{residue classes of vectors in } V \text{ modulo } W\}. \end{aligned}$$

("Residue class" is just a synonym for "coset" here.) Let us write $\overline{v}$ for the coset $v + W \in V/W$ whenever $v \in V$.

Now, the $\mathbf{k}$-module $V/W$ becomes a representation of $G$ in a natural way, by defining the left $G$-action on $V/W$ by

$$g \rightharpoonup \overline{v} = \overline{g \rightharpoonup v} \qquad \text{for all } g \in G \text{ and } v \in V.$$

Alternatively, this left $G$-action can be obtained as a particular case of Definition 4.1.45 **(b)**: Namely, let $\overset{W}{\equiv}$ be the binary relation "congruent modulo $W$" on $V$ (that is, the relation defined by

$$\left(v_1 \overset{W}{\equiv} v_2\right) \iff (v_1 - v_2 \in W)$$

for all $v_1, v_2 \in V$). This relation $\overset{W}{\equiv}$ is an equivalence relation and is $G$-invariant (since $v_1 \overset{W}{\equiv} v_2$ yields $v_1 - v_2 \in W$ and thus $g \rightharpoonup (v_1 - v_2) \in W$ for each $g \in G$, so that

$$\begin{aligned} g \rightharpoonup v_1 - g \rightharpoonup v_2 &= g \rightharpoonup (v_1 - v_2) \qquad \text{(by the } \mathbf{k}\text{-linearity of } \rightharpoonup) \\ &\in W \end{aligned}$$

and therefore $g \rightharpoonup v_1 \overset{W}{\equiv} g \rightharpoonup v_2$). Thus, the quotient set $V/\overset{W}{\equiv}$ itself acquires a left $G$-action, according to Definition 4.1.45 **(b)**. This is precisely the left $G$-action on $V/W$ defined above (since $V/W$ is precisely $V/\overset{W}{\equiv}$).

Now, let us study the quotients of the natural representation $\mathbf{k}^n$ of $S_n$ by its two subrepresentations $D(\mathbf{k}^n)$ and $R(\mathbf{k}^n)$. Here, we are using the notations introduced in Subsection 4.2.6.

**Theorem 4.2.30. (a)** We have $\mathbf{k}^n / D(\mathbf{k}^n) \cong R(\mathbf{k}^n)$ as $S_n$-representations if $n$ is invertible in $\mathbf{k}$.

**(b)** We always have $\mathbf{k}^n / D(\mathbf{k}^n) \cong R(\mathbf{k}^n)$ as $\mathbf{k}$-modules.

**(c)** For $n \geq 3$, we don't always have $\mathbf{k}^n / D(\mathbf{k}^n) \cong R(\mathbf{k}^n)$ as $S_n$-representations.

**(d)** We always have $\mathbf{k}^n / R(\mathbf{k}^n) \cong D(\mathbf{k}^n)$ as $S_n$-representations.

*Proof of Theorem 4.2.30.* WLOG assume that $n \neq 0$ (since the $n = 0$ case is obvious). Let $(e_1, e_2, \ldots, e_n)$ be the standard basis of $\mathbf{k}^n$.

Some intuition first: We have

$$D(\mathbf{k}^n) = \operatorname{span}_{\mathbf{k}} \{(1, 1, \ldots, 1)\} = \{(a, a, \ldots, a) \mid a \in \mathbf{k}\}$$

(where the tuples are $n$-tuples). The elements of $\mathbf{k}^n / D(\mathbf{k}^n)$ are residue classes of vectors in $\mathbf{k}^n$ modulo this submodule $D(\mathbf{k}^n)$. Thus, roughly speaking, they are vectors in $\mathbf{k}^n$ defined up to adding a vector of the form $(a, a, \ldots, a)$ with $a \in \mathbf{k}$. In other words, they are vectors in $\mathbf{k}^n$ defined up to adding the same scalar to each entry.

In the following, for any $v \in \mathbf{k}^n$, we shall let $\overline{v}$ denote the residue class of $v$ in $\mathbf{k}^n / D(\mathbf{k}^n)$ or in $\mathbf{k}^n / R(\mathbf{k}^n)$ depending on the context. (Specifically, in the proofs of part **(a)**, **(b)** and **(c)**, it will mean the residue class in $\mathbf{k}^n / D(\mathbf{k}^n)$, whereas in the proof of part **(d)** it will mean the residue class in $\mathbf{k}^n / R(\mathbf{k}^n)$.)

**(a)** Assume that $n$ is invertible in $\mathbf{k}$. Consider the map

$$\begin{aligned} f : R(\mathbf{k}^n) &\to \mathbf{k}^n / D(\mathbf{k}^n), \\ v &\mapsto \overline{v}. \end{aligned}$$

This map $f$ is $\mathbf{k}$-linear (obviously) and $S_n$-equivariant (since $g \rightharpoonup \overline{v} = \overline{g \rightharpoonup v}$ for all $g \in S_n$ and $v \in \mathbf{k}^n$), so it is a morphism of $S_n$-representations. Moreover:

- The map $f$ is injective.

  [*Proof:* Let $v$ lie in the kernel of the $\mathbf{k}$-linear map $f$. Thus, $v \in R(\mathbf{k}^n)$ and $\overline{v} = \overline{0}$. From $v \in D(\mathbf{k}^n)$, we obtain $v = (a, a, \ldots, a)$ for some $a \in \mathbf{k}$. Consider this $a$. Then, $(a, a, \ldots, a) = v \in R(\mathbf{k}^n)$, so that $\underbrace{a + a + \cdots + a}_{n \text{ times}} = 0$ (by the definition of $R(\mathbf{k}^n)$). Hence, $na = \underbrace{a + a + \cdots + a}_{n \text{ times}} = 0$. Since $n$ is invertible in $\mathbf{k}$, we can multiply this equality by $n^{-1}$ and obtain $a = 0$. Hence, we can rewrite $v = (a, a, \ldots, a)$ as $v = (0, 0, \ldots, 0) = 0$.

Forget that we fixed $v$. We thus have shown that each $v$ in the kernel of the $\mathbf{k}$-linear map $f$ is 0. In other words, the kernel of the $\mathbf{k}$-linear map $f$ is $\{0\}$. Hence, $f$ is injective.]

- The map $f$ is surjective.

  [*Proof:* Let $x \in \mathbf{k}^n / D(\mathbf{k}^n)$. We want to find a $y \in R(\mathbf{k}^n)$ such that $f(y) = x$.

  We have $x \in \mathbf{k}^n / D(\mathbf{k}^n)$, so that $x = \overline{v}$ for some $v \in \mathbf{k}^n$. Consider this $v$. This vector $v$ may or may not lie in $R(\mathbf{k}^n)$. However, we can find a vector $y \in R(\mathbf{k}^n)$ such that $\overline{y} = \overline{v}$. To this purpose, we write $v$ as $v = (v_1, v_2, \ldots, v_n)$, and we let $v_\varnothing := \dfrac{1}{n}(v_1 + v_2 + \cdots + v_n)$ denote the average entry of $v$ (this is well-defined, since $n$ is invertible in $\mathbf{k}$). Then, we set $y := v - (v_\varnothing, v_\varnothing, \ldots, v_\varnothing)$. Then, $v - y = (v_\varnothing, v_\varnothing, \ldots, v_\varnothing) \in D(\mathbf{k}^n)$, so that $\overline{v} = \overline{y}$ and therefore $\overline{y} = \overline{v} = x$. Moreover, from

  $$\begin{aligned} y &= \underbrace{v}_{=(v_1, v_2, \ldots, v_n)} - (v_\varnothing, v_\varnothing, \ldots, v_\varnothing) = (v_1, v_2, \ldots, v_n) - (v_\varnothing, v_\varnothing, \ldots, v_\varnothing) \\ &= (v_1 - v_\varnothing,\ v_2 - v_\varnothing,\ \ldots,\ v_n - v_\varnothing), \end{aligned}$$

  we see that the sum of all entries of $y$ is

  $$\begin{aligned} &(v_1 - v_\varnothing) + (v_2 - v_\varnothing) + \cdots + (v_n - v_\varnothing) \\ &= (v_1 + v_2 + \cdots + v_n) - n \underbrace{v_\varnothing}_{= \frac{1}{n}(v_1 + v_2 + \cdots + v_n)} \\ &= (v_1 + v_2 + \cdots + v_n) - n \cdot \frac{1}{n}(v_1 + v_2 + \cdots + v_n) = 0. \end{aligned}$$

  Thus, $y \in R(\mathbf{k}^n)$. Hence, $f(y)$ is well-defined and equals $f(y) = \overline{y} = x$.

  Forget that we fixed $x$. Thus, for each $x \in \mathbf{k}^n / D(\mathbf{k}^n)$, we have found a $y \in R(\mathbf{k}^n)$ such that $f(y) = x$. In other words, we have shown that the map $f$ is surjective.]

Now we know that the map $f$ is both injective and surjective, thus bijective, i.e., invertible. Hence, $f$ is an invertible morphism of $S_n$-representations, and therefore an isomorphism of $S_n$-representations (by Proposition 4.2.24). This proves Theorem 4.2.30 **(a)**.

**(b)** This is an exercise in linear algebra. We shall show that the two $\mathbf{k}$-modules $\mathbf{k}^n / D(\mathbf{k}^n)$ and $R(\mathbf{k}^n)$ are free of the same rank. This will automatically entail their isomorphism (as $\mathbf{k}$-modules, not as $S_n$-representations, of course), because any two free $\mathbf{k}$-modules of the same rank are isomorphic[40].

[40] *Proof.* Let $M$ and $N$ be two free $\mathbf{k}$-modules of the same rank $k$. Then, $M$ and $N$ have bases $(m_1, m_2, \ldots, m_k)$ and $(n_1, n_2, \ldots, n_k)$, respectively. Thus, we can define a $\mathbf{k}$-linear map $f : M \to N$ that sends each $m_i$ to the corresponding $n_i$. This $\mathbf{k}$-linear map $f$ is easily seen to be an isomorphism. Hence, $M$ and $N$ are isomorphic.

We already know that the $\mathbf{k}$-module $R\left(\mathbf{k}^n\right)$ is free of rank $n-1$, with basis $\left(e_1 - e_n,\ e_2 - e_n,\ \ldots,\ e_{n-1} - e_n\right)$.

Now, we claim that the $\mathbf{k}$-module $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$ is free of rank $n-1$, with basis $\left(\overline{e_1}, \overline{e_2}, \ldots, \overline{e_{n-1}}\right)$. Indeed:

- The residue classes $\overline{e_1}, \overline{e_2}, \ldots, \overline{e_{n-1}}$ in $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$ are $\mathbf{k}$-linearly independent.

  [*Proof:* Let $a_1, a_2, \ldots, a_{n-1} \in \mathbf{k}$ be scalars such that $a_1\overline{e_1} + a_2\overline{e_2} + \cdots + a_{n-1}\overline{e_{n-1}} = \overline{0}$. We must show that $a_1 = a_2 = \cdots = a_{n-1} = 0$.

  By assumption, we have $a_1\overline{e_1} + a_2\overline{e_2} + \cdots + a_{n-1}\overline{e_{n-1}} = \overline{0}$. In view of

  $$\begin{aligned} a_1\overline{e_1} + a_2\overline{e_2} + \cdots + a_{n-1}\overline{e_{n-1}} &= \overline{a_1e_1 + a_2e_2 + \cdots + a_{n-1}e_{n-1}} \\ &= \overline{(a_1, a_2, \ldots, a_{n-1}, 0)}, \end{aligned}$$

  we can rewrite this as $\overline{(a_1, a_2, \ldots, a_{n-1}, 0)} = \overline{0}$. In other words, $(a_1, a_2, \ldots, a_{n-1}, 0) \in D\left(\mathbf{k}^n\right)$. By the definition of $D\left(\mathbf{k}^n\right)$, this means that $a_1 = a_2 = \cdots = a_{n-1} = 0$. But this is precisely what we needed to prove. Thus, the $\mathbf{k}$-linear independence of $\overline{e_1}, \overline{e_2}, \ldots, \overline{e_{n-1}}$ in $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$ is proved.]

- The residue classes $\overline{e_1}, \overline{e_2}, \ldots, \overline{e_{n-1}}$ span the $\mathbf{k}$-module $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$.

  [*Proof:* We have

  $$\left(\overline{e_1} + \overline{e_2} + \cdots + \overline{e_{n-1}}\right) + \overline{e_n} = \overline{e_1} + \overline{e_2} + \cdots + \overline{e_n} = \overline{e_1 + e_2 + \cdots + e_n} = \overline{0}$$

  (since $e_1 + e_2 + \cdots + e_n = (1, 1, \ldots, 1) \in D\left(\mathbf{k}^n\right)$), so that

  $$\overline{e_n} = -\left(\overline{e_1} + \overline{e_2} + \cdots + \overline{e_{n-1}}\right). \tag{101}$$

  However, the standard basis vectors $e_1, e_2, \ldots, e_n$ clearly span the $\mathbf{k}$-module $\mathbf{k}^n$. Thus, their residue classes $\overline{e_1}, \overline{e_2}, \ldots, \overline{e_n}$ span the quotient $\mathbf{k}$-module $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$. Moreover, the last of the latter $n$ classes (that is, $\overline{e_n}$) is actually redundant (since the equality (101) expresses it as a $\mathbf{k}$-linear combination of the remaining $n-1$ classes $\overline{e_1}, \overline{e_2}, \ldots, \overline{e_{n-1}}$). Hence, the other residue classes $\overline{e_1}, \overline{e_2}, \ldots, \overline{e_{n-1}}$ already span the $\mathbf{k}$-module $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$.]

Thus, we know that the residue classes $\overline{e_1}, \overline{e_2}, \ldots, \overline{e_{n-1}}$ are $\mathbf{k}$-linearly independent and span the $\mathbf{k}$-module $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$. Hence, they form a basis of $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$. Thus, the $\mathbf{k}$-module $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$ is free of rank $n-1$.

Now, we know that the two $\mathbf{k}$-modules $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$ and $R\left(\mathbf{k}^n\right)$ are both free of rank $n-1$. Hence, they are free of the same rank, and thus isomorphic. This proves Theorem 4.2.30 **(b)**.

**(c)** Assume that $n \geq 3$, and that $\mathbf{k}$ is a field of characteristic $p$ with $p \mid n$. We shall show that the $S_n$-representations $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$ and $R\left(\mathbf{k}^n\right)$ are not isomorphic.

Note that it does **not** suffice that the map $f$ from the above proof of Theorem 4.2.30 **(a)** fails to be an isomorphism, or that the $\mathbf{k}$-module isomorphism from Theorem 4.2.30 **(b)** fails to be $S_n$-equivariant. After all, perhaps these two maps are just the wrong ones for the job, but some other map would qualify?

We must instead show that **no** $S_n$-representation isomorphism from $\mathbf{k}^n/D(\mathbf{k}^n)$ to $R(\mathbf{k}^n)$ exists.

Here is a general trick for proving non-isomorphy:

If $G$ is any group and $V$ is any representation of $G$, then we set

$$V^{G-\mathrm{fix}} := \{v \in V \mid g \rightharpoonup v = v \text{ for all } g \in G\}.$$

This is called the *space of $G$-fixed points in $V$*. It is a $\mathbf{k}$-submodule of $V$ and even a $G$-subrepresentation of $V$ (a trivial one, by its definition). Its elements are called the *$G$-fixed points in $V$*. (Classically, it is denoted by $V^G$, but this notation can also mean the set of all maps from $G$ to $V$, which is different.)

Obviously, if $V$ and $W$ are two isomorphic representations of $G$, then $V^{G-\mathrm{fix}} \cong W^{G-\mathrm{fix}}$ as well (as $\mathbf{k}$-modules or as $G$-representations, as you prefer). Thus, if two $G$-representations have non-isomorphic spaces of $G$-fixed points, then they themselves are not isomorphic.

We shall now show that the $S_n$-representations $\mathbf{k}^n/D(\mathbf{k}^n)$ and $R(\mathbf{k}^n)$ have non-isomorphic spaces of $S_n$-fixed points:

- The only $S_n$-fixed point in $\mathbf{k}^n/D(\mathbf{k}^n)$ is the zero vector $\overline{0}$.

  [*Proof:* Clearly, the zero vector $\overline{0}$ is an $S_n$-fixed point. Now let us prove the converse. Let $\overline{(a_1, a_2, \ldots, a_n)} \in \mathbf{k}^n/D(\mathbf{k}^n)$ be an $S_n$-fixed point. Then, by the definition of an $S_n$-fixed point, we have

  $$s_1 \rightharpoonup \overline{(a_1, a_2, \ldots, a_n)} = \overline{(a_1, a_2, \ldots, a_n)}.$$

  In view of

  $$s_1 \rightharpoonup \overline{(a_1, a_2, \ldots, a_n)} = \overline{s_1 \rightharpoonup (a_1, a_2, \ldots, a_n)} = \overline{(a_2, a_1, a_3, a_4, \ldots, a_n)},$$

  this rewrites as

  $$\overline{(a_2, a_1, a_3, a_4, \ldots, a_n)} = \overline{(a_1, a_2, \ldots, a_n)}.$$

  In other words,

  $$(a_2, a_1, a_3, a_4, \ldots, a_n) - (a_1, a_2, \ldots, a_n) \in D(\mathbf{k}^n).$$

  In view of

  $$\begin{aligned}&(a_2, a_1, a_3, a_4, \ldots, a_n) - (a_1, a_2, \ldots, a_n)\\ &= (a_2 - a_1,\ a_1 - a_2,\ 0,\ 0,\ \ldots,\ 0),\end{aligned}$$

this rewrites further as

$$(a_2 - a_1,\ a_1 - a_2,\ 0,\ 0,\ \ldots,\ 0) \in D\left(\mathbf{k}^n\right).$$

By the definition of $D\left(\mathbf{k}^n\right)$, we can rewrite this as

$$a_2 - a_1 = a_1 - a_2 = 0 = 0 = \cdots = 0.$$

Since $n \geq 3$, there is at least one 0 in this equality, so we obtain $a_1 - a_2 = 0$ and therefore $a_1 = a_2$. A similar argument (using $s_2$ instead of $s_1$) yields $a_2 = a_3$, and likewise we find $a_3 = a_4$ and $a_4 = a_5$ and so on. Combining all these equalities, we obtain

$$a_1 = a_2 = \cdots = a_n,$$

so that $(a_1, a_2, \ldots, a_n) \in D\left(\mathbf{k}^n\right)$ and thus $\overline{(a_1, a_2, \ldots, a_n)} = \overline{0}$.

Forget that we fixed $\overline{(a_1, a_2, \ldots, a_n)}$. We thus have shown that any $S_n$-fixed point $\overline{(a_1, a_2, \ldots, a_n)} \in \mathbf{k}^n / D\left(\mathbf{k}^n\right)$ equals $\overline{0}$. In other words, the only $S_n$-fixed point in $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$ is the zero vector $\overline{0}$.]

- I claim that the vector $(1, 1, \ldots, 1) \in \mathbf{k}^n$ is a nonzero $S_n$-fixed point in $R\left(\mathbf{k}^n\right)$.

  [*Proof:* The sum of all entries of the vector $(1, 1, \ldots, 1) \in \mathbf{k}^n$ is

  $$\underbrace{1 + 1 + \cdots + 1}_{n \text{ times}} = n \cdot 1 = 0 \qquad \text{in } \mathbf{k}$$

  (since $\operatorname{char} \mathbf{k} = p \mid n$). Thus, this vector $(1, 1, \ldots, 1)$ belongs to $R\left(\mathbf{k}^n\right)$ (by the definition of $R\left(\mathbf{k}^n\right)$). Moreover, this vector $(1, 1, \ldots, 1)$ is nonzero (since $\mathbf{k}$ is a field, so that $1 \neq 0$ in $\mathbf{k}$), and is an $S_n$-fixed point (since the action of a permutation $g \in S_n$ just permutes its entries, but all its entries are equal and therefore do not change under permutation). Hence, $(1, 1, \ldots, 1) \in \mathbf{k}^n$ is a nonzero $S_n$-fixed point in $R\left(\mathbf{k}^n\right)$.]

Thus, the $S_n$-representation $R\left(\mathbf{k}^n\right)$ has a nonzero $S_n$-fixed point (namely, $(1, 1, \ldots, 1)$), whereas the $S_n$-representation $\mathbf{k}^n / D\left(\mathbf{k}^n\right)$ does not. Therefore, these two representations cannot be isomorphic. This proves Theorem 4.2.30 **(c)**.

**(d)** Consider the $n$ residue classes $\overline{e_1}, \overline{e_2}, \ldots, \overline{e_n}$ in the quotient $\mathbf{k}$-module $\mathbf{k}^n / R\left(\mathbf{k}^n\right)$. We claim that all these classes are equal, i.e., that we have

$$\overline{e_1} = \overline{e_2} = \cdots = \overline{e_n}. \tag{102}$$

[*Proof:* For each $i \in \{2, 3, \ldots, n\}$, we have

$$e_i - e_1 = (-1, 0, 0, \ldots, 0, 1, 0, 0, \ldots, 0) \in R\left(\mathbf{k}^n\right)$$

(since the sum of the entries of the vector $(-1,0,0,\ldots,0,1,0,0,\ldots,0)$ is $(-1)+1=0$) and thus $\overline{e_i}=\overline{e_1}$. In other words, $\overline{e_1}=\overline{e_2}=\cdots=\overline{e_n}$. This proves (102).]

We shall now show that the $\mathbf{k}$-module $\mathbf{k}^n/R(\mathbf{k}^n)$ is free of rank 1, with basis $(\overline{e_1})$. Indeed:

- The residue class $\overline{e_1}$ in $\mathbf{k}^n/R(\mathbf{k}^n)$ is $\mathbf{k}$-linearly independent.

  [*Proof:* Let $a_1\in\mathbf{k}$ be a scalar such that $a_1\overline{e_1}=\overline{0}$. We must show that $a_1=0$.

  By assumption, we have $a_1\overline{e_1}=\overline{0}$. In view of

  $$a_1\overline{e_1}=\overline{a_1e_1}=\overline{(a_1,0,0,\ldots,0)},$$

  we can rewrite this as $\overline{(a_1,0,0,\ldots,0)}=\overline{0}$. In other words, $(a_1,0,0,\ldots,0)\in R(\mathbf{k}^n)$. By the definition of $R(\mathbf{k}^n)$, this means that $a_1+0+0+\cdots+0=0$. This obviously simplifies to $a_1=0$. But this is precisely what we needed to prove. Thus, the $\mathbf{k}$-linear independence of $\overline{e_1}$ in $\mathbf{k}^n/R(\mathbf{k}^n)$ is proved.]

- The residue class $\overline{e_1}$ spans the $\mathbf{k}$-module $\mathbf{k}^n/R(\mathbf{k}^n)$.

  [*Proof:* The standard basis vectors $e_1,e_2,\ldots,e_n$ clearly span the $\mathbf{k}$-module $\mathbf{k}^n$. Thus, their residue classes $\overline{e_1},\overline{e_2},\ldots,\overline{e_n}$ span the quotient $\mathbf{k}$-module $\mathbf{k}^n/R(\mathbf{k}^n)$. But all these $n$ residue classes are equal (by (102)), so that we need only one of them to span $\mathbf{k}^n/R(\mathbf{k}^n)$. Thus, the single residue class $\overline{e_1}$ already spans the $\mathbf{k}$-module $\mathbf{k}^n/R(\mathbf{k}^n)$.]

Thus, we have shown that the residue class $\overline{e_1}$ is $\mathbf{k}$-linearly independent and spans the $\mathbf{k}$-module $\mathbf{k}^n/R(\mathbf{k}^n)$. Hence, it forms a basis of $\mathbf{k}^n/R(\mathbf{k}^n)$. Thus, the $\mathbf{k}$-module $\mathbf{k}^n/R(\mathbf{k}^n)$ is free of rank 1.

Now, we know that the two $\mathbf{k}$-modules $\mathbf{k}^n/R(\mathbf{k}^n)$ and $D(\mathbf{k}^n)$ are both free of rank 1. Hence, they are free of the same rank, and thus isomorphic as $\mathbf{k}$-modules (for the same reason as in the proof of Theorem 4.2.30 **(b)**).

Are they also isomorphic as $S_n$-representations? To find out, we analyze their $S_n$-actions:

- The $S_n$-action on $\mathbf{k}^n/R(\mathbf{k}^n)$ is trivial (i.e., we have $g\rightharpoonup v=v$ for all $g\in S_n$ and $v\in\mathbf{k}^n/R(\mathbf{k}^n)$).

  [*Proof:* Let $g\in S_n$ and $v\in\mathbf{k}^n/R(\mathbf{k}^n)$. We must prove that $g\rightharpoonup v=v$.

  Since both sides of this equality are $\mathbf{k}$-linear in $v$, we can WLOG assume that $v=\overline{e_1}$ (since $\overline{e_1}$ spans the $\mathbf{k}$-module $\mathbf{k}^n/R(\mathbf{k}^n)$). Assume this. Then, $g\rightharpoonup v=g\rightharpoonup\overline{e_1}=\overline{g\rightharpoonup e_1}=\overline{e_{g(1)}}$ (since (100) yields $g\rightharpoonup e_1=e_{g(1)}$). However, (102) yields $\overline{e_{g(1)}}=\overline{e_1}$, so that $g\rightharpoonup v=\overline{e_{g(1)}}=\overline{e_1}=v$. This completes the proof that the $S_n$-action on $\mathbf{k}^n/R(\mathbf{k}^n)$ is trivial.]

- The $S_n$-action on $D(\mathbf{k}^n)$ is trivial (i.e., we have $g \rightharpoonup v = v$ for all $g \in S_n$ and $v \in D(\mathbf{k}^n)$).

  [*Proof:* Let $g \in S_n$ and $v \in D(\mathbf{k}^n)$. We must prove that $g \rightharpoonup v = v$.

  We have $v \in D(\mathbf{k}^n)$, so that $v = (a, a, \ldots, a)$ for some $a \in \mathbf{k}$. Thus, all entries of $v$ are equal. Hence, permuting the entries of $v$ does not change $v$. Therefore, $g \rightharpoonup v = v$ (since $g$ acts on $\mathbf{k}^n$ by permuting the entries). This completes the proof that the $S_n$-action on $D(\mathbf{k}^n)$ is trivial.]

Now, recall that the two $\mathbf{k}$-modules $\mathbf{k}^n / R(\mathbf{k}^n)$ and $D(\mathbf{k}^n)$ are isomorphic as $\mathbf{k}$-modules. In other words, there exists a $\mathbf{k}$-module isomorphism $f : \mathbf{k}^n / R(\mathbf{k}^n) \to D(\mathbf{k}^n)$. Since the $S_n$-actions on both $\mathbf{k}^n / R(\mathbf{k}^n)$ and $D(\mathbf{k}^n)$ are trivial, this $f$ is automatically $S_n$-equivariant (by Example 4.1.30, applied to $G = S_n$, $X = \mathbf{k}^n / R(\mathbf{k}^n)$ and $Y = D(\mathbf{k}^n)$), hence is a morphism of $S_n$-representations, and thus is an isomorphism of $S_n$-representations (by Proposition 4.2.24). This proves Theorem 4.2.30 **(d)**. □

Theorem 4.2.30 describes $\mathbf{k}^n / R(\mathbf{k}^n)$ always and $\mathbf{k}^n / D(\mathbf{k}^n)$ in most cases. (It is somewhat surprising that $\mathbf{k}^n / R(\mathbf{k}^n) \cong D(\mathbf{k}^n)$ is always true but $\mathbf{k}^n / D(\mathbf{k}^n) \cong R(\mathbf{k}^n)$ is not!)

**Exercise 4.2.3.** Consider the left action of the symmetric group $S_n$ on the matrix ring $\mathbf{k}^{n \times n}$ defined by

$$g \rightharpoonup \left(a_{i,j}\right)_{i \in [n],\, j \in [n]} = \left(a_{g^{-1}(i), g^{-1}(j)}\right)_{i \in [n],\, j \in [n]}$$
$$\text{for all } g \in S_n \text{ and } \left(a_{i,j}\right)_{i \in [n],\, j \in [n]}.$$

(This is the restriction of the action of $S_n \times S_n$ on $\mathbf{k}^{n \times n}$ defined in Example 4.1.18 along the morphism $\alpha : S_n \to S_n \times S_n,\ g \mapsto (g, g)$. Unlike the latter action, which permutes the rows and the columns by two different permutations, this one permutes the rows and the columns by the **same** permutation.) This left action is $\mathbf{k}$-linear, thus making $\mathbf{k}^{n \times n}$ into a representation of $S_n$.

**(a)** $\boxed{1}$ Prove that the set of all symmetric matrices in $\mathbf{k}^{n \times n}$ is a subrepresentation of $\mathbf{k}^{n \times n}$.

**(b)** $\boxed{1}$ Prove that the set of all matrices in $\mathbf{k}^{n \times n}$ whose columns sum to the zero vector is a subrepresentation of $\mathbf{k}^{n \times n}$.

**(c)** $\boxed{2}$ Find four more subrepresentations of $\mathbf{k}^{n \times n}$, not counting $\{0\}$ and $\mathbf{k}^{n \times n}$ itself. (The representations should be different for $\mathbf{k} = \mathbb{Q}$ and sufficiently large $n$. Proofs are not required in this part.)

**Exercise 4.2.4.** Let $V$ be the representation $\mathbf{k}[x_1, x_2, \ldots, x_n]_2$ of $S_n$ constructed in Example 4.2.20 **(b)** (applied to $d = 2$). This representation $V$ has basis

$\left(x_i x_j\right)_{1 \leq i \leq j \leq n}$, with $S_n$-action given by

$$g \rightharpoonup \left(x_i x_j\right) = x_{g(i)} x_{g(j)} \qquad \text{for all } i, j \in [n] .$$

Consider the two $\mathbf{k}$-submodules

$$W_1 := \operatorname{span}_{\mathbf{k}} \left\{x_1^2, x_2^2, \ldots, x_n^2\right\} \qquad \text{and}$$
$$W_2 := \operatorname{span}_{\mathbf{k}} \left\{x_i x_j \mid 1 \leq i < j \leq n\right\}$$

of $V$.

**(a)** $\boxed{1}$ Show that $W_1$ and $W_2$ are subrepresentations of $V$ such that $V = W_1 \oplus W_2$ (as a $\mathbf{k}$-module).

**(b)** $\boxed{1}$ One of $W_1$ and $W_2$ is isomorphic to an $S_n$-representation we have thoroughly studied. Which one?

**(c)** $\boxed{1}$ Prove that $W_1 \cong W_2$ (as $S_n$-representations) when $n = 3$.

**(d)** $\boxed{1}$ Show that each of $W_1$ and $W_2$ contains a nonzero trivial subrepresentation if $n \geq 2$.

## 4.3. Modules over rings, and a different way of thinking about representations

### 4.3.1. Definitions

In the preceding section (Corollary 4.2.15 in particular), we have seen two equivalent ways to view a representation of a group $G$ over $\mathbf{k}$: as a $\mathbf{k}$-linear left $G$-action on a $\mathbf{k}$-module $V$, or as a group morphism $\rho : G \to \mathrm{GL}(V)$ (its curried form). Soon we will see another: as a left $\mathbf{k}[G]$-module $V$. This is where the group algebra $\mathbf{k}[G]$ finally shows its use.

First, let us recall the notion of a left module (copied from [23wa, §3.1]):

**Definition 4.3.1.** Let $R$ be a ring. A *left $R$-module* (or a *left module over $R$*) means a set $M$ equipped with

- a binary operation $+$ (that is, a map from $M \times M$ to $M$) that is called *addition*;
- an element $0_M \in M$ that is called the *zero element* or the *zero vector* or just the *zero*, and is just denoted by $0$ when no confusion is likely;
- a map from $R \times M$ to $M$ that is called the *action of $R$ on $M$* (or *scaling of $M$ by $R$*), and is written as multiplication (i.e., we denote the image of a pair $(r, m) \in R \times M$ under this map by $rm$ or $r \cdot m$)

such that the following properties (the "*module axioms*") hold:

- $(M, +, 0)$ is an abelian group.
- The *right distributivity law* holds: We have $(r+s)\, m = rm + sm$ for all $r, s \in R$ and $m \in M$.
- The *left distributivity law* holds: We have $r\,(m+n) = rm + rn$ for all $r \in R$ and $m, n \in M$.
- The *associativity law* holds: We have $(rs)\, m = r\,(sm)$ for all $r, s \in R$ and $m \in M$.
- We have $0_R m = 0_M$ for every $m \in M$.
- We have $r \cdot 0_M = 0_M$ for every $r \in R$.
- We have $1m = m$ for every $m \in M$. (Here, "1" means the unity of $R$.)

When $M$ is a left $R$-module, the elements of $M$ are called *vectors*, and the elements of $R$ are called *scalars*. (Note that this sometimes flies in the face of standard intuition for scalars and vectors; see Example 4.3.7 below.)

As the name "left $R$-module" suggests, there is an analogous notion of a *right $R$-module*:

**Definition 4.3.2.** Let $R$ be a ring. A *right $R$-module* is defined just as a left $R$-module was defined in Definition 4.3.1, but with the following changes:

- For a right $R$-module $M$, the action is not a map from $R \times M$ to $M$, but rather a map from $M \times R$ to $M$.
- Accordingly, we use the notation $mr$ (rather than $rm$) for the image of a pair $(m, r)$ under this map.
- The axioms for a right $R$-module are similar to the module axioms for a left $R$-module, accounting for the different form of the action. For example, the associativity law for a right $R$-module is saying that $m\,(rs) = (mr)\, s$ for all $r, s \in R$ and $m \in M$.

When $R$ is commutative, there is a natural way to turn any left $R$-module into a right $R$-module and vice versa:

**Proposition 4.3.3.** Let $R$ be a commutative ring. Then, we can make any left $R$-module $M$ into a right $R$-module by setting

$$mr = rm \qquad \text{for all } r \in R \text{ and } m \in M. \tag{103}$$

Similarly, we can make any right $R$-module $M$ into a left $R$-module by setting

$$rm = mr \qquad \text{for all } r \in R \text{ and } m \in M. \tag{104}$$

These two transformations are mutually inverse, so we shall use them to identify left $R$-modules with right $R$-modules. Thus, we use the words "left $R$-module" and "right $R$-module" interchangeably, and just speak of "*$R$-modules*" instead (without specifying whether they are left or right).

Of course, this is precisely the notion of $\mathbf{k}$-modules that we have always been using, since $\mathbf{k}$ is assumed to be commutative.

**Remark 4.3.4.** Even when the ring $R$ is noncommutative, we can translate left modules into right modules, but not over the same ring: A left $R$-module naturally becomes a right $R^{\mathrm{op}}$-module, and vice versa, where $R^{\mathrm{op}}$ is the *opposite ring* of $R$ (that is, the ring $R$ with its multiplication reversed – i.e., the ring $R$ in which all products $rs$ have been renamed as $sr$).

When $R$ is a field, the $R$-modules are also known as *$R$-vector spaces*.

Most of the rings that we shall consider in this course are not just rings, but $\mathbf{k}$-algebras. If $R$ is a $\mathbf{k}$-algebra, then any $R$-module becomes a $\mathbf{k}$-module in a natural way:

**Definition 4.3.5.** Let $R$ be a $\mathbf{k}$-algebra. Then, any left $R$-module $M$ automatically becomes a $\mathbf{k}$-module by setting

$$\lambda v := \underbrace{\left(\lambda \cdot 1_R\right)}_{\in R} v \qquad \text{for all } \lambda \in \mathbf{k} \text{ and } v \in M.$$

This way of transforming $R$-modules into $\mathbf{k}$-modules is called *restriction of scalars*. The original action of $R$ on $M$ is easily seen to be $\mathbf{k}$-bilinear as a map from $R \times M$ to $M$.

Of course, the same holds for right $R$-modules. In the following, we shall rarely talk about right $R$-modules, since all of their theory is a carbon copy of the theory of left $R$-modules. Nevertheless, right $R$-modules sometimes appear in nature, and you should be ready to formulate and prove their properties as the need for them arises.

### 4.3.2. Examples

Let us see some examples of $R$-modules for various rings $R$. We do not linger on the case of commutative $R$, since we have already been discussing $\mathbf{k}$-modules for a long time. We recall that $\mathbf{k}$-modules are the same as left $\mathbf{k}$-modules (and also the same as right $\mathbf{k}$-modules, if we write the scalars on the right). Let us now see some examples of $R$-modules where $R$ is not necessarily commutative:

**Example 4.3.6.** Any ring $R$ is automatically a left $R$-module, if we define the action of $R$ on $R$ to be just the multiplication map of $R$ (that is, $rm$ will just be the product of $r$ and $m$ in $R$). This is called the *left regular $R$-module*, and its action is called the *left regular $R$-action*.

**Example 4.3.7.** The space $\mathbf{k}^{n\times 1}$ of column vectors is not only a $\mathbf{k}$-module, but also a left $\mathbf{k}^{n\times n}$-module. Here, the action of $\mathbf{k}^{n\times n}$ on $\mathbf{k}^{n\times 1}$ is matrix-vector multiplication.

(Thus, Definition 4.3.1 tells us here that we can refer to the elements of $\mathbf{k}^{n\times n}$ as "scalars" and to the elements of $\mathbf{k}^{n\times 1}$ as "vectors". This is rather baroque, since the "scalars" have $n^2$ degrees of freedom while the "vectors" have only $n$. In other words, our "scalars" are "large" while our "vectors" are "small", quite contrary to the intuition we might have from linear algebra. But such is the price of working with arbitrary rings instead of fields. For another example, $\mathbb{Z}/2$ is a $\mathbb{Z}$-module, so that the elements of $\mathbb{Z}$ can be "scalars" while those of $\mathbb{Z}/2$ are "vectors".)

**Example 4.3.8.** More generally, for any $n, m \in \mathbb{N}$, the matrix space $\mathbf{k}^{n\times m}$ is a left $\mathbf{k}^{n\times n}$-module. The action is again given by matrix multiplication:

$$(n \times n\text{-matrix}) \cdot (n \times m\text{-matrix}) = (n \times m\text{-matrix}).$$

The next example is particularly important for us, as it defines one of the simplest types of $\mathbf{k}[G]$-modules for a group algebra $\mathbf{k}[G]$:

**Example 4.3.9.** Let $G$ be a group, and let $X$ be a left $G$-set. Then, the free $\mathbf{k}$-module $\mathbf{k}^{(X)}$ becomes a left $\mathbf{k}[G]$-module by defining the action of $\mathbf{k}[G]$ on $\mathbf{k}^{(X)}$ as follows:

$$\left(\sum_{g\in G} \alpha_g e_g\right)\left(\sum_{x\in X} \beta_x e_x\right) = \sum_{g\in G}\sum_{x\in X} \alpha_g\beta_x e_{g\rightharpoonup x} \tag{105}$$
$$\text{for any } (\alpha_g)_{g\in G} \in \mathbf{k}[G] \text{ and } (\beta_x)_{x\in X} \in \mathbf{k}^{(X)}.$$

Using our convention to write $e_g$ as $g$ for each $g \in G$ (Convention 3.1.8), and a similar convention to write $e_x$ as $x$ for each $x \in X$, we can rewrite this equality as

$$\left(\sum_{g\in G} \alpha_g g\right)\left(\sum_{x\in X} \beta_x x\right) = \sum_{g\in G}\sum_{x\in X} \alpha_g\beta_x (g \rightharpoonup x). \tag{106}$$

This left $\mathbf{k}[G]$-module $\mathbf{k}^{(X)}$ is called the *permutation module* corresponding to the left $G$-set $X$. (It is named so because the elements of $G$ "permute" the basis vectors of $\mathbf{k}^{(X)}$.)

*Proof of Example 4.3.9 (sketched).* We must show that the module axioms hold. All of them are straightforward, and most are obvious. The only one that takes a bit of effort is the associativity law. So let us prove it. We must show that $(rs)\,m = r\,(sm)$ for all $r, s \in \mathbf{k}[G]$ and $m \in \mathbf{k}^{(X)}$.

Fix $r, s \in \mathbf{k}[G]$ and $m \in \mathbf{k}^{(X)}$. We must prove the equality $(rs)\,m = r\,(sm)$. However, the action of $\mathbf{k}[G]$ on $\mathbf{k}^{(X)}$ (more precisely, the map from $\mathbf{k}[G] \times \mathbf{k}^{(X)}$ to $\mathbf{k}^{(X)}$ that is defined by (105) and that we are claiming to be an action of $\mathbf{k}[G]$ on $\mathbf{k}^{(X)}$, even though we have not proved this yet) is easily seen to be $\mathbf{k}$-bilinear, because each addend $\alpha_g \beta_x e_{g \rightharpoonup x}$ on the right hand side of (105) depends $\mathbf{k}$-linearly on $\left(\alpha_g\right)_{g \in G}$ and depends $\mathbf{k}$-linearly on $\left(\beta_x\right)_{x \in X} \in \mathbf{k}^{(X)}$. Thus, both sides of the equality $(rs)\,m = r\,(sm)$ depend $\mathbf{k}$-linearly on $r$. Hence, in proving this equality, we can WLOG assume that $r$ is an element of the standard basis $\left(e_g\right)_{g \in G}$ of $\mathbf{k}[G]$. For similar reasons, we can WLOG assume that $s$ is an element of this standard basis as well, and that $m$ is an element of the standard basis $\left(e_x\right)_{x \in X}$ of $\mathbf{k}^{(X)}$. Let us make all these three assumptions. Hence,

$$r = e_g, \qquad s = e_h, \qquad \text{and} \qquad m = e_x \tag{107}$$

for some $g \in G$, $h \in G$ and $x \in X$. Consider these $g$, $h$ and $x$. Now, (105) easily yields

$$e_k e_y = e_{k \rightharpoonup y} \qquad \text{for all } k \in G \text{ and } y \in X \tag{108}$$

(indeed, this follows by applying (105) to $\alpha_g = \delta_{g,k}$ and $\beta_x = \delta_{x,y}$, where $\delta$ is the Kronecker delta). Now, from (107), we obtain

$$r\,(sm) = e_g \underbrace{\left(e_h e_x\right)}_{\substack{= e_{h \rightharpoonup x} \\ \text{(by (108))}}} = e_g e_{h \rightharpoonup x} = e_{g \rightharpoonup (h \rightharpoonup x)} \qquad \text{(by (108))}$$

and

$$\begin{aligned}(rs)\,m &= \underbrace{\left(e_g e_h\right)}_{\substack{= e_{gh} \\ \text{(by the definition of} \\ \text{multiplication in } \mathbf{k}[G]\text{)}}} e_x = e_{gh} e_x = e_{(gh) \rightharpoonup x} \qquad \text{(by (108))} \\ &= e_{g \rightharpoonup (h \rightharpoonup x)} \qquad \left(\begin{array}{c}\text{since } (gh) \rightharpoonup x = g \rightharpoonup (h \rightharpoonup x) \\ \text{(because } X \text{ is a left } G\text{-set)}\end{array}\right).\end{aligned}$$

Comparing these two equalities, we obtain $r\,(sm) = (rs)\,m$, which is precisely what we wanted to show. Thus, the associativity law has been verified, and the proof of Example 4.3.9 is complete. □

**Example 4.3.10.** In Example 4.3.9, we have defined permutation modules in general; let us now show a specific hands-on instance of this construction.

Let $A$ be any set. Let $k \in \mathbb{N}$. As we saw in Example 4.1.14, the symmetric group $S_A$ acts on the set $\mathcal{P}_k(A)$ of all $k$-element subsets of $A$. Thus, $\mathcal{P}_k(A)$ is a left $S_A$-set. For instance, taking $A = [4]$ and $k = 2$, we conclude that

$$\mathcal{P}_2([4]) = \{\{1,2\},\ \{1,3\},\ \{1,4\},\ \{2,3\},\ \{2,4\},\ \{3,4\}\}$$

is a left $S_{[4]}$-set, i.e., a left $S_4$-set, with action given by $g \rightharpoonup U = g(U)$ for all $g \in S_4$ and $U \in \mathcal{P}_2([4])$ (that is, $g \rightharpoonup \{u,v\} = \{g(u), g(v)\}$ for all $g \in S_4$ and $u \neq v$ in $[4]$).

The permutation module corresponding to this left $S_4$-set is a left $\mathbf{k}[S_4]$-module $\mathbf{k}^{(\mathcal{P}_2([4]))}$. As a $\mathbf{k}$-module, it is free with basis $\left(e_{\{1,2\}}, e_{\{1,3\}}, e_{\{1,4\}}, e_{\{2,3\}}, e_{\{2,4\}}, e_{\{3,4\}}\right)$. The action of $\mathbf{k}[S_4]$ is defined by (105). Thus, for instance, for any two permutations $g, h \in S_4$ and any two 2-element subsets $\{u,v\}$ and $\{x,y\}$ of $[4]$, we have

$$\begin{aligned}
&\left(e_g + e_h\right)\left(e_{\{u,v\}} + e_{\{x,y\}}\right) \\
&= e_{g \rightharpoonup \{u,v\}} + e_{g \rightharpoonup \{x,y\}} + e_{h \rightharpoonup \{u,v\}} + e_{h \rightharpoonup \{x,y\}} \\
&= e_{\{g(u),g(v)\}} + e_{\{g(x),g(y)\}} + e_{\{h(u),h(v)\}} + e_{\{h(x),h(y)\}}.
\end{aligned}$$

Using our convention to write $e_g$ as $g$ for each $g \in G$ (Convention 3.1.8), and a similar convention to write $e_x$ as $x$ for each $x \in X$, we can express this in the more pleasant form

$$\begin{aligned}
&(g+h)(\{u,v\} + \{x,y\}) \\
&= \{g(u), g(v)\} + \{g(x), g(y)\} + \{h(u), h(v)\} + \{h(x), h(y)\};
\end{aligned}$$

of course, it needs to be kept in mind that we are not adding the literal permutations $g, h$ or the literal subsets $\{u,v\}, \{x,y\}$ but rather the corresponding basis vectors of $\mathbf{k}[S_4]$ or $\mathbf{k}^{(\mathcal{P}_2([4]))}$.

Here is a simpler example of a left $\mathbf{k}[G]$-module:

**Example 4.3.11.** Let $G$ be a group. Any $\mathbf{k}$-module $V$ becomes a left $\mathbf{k}[G]$-module if we define the action of $\mathbf{k}[G]$ on $V$ by

$$\mathbf{a}v := \varepsilon(\mathbf{a})\, v \qquad \text{for all } \mathbf{a} \in \mathbf{k}[G] \text{ and } v \in V$$

(where $\varepsilon : \mathbf{k}[G] \to \mathbf{k}$ is the $\mathbf{k}$-algebra morphism that sends each $g \in G$ to 1, as suggested in Remark 3.12.4).

### 4.3.3. Properties and features

Many of the basic concepts related to left $R$-modules are not much different from their analogues for $\mathbf{k}$-modules. In particular, the following concepts can be defined just as in the case of $\mathbf{k}$-modules (with the obvious insertion of the word "left"):

- $R$-submodules of an $R$-module ([23wa, Definition 3.1.4]);
- quotients of $R$-modules by $R$-submodules ([23wa, Definition 3.6.1]);

- direct sums and direct products of $R$-modules ([23wa, §3.3.1]);
- morphisms of $R$-modules, also known as *left $R$-linear maps* ([23wa, Definition 3.5.1]);
- isomorphisms of $R$-modules ([23wa, Definition 3.5.1]);
- kernels and images of morphisms of $R$-modules ([23wa, §3.5.6]).

We refer to [23wa, Chapter 3] for more about these concepts. We shall also use some basic properties of these constructions, such as:

- The kernel and the image of an $R$-module morphism $f : V \to W$ are always $R$-submodules of $V$ and $W$, respectively ([23wa, §3.5.6]). We denote them by $\operatorname{Ker} f$ and $\operatorname{Im} f$, respectively (but usually we just write $f(V)$ for $\operatorname{Im} f$).
- The *first isomorphism theorem*[41]: If $f : V \to W$ is an $R$-module morphism, then $V/\operatorname{Ker} f \cong f(V)$ as $R$-modules ([23wa, Theorem 3.6.8]).

### 4.3.4. Representations of $G$ as left $\mathbf{k}[G]$-modules

Now we take a closer look at left $\mathbf{k}[G]$-modules when $G$ is a group. It turns out that they are just representations of $G$ over $\mathbf{k}$ in disguise:

**Theorem 4.3.12.** Let $G$ be a group.

**(a)** A left $\mathbf{k}[G]$-module $V$ is always a $\mathbf{k}$-module (by restriction of scalars, as defined in Definition 4.3.5) and (at the same time) a left $G$-set (with the left $G$-action defined by $g \rightharpoonup v = e_g v$ for all $g \in G$ and $v \in V$). Moreover, the left $G$-action on this left $G$-set is $\mathbf{k}$-linear, so that $V$ becomes a representation of $G$ over $\mathbf{k}$.

**(b)** Conversely: If $V$ is a representation of $G$ over $\mathbf{k}$, then we can turn $V$ into a left $\mathbf{k}[G]$-module by defining the action of $\mathbf{k}[G]$ on $V$ by

$$\left(\sum_{g\in G} \alpha_g g\right) v := \sum_{g\in G} \alpha_g (g \rightharpoonup v) \qquad (109)$$
$$\text{for all } \left(\alpha_g\right)_{g\in G} \in \mathbf{k}[G] \text{ and } v \in V.$$

**(c)** These two transformations (from left $\mathbf{k}[G]$-modules to representations of $G$ and backwards) are mutually inverse. That is:

- If we convert a left $\mathbf{k}[G]$-module $V$ into a representation of $G$ (according to Theorem 4.3.12 **(a)**), and then convert the resulting representation back into a left $\mathbf{k}[G]$-module (according to Theorem 4.3.12 **(b)**), then we recover the original left $\mathbf{k}[G]$-module $V$.

[41] also known as the *first homomorphism theorem*

- If we convert a representation $V$ of $G$ into a left $\mathbf{k}[G]$-module (according to Theorem 4.3.12 **(b)**), and then convert the resulting left $\mathbf{k}[G]$-module back into a representation of $G$ (according to Theorem 4.3.12 **(a)**), then we recover the original representation $V$.

*Proof sketch.* **(a)** This is straightforward. For instance, the associativity axiom $g \rightharpoonup (h \rightharpoonup x) = (gh) \rightharpoonup x$ follows from $e_g (e_h x) = \underbrace{(e_g e_h)}_{\substack{=e_{gh} \\ \text{(by the definition of} \\ \text{multiplication in } \mathbf{k}[G])}} x = e_{gh} x$.

**(b)** This is a fairly easy linearity argument, similar to the proof of Example 4.3.9. Here are the details:

Let $V$ be a representation of $G$ over $\mathbf{k}$. Define an action of $\mathbf{k}[G]$ on $V$ by (109). We must prove that it satisfies the module axioms. As usual, the hardest of them is the associativity law. So let us check it.

We must show that $(rs) m = r (sm)$ for all $r, s \in \mathbf{k}[G]$ and $m \in V$. The action of $\mathbf{k}[G]$ on $V$ (more precisely, the map from $\mathbf{k}[G] \times V$ to $V$ that is defined by (109) and that we are claiming to be an action of $\mathbf{k}[G]$ on $V$, even though we have not proved this yet) is easily seen to be $\mathbf{k}$-bilinear (indeed, (109) shows that it is $\mathbf{k}$-linear in $(\alpha_g)_{g \in G}$, whereas the $\mathbf{k}$-linearity of the $G$-action on $V$ entails that it is $\mathbf{k}$-linear in $v$). Thus, both sides of the equality $(rs) m = r (sm)$ (which we want to prove) depend $\mathbf{k}$-linearly on $r$. Hence, in proving this equality, we can WLOG assume that $r$ is an element of the standard basis $(e_g)_{g \in G}$ of $\mathbf{k}[G]$. For similar reasons, we can WLOG assume that $s$ is an element of this standard basis as well. Let us make both assumptions. Hence, $r = e_g$ and $s = e_h$ for some $g \in G$ and $h \in G$. Consider these $g$ and $h$. Following Convention 3.1.8, we rewrite $r = e_g$ and $s = e_h$ as $r = g$ and $s = h$.

Now, (109) easily yields

$$kv = k \rightharpoonup v \qquad \text{for all } k \in G \text{ and } v \in V \tag{110}$$

(indeed, this follows by applying (109) to $\alpha_g = \delta_{g,k}$, where $\delta$ is the Kronecker delta). Now, from $r = g$ and $s = h$, we obtain

$$\begin{aligned} (rs) m &= (gh) m = (gh) \rightharpoonup m \qquad \text{(by (110))} \\ &= g \rightharpoonup (h \rightharpoonup m) \qquad \text{(since } \rightharpoonup \text{ is a left } G\text{-action on } V\text{)} \end{aligned}$$

and

$$\begin{aligned} r (sm) &= g (hm) = g \rightharpoonup \underbrace{(hm)}_{\substack{=h \rightharpoonup m \\ \text{(by (110))}}} \qquad \text{(by (110))} \\ &= g \rightharpoonup (h \rightharpoonup m) . \end{aligned}$$

Comparing these two equalities, we obtain $(rs) m = r (sm)$, as desired. This completes the proof of Theorem 4.3.12 **(b)**.

**(c)** This is again easy. The "representation $\to$ $\mathbf{k}[G]$-module $\to$ representation" part is obvious. The "$\mathbf{k}[G]$-module $\to$ representation $\to$ $\mathbf{k}[G]$-module" direction requires a little linearity argument: Let $V$ be a left $\mathbf{k}[G]$-module. Convert it into a representation of $G$ (according to Theorem 4.3.12 **(a)**), denoting the $G$-action by $\rightharpoonup$ as usual. Then, convert the resulting representation back into a left $\mathbf{k}[G]$-module (according to Theorem 4.3.12 **(b)**). The action of $\mathbf{k}[G]$ on this new $\mathbf{k}[G]$-module is then given by (109). But the action of $\mathbf{k}[G]$ on the **original** $\mathbf{k}[G]$-module $V$ is also given by (109), since it satisfies

$$\begin{aligned}\left(\sum_{g\in G}\alpha_g g\right)v &= \sum_{g\in G}\alpha_g \underbrace{gv}_{\substack{=g\rightharpoonup v \\ \text{(by the definition of} \\ \text{the } G\text{-action } \rightharpoonup\text{)}}} && \text{(by } \mathbf{k}\text{-bilinearity)} \\ &= \sum_{g\in G}\alpha_g\,(g\rightharpoonup v) && \text{for all } \left(\alpha_g\right)_{g\in G}\in\mathbf{k}[G] \text{ and } v\in V.\end{aligned}$$

Hence, the two actions of $\mathbf{k}[G]$ are the same. But this means that the new left $\mathbf{k}[G]$-module is precisely the original $\mathbf{k}[G]$-module $V$. This proves Theorem 4.3.12 **(c)**. $\square$

A good way to think of Theorem 4.3.12 **(b)** is as "packing" a $\mathbf{k}$-module structure and a ($\mathbf{k}$-linear) left $G$-action into a left $\mathbf{k}[G]$-module structure (which combines them both, since $\mathbf{k}[G]$ includes both a "copy" of $\mathbf{k}$ and a "copy" of $G$). Likewise, Theorem 4.3.12 **(a)** "unpacks" a left $\mathbf{k}[G]$-module structure into a $\mathbf{k}$-module structure and a left $G$-action. Thus, in short, we have proved the following:

**Corollary 4.3.13.** Let $G$ be a group. A representation of $G$ over $\mathbf{k}$ is "the same as" a left $\mathbf{k}[G]$-module (since Theorem 4.3.12 allows us to convert either one into the other).

**Convention 4.3.14.** We shall thus use the notions of "representation of $G$ over $\mathbf{k}$" and "left $\mathbf{k}[G]$-module" interchangeably, trusting the reader to convert from one to the other at the slightest need. As a consequence, the notations "$g\rightharpoonup v$" and "$gv$" and "$e_g v$" (where $g$ belongs to the group $G$, and where $v$ belongs to the left $\mathbf{k}[G]$-module $V$) all mean the same thing and will be used as synonyms.

**Example 4.3.15.** Let $G$ be a group. Let $V$ be a trivial representation of $G$ (that is, a left $\mathbf{k}$-module equipped with a trivial $G$-action, as defined in Example 4.1.10). Thus, $V$ is a left $\mathbf{k}[G]$-module. Let us see explicitly how an element $\mathbf{a}=\sum_{g\in G}\alpha_g g\in\mathbf{k}[G]$ (with $\left(\alpha_g\right)_{g\in G}\in\mathbf{k}^{(G)}$) acts on a vector $v\in V$: Namely,

from $\mathbf{a} = \sum_{g \in G} \alpha_g g$, we obtain

$$\mathbf{a}v = \left( \sum_{g \in G} \alpha_g g \right) v = \sum_{g \in G} \alpha_g \underbrace{(g \rightharpoonup v)}_{\substack{=v \\ \text{(since the } G\text{-action} \\ \text{on } V \text{ is trivial)}}} \qquad \text{(by (109))}$$

$$= \sum_{g \in G} \alpha_g v = \left( \sum_{g \in G} \alpha_g \right) v. \tag{111}$$

We can rewrite this in a nicer form: Indeed, consider the $\mathbf{k}$-algebra morphism $\varepsilon : \mathbf{k}[G] \to \mathbf{k}$ that sends each $g \in G$ to 1 (as suggested in Remark 3.12.4). Then, for each $\mathbf{a} = \sum_{g \in G} \alpha_g g \in \mathbf{k}[G]$ (with $(\alpha_g)_{g \in G} \in \mathbf{k}^{(G)}$), we have

$$\varepsilon(\mathbf{a}) = \varepsilon \left( \sum_{g \in G} \alpha_g g \right) = \sum_{g \in G} \alpha_g \underbrace{\varepsilon(g)}_{=1} = \sum_{g \in G} \alpha_g.$$

Hence, we can rewrite (111) as

$$\mathbf{a}v = \varepsilon(\mathbf{a})\, v.$$

Thus, our action of $\mathbf{k}[G]$ on $V$ is the same action that we defined in Example 4.3.11. Thus, the latter action is just a trivial representation of $G$ in disguise.

Let us now define a more useful representation of any group $G$:

**Definition 4.3.16.** Let $G$ be a group. The *left regular representation* of $G$ is the left regular $\mathbf{k}[G]$-module $\mathbf{k}[G]$ (where the action is just the multiplication in $\mathbf{k}[G]$, as in Example 4.3.6). Thus, as a representation of $G$, it has its left $G$-action given by

$$g \rightharpoonup \mathbf{a} = g\mathbf{a} = e_g \mathbf{a} \qquad \text{for all } g \in G \text{ and } \mathbf{a} \in \mathbf{k}[G].$$

For example, the left regular representation of $S_n$ is $\mathbf{k}[S_n]$, with action by left multiplication:

$$g \rightharpoonup \mathbf{a} = g\mathbf{a} = e_g \mathbf{a} \qquad \text{for all } g \in S_n \text{ and } \mathbf{a} \in \mathbf{k}[S_n].$$

Convention 4.3.14 gives us a new way to think about representations of a group $G$. In particular, all concepts that are defined for left $\mathbf{k}[G]$-modules (e.g., submodules, quotients, morphisms, direct sums, direct products) can thus be automatically transferred to representations of $G$. Thus, we obtain concepts of subrepresentation, quotient of a representation, morphism of representations,

direct sum, direct product, etc. of representations of $G$ (by just viewing these representations as left $\mathbf{k}[G]$-modules).

However, some of these concepts have already been defined for representations before. To be safe, we need to verify that the concepts transferred from left $\mathbf{k}[G]$-modules agree with these formerly defined concepts for representations of $G$. But this is easy. For instance, we can easily convince ourselves that subrepresentations are the same thing as $\mathbf{k}[G]$-submodules:

**Proposition 4.3.17.** Let $G$ be a group. Let $V$ be a left $\mathbf{k}[G]$-module (i.e., a representation of $G$ over $\mathbf{k}$). Then, the left $\mathbf{k}[G]$-submodules of $V$ are precisely the subrepresentations of $V$.

*Proof sketch.* It is clear that any left $\mathbf{k}[G]$-submodule of $V$ is a subrepresentation of $V$. It remains to prove the converse, i.e., that any subrepresentation of $V$ is a left $\mathbf{k}[G]$-submodule of $V$.

Let $W$ be a subrepresentation of $V$. We must show that $W$ is a left $\mathbf{k}[G]$-submodule of $V$.

We know that $W$ is a subrepresentation of $V$. In other words, $W$ is a $\mathbf{k}$-submodule of $V$ that is simultaneously a $G$-subset of $V$. In particular, $W$ is a $G$-subset of $V$; hence,

$$g \rightharpoonup w \in W \qquad \text{for each } g \in G \text{ and } w \in W. \tag{112}$$

Let $\mathbf{a} \in \mathbf{k}[G]$ and $w \in W$. Write $\mathbf{a}$ as $\mathbf{a} = \sum_{g \in G} \alpha_g g$, where $(\alpha_g)_{g \in G} \in \mathbf{k}[G]$. Then,

$$\mathbf{a}w = \left( \sum_{g \in G} \alpha_g g \right) w = \sum_{g \in G} \alpha_g \underbrace{(g \rightharpoonup w)}_{\substack{\in W \\ \text{(by (112))}}} \qquad \text{(by (109))}$$
$$= (\text{a } \mathbf{k}\text{-linear combination of elements of } W) \in W$$

(since $W$ is a $\mathbf{k}$-submodule of $V$).

Forget that we fixed $\mathbf{a}$ and $w$. We thus have shown that $\mathbf{a}w \in W$ for each $\mathbf{a} \in \mathbf{k}[G]$ and $w \in W$. In other words, $W$ is a left $\mathbf{k}[G]$-submodule of $V$. This completes the proof of Proposition 4.3.17. □

Similarly, morphisms of representations of $G$ are the same thing as morphisms of left $\mathbf{k}[G]$-modules:

**Proposition 4.3.18.** Let $G$ be a group. Let $V$ and $W$ be two left $\mathbf{k}[G]$-modules (i.e., representations of $G$ over $\mathbf{k}$). Let $f : V \to W$ be a map. Then, $f$ is a morphism of left $\mathbf{k}[G]$-modules if and only if $f$ is a morphism of representations.

*Proof sketch.* $\Longrightarrow$: This is easy.

$\Longleftarrow$: Assume that $f$ is a morphism of representations. We must show that $f$ is a morphism of left $\mathbf{k}[G]$-modules.

We have assumed that $f$ is a morphism of representations. In other words, $f$ is both $\mathbf{k}$-linear (i.e., a morphism of $\mathbf{k}$-modules) and $G$-equivariant (i.e., a morphism of $G$-sets).

Now, let $\mathbf{a} \in \mathbf{k}[G]$ and $v \in V$. We shall show that $f(\mathbf{a}v) = \mathbf{a}f(v)$.
Write $\mathbf{a}$ as $\mathbf{a} = \sum_{g \in G} \alpha_g g$, where $(\alpha_g)_{g \in G} \in \mathbf{k}[G]$. Then,

$$\mathbf{a}v = \left( \sum_{g \in G} \alpha_g g \right) v = \sum_{g \in G} \alpha_g (g \rightharpoonup v) \qquad \text{(by (109))}.$$

Applying the map $f$ to both sides of this equality, we obtain

$$f(\mathbf{a}v) = f\left( \sum_{g \in G} \alpha_g (g \rightharpoonup v) \right) = \sum_{g \in G} \alpha_g \underbrace{f(g \rightharpoonup v)}_{\substack{= g \rightharpoonup f(v) \\ \text{(since } f \text{ is } G\text{-equivariant)}}} \qquad \text{(since } f \text{ is } \mathbf{k}\text{-linear)}$$
$$= \sum_{g \in G} \alpha_g (g \rightharpoonup f(v)).$$

Comparing this with

$$\mathbf{a}f(v) = \left( \sum_{g \in G} \alpha_g g \right) f(v) \qquad \left( \text{since } \mathbf{a} = \sum_{g \in G} \alpha_g g \right)$$
$$= \sum_{g \in G} \alpha_g (g \rightharpoonup f(v)) \qquad \left( \begin{array}{c} \text{by (109), applied to } W \text{ and } f(v) \\ \text{instead of } V \text{ and } v \end{array} \right),$$

we obtain $f(\mathbf{a}v) = \mathbf{a}f(v)$.

Forget that we fixed $\mathbf{a}$ and $v$. We have shown that every $\mathbf{a} \in \mathbf{k}[G]$ and $v \in V$ satisfy $f(\mathbf{a}v) = \mathbf{a}f(v)$. In other words, the map $f$ is a morphism of left $\mathbf{k}[G]$-modules (since $f$ is $\mathbf{k}$-linear and thus respects addition). Thus, the proof of Proposition 4.3.18 is complete. □

**Corollary 4.3.19.** Let $G$ be a group. Let $V$ and $W$ be two left $\mathbf{k}[G]$-modules (i.e., representations of $G$ over $\mathbf{k}$). Let $f : V \to W$ be a map. Then, $f$ is an isomorphism of left $\mathbf{k}[G]$-modules if and only if $f$ is an isomorphism of representations.

*Proof sketch.* This follows easily from Proposition 4.3.18. □

The reader is invited to convince himself that quotients of a representation of $G$ are the same thing as quotients of a left $\mathbf{k}[G]$-module.

**Exercise 4.3.1.** [1] Let $G$ be a group, and let $X$ be a left $G$-set. Consider the permutation module $\mathbf{k}^{(X)}$ defined in Example 4.3.9. Let $s := \sum_{x \in X} e_x \in \mathbf{k}^{(X)}$. Prove that $\operatorname{span}_{\mathbf{k}} \{s\}$ is a trivial subrepresentation of the $G$-representation $\mathbf{k}^{(X)}$. This subrepresentation is 1-dimensional (i.e., a free $\mathbf{k}$-module of rank 1) unless $X$ is empty.

### 4.3.5. Irreducible representations

Clearly, any representation $V$ of any group $G$ has two subrepresentations: the subset $\{0\}$ and $V$ itself. (These are distinct unless $V$ itself is $\{0\}$.) Sometimes, these are the only subrepresentations of $V$. In this case, we call $V$ *simple*:

**Definition 4.3.20.** Let $G$ be a group. A representation $V$ of $G$ (or, equivalently, a left $\mathbf{k}[G]$-module $V$) is said to be *simple* (aka *irreducible*) if it is not $\{0\}$, but its only subrepresentations are $V$ and $\{0\}$.

More generally, we can make the same definition for modules over a ring:

**Definition 4.3.21.** Let $R$ be a ring. A left $R$-module $V$ is said to be *simple* (aka *irreducible*) if it is not $\{0\}$, but its only $R$-submodules are $V$ and $\{0\}$.

Clearly, Definition 4.3.20 is the particular case of Definition 4.3.21 for $R = \mathbf{k}[G]$. Traditionally, the word "simple" is used for left $\mathbf{k}[G]$-modules, whereas the word "irreducible" applies to representations. But the two concepts are equivalent, and there is no reason other than habit to distinguish between them.

The following abbreviations are particularly popular with physicists: Irreducible representations are called *irreps*, whereas representations in general are called *reps*.

The irreducibility of a representation depends heavily on the base ring $\mathbf{k}$. The most well-behaved case is when $\mathbf{k}$ is a field:

**Example 4.3.22.** Let $\mathbf{k}$ be a field. The representations of the trivial group $\{1\}$ over $\mathbf{k}$ are just the $\mathbf{k}$-vector spaces (by Example 4.2.7). Their subrepresentations are simply their $\mathbf{k}$-vector subspaces. Thus, an irreducible representation of $\{1\}$ is simply a $\mathbf{k}$-vector space $V$ that is not $\{0\}$ but whose only vector subspaces are $V$ and $\{0\}$. What are these spaces?

If a $\mathbf{k}$-vector space $V$ has dimension 1, then its only vector subspaces are $V$ and $\{0\}$, and thus $V$ is an irreducible representation of $\{1\}$. But any $\mathbf{k}$-vector space $V$ of dimension $> 1$ has a subspace of dimension 1, and thus cannot be irreducible (since such a subspace is neither $V$ nor $\{0\}$). Thus, the irreducible representations of $\{1\}$ are exactly the 1-dimensional $\mathbf{k}$-vector spaces. Up to isomorphism, there is only one such ($\mathbf{k}$ itself).

**Example 4.3.23.** Let $\mathbf{k}$ be a field. Then, the natural representation $\mathbf{k}^n$ of $S_n$ is not irreducible (unless $n = 1$), but (for $n \geq 1$) its subrepresentation $D(\mathbf{k}^n)$ (defined in Subsection 4.2.6) is (since it is 1-dimensional as a $\mathbf{k}$-vector space, and thus has no subspaces other than itself and $\{0\}$). The subrepresentation $R(\mathbf{k}^n)$ is also irreducible if $n \geq 2$ and $\operatorname{char}\mathbf{k} \nmid n$ (this follows easily from Theorem 4.2.27), but not if $n > 2$ and $\operatorname{char}\mathbf{k} \mid n$ (since Proposition 4.2.28 **(b)** yields $D(\mathbf{k}^n) \subseteq R(\mathbf{k}^n)$ in this case).

Another irreducible representation of $S_n$ is the sign representation $\mathbf{k}_{\mathrm{sign}}$ (as defined in Example 4.2.18). This is again because this representation is 1-dimensional as a $\mathbf{k}$-vector space. (Note, however, that the sign representation is isomorphic to $R(\mathbf{k}^n)$ if $n = 2$.)

Are there more irreducible representations of $S_n$ ? Yes, and we will get to meet them.

**Example 4.3.24.** Let $\mathbf{k}$ be a field. The representations of the cyclic group $C_2 = \{1, g\}$ over $\mathbf{k}$ are $\mathbf{k}$-vector spaces equipped with a $\mathbf{k}$-linear involution $w$ (as we saw in Example 4.2.8). A subrepresentation of such a representation is a vector subspace $W$ that is preserved by this involution (i.e., that satisfies $w(W) \subseteq W$). Thus, the irreducible representations of $C_2$ are the nonzero $\mathbf{k}$-vector spaces $V$ equipped with a $\mathbf{k}$-linear involution $w : V \to V$ that preserves no vector subspaces of $V$ other than $V$ and $\{0\}$. What are these?

We claim the following:

**(a)** All irreducible representations of $C_2$ are 1-dimensional. Moreover, the $\mathbf{k}$-linear involution $w$ (that is, the action of the generator $g \in C_2$) on any such representation is either the identity map id or its negative $-\mathrm{id}$.

**(b)** If $\operatorname{char} \mathbf{k} \neq 2$, then $C_2$ has (up to isomorphism) exactly two different irreducible representations: the trivial representation (the vector space $\mathbf{k}$, with $g$ acting as id) and the "sign representation" (the vector space $\mathbf{k}$, with $g$ acting as $-\mathrm{id}$).

**(c)** If $\operatorname{char} \mathbf{k} = 2$, then $C_2$ has (up to isomorphism) only one irreducible representation: the trivial representation (the vector space $\mathbf{k}$, with $g$ acting as id).

*Proof of Example 4.3.24 (sketched).* **(a)** Let $V$ be an irreducible representation of $C_2$, and let $w$ be the corresponding $\mathbf{k}$-linear involution of $V$ (that is, the action of $g$). We must prove that $\dim V = 1$ and that $w$ is either id or $-\mathrm{id}$.

Since $w$ is an involution, we have $w^2 = \mathrm{id} = 1$, so that $w^2 - 1 = 0$. But $w^2 - 1 = (w-1)(w+1)$ (since $w$ is a $\mathbf{k}$-linear map). Thus, $(w-1)(w+1) = 0$.

We are in one of the two cases:

*Case 1:* We have $\operatorname{Ker}(w-1) = \{0\}$.

*Case 2:* We have $\operatorname{Ker}(w-1) \neq \{0\}$.

Let us first consider Case 1. In this case, we have $\operatorname{Ker}(w-1) = \{0\}$. In other words, the $\mathbf{k}$-linear map $w - 1$ is injective. Hence, we can cancel $w - 1$ from $(w-1)(w+1) = 0$, obtaining $w + 1 = 0$. Thus, $w = -1 = -\mathrm{id}$. Hence, any vector subspace of $V$ is automatically preserved by $w$ (since $-\mathrm{id}$ preserves any subspace), and thus is a subrepresentation of $V$. If $\dim V > 1$, then $V$ has at least one vector subspace that is neither $V$ nor $\{0\}$, and thus (by the preceding sentence) at least one subrepresentation that is neither $V$ nor $\{0\}$; but this contradicts the irreducibility of $V$. Hence, we cannot have $\dim V > 1$. Thus, $\dim V \leq 1$. But the irreducibility of $V$ also yields $V \neq \{0\}$ and therefore $\dim V \geq 1$. Combining, we find $\dim V = 1$. Hence, we have shown that $\dim V = 1$ and $w = -\mathrm{id}$. Thus, Example 4.3.24 **(a)** is proved in Case 1.

Let us now consider Case 2. In this case, we have $\operatorname{Ker}(w-1) \neq \{0\}$. Hence, there exists a nonzero vector $v \in \operatorname{Ker}(w-1)$. Consider this $v$. From $v \in \operatorname{Ker}(w-1)$, we obtain $(w-1)(v) = 0$, so that $w(v) - v = 0$ and thus $w(v) = v$. Therefore, the 1-dimensional vector subspace $\operatorname{span}_{\mathbf{k}}\{v\}$ of $V$ is preserved by $w$. In other words, $\operatorname{span}_{\mathbf{k}}\{v\}$ is a subrepresentation of $V$. By the irreducibility of $V$, this subrepresentation must be either $V$ or $\{0\}$. Hence, $\operatorname{span}_{\mathbf{k}}\{v\}$ must be $V$ (since it cannot be $\{0\}$). But this necessarily entails that $\dim V = 1$ and $w = \mathrm{id}$ (since $w(v) = v$). Thus, Example 4.3.24 **(a)** is proved in Case 2.

We have now proved Example 4.3.24 **(a)** in both Cases 1 and 2.

**(b)** From $\operatorname{char}\mathbf{k} \neq 2$, we obtain $1 \neq -1$ in $\mathbf{k}$. Thus, the trivial representation and the "sign representation" of $C_2$ are not isomorphic (since $w$ acts as the identity on the former but not on the latter). Moreover, Example 4.3.24 **(a)** shows that any irreducible representation of $C_2$ is isomorphic to one of the two. This proves Example 4.3.24 **(b)**.

**(c)** This is just like part **(b)**, but with one difference: Since $\operatorname{char}\mathbf{k} = 2$, we now have $1 = -1$ in $\mathbf{k}$, so that the trivial representation and the "sign representation" are the same thing. □

This last example shows that the description of the irreps of a group $G$ over $\mathbf{k}$ can depend on $\operatorname{char}\mathbf{k}$. The situation where $\mathbf{k}$ is not a field is even weirder:

**Example 4.3.25.** The irreducible representations of the trivial group $\{1\}$ over $\mathbb{Z}$ are the simple $\mathbb{Z}[\{1\}]$-modules, i.e., the simple $\mathbb{Z}$-modules (since $\mathbb{Z}[\{1\}] \cong \mathbb{Z}$). They are (up to isomorphism) the $\mathbb{Z}$-modules $\mathbb{Z}/p$ for all primes $p$.

**Exercise 4.3.2.** Consider the cyclic group $C_3$.

**(a)** $\boxed{2}$ Prove that $C_3$ has (up to isomorphism) three different irreducible representations over $\mathbb{C}$, and each of them has dimension 1.

**(b)** $\boxed{2}$ Prove that $C_3$ has (up to isomorphism) two different irreducible representations over $\mathbb{R}$: one of dimension 1 and one of dimension 2.

[**Hint:** What eigenvalues can a $\mathbb{C}$-linear map (i.e., complex matrix) $w$ satisfying $w^3 = \mathrm{id}$ have?]

### 4.3.6. Some further properties of representations/modules

We recall a basic concept from linear algebra ([23wa, Exercise 3.11.3]):

**Definition 4.3.26.** Let $V$ be a $\mathbf{k}$-module. Then, we set

$$\operatorname{End}_{\mathbf{k}} V := \{\mathbf{k}\text{-module endomorphisms of } V\} = \{\mathbf{k}\text{-linear maps } V \to V\}.$$

This is a $\mathbf{k}$-algebra (with addition and scaling being pointwise[42], and with multiplication being composition of maps). It is called the *endomorphism ring* of $V$. Its group of units (i.e., invertible elements) is the general linear group $\operatorname{GL}(V)$.

Note that when $V$ is the free $\mathbf{k}$-module $\mathbf{k}^n$ for some $n \in \mathbb{N}$, the endomorphism ring $\operatorname{End}_{\mathbf{k}} V$ is isomorphic to the matrix ring $\mathbf{k}^{n \times n}$ (as a $\mathbf{k}$-algebra). The isomorphism simply sends a $\mathbf{k}$-linear map $V \to V$ to the matrix that represents this map in the standard basis of $V$.

We can now define an analogue of currying for modules over any $\mathbf{k}$-algebra:

**Theorem 4.3.27.** Let $R$ be a $\mathbf{k}$-algebra. Let $V$ be a left $R$-module. Then, for each $r \in R$, the map

$$\begin{aligned} \tau_r : V &\to V, \\ v &\mapsto rv \end{aligned}$$

is $\mathbf{k}$-linear and thus belongs to $\operatorname{End}_{\mathbf{k}} V$. Moreover, the map

$$\begin{aligned} \rho : R &\to \operatorname{End}_{\mathbf{k}} V, \\ r &\mapsto \tau_r \end{aligned}$$

is a $\mathbf{k}$-algebra morphism.

*Proof sketch.* Straightforward and left to the reader. (For instance, $\tau_{rs} = \tau_r \tau_s$ follows from the associativity law $(rs)\, m = r\,(sm)$ for the left $R$-module $V$. Most other properties follow from the $\mathbf{k}$-bilinearity of the action $R \times V \to V$, $(r, v) \mapsto rv$.) □

**Definition 4.3.28.** The morphism $\rho$ in Theorem 4.3.27 is called the *curried form* of the left $R$-action on $V$ (as a $\mathbf{k}$-module).

The map $\tau_r$ in Theorem 4.3.27 is called the *action* of $r$ on $V$.

This notion of "action" is perfectly aligned with the notion of "action" for a $G$-set (see Definition 4.1.20): Indeed, if $G$ is a group and if $V$ is a left $\mathbf{k}[G]$-module (i.e., a representation of $G$), then the action of a group element $g \in G$ on $V$ in the sense of Definition 4.1.20 is precisely the action of the corresponding standard basis vector $g = e_g \in \mathbf{k}[G]$ on $V$ in the sense of Definition 4.3.28 (since $g \rightharpoonup v = e_g v$ for each $v \in V$).

Currying lets us define the notion of a faithful module:

[42] That is:

- The sum $f + g$ of two endomorphisms $f, g \in \operatorname{End}_{\mathbf{k}} V$ is defined by
$$(f+g)(v) = f(v) + g(v) \qquad \text{for all } v \in V.$$
- Scaling an endomorphism $f \in \operatorname{End}_{\mathbf{k}} V$ by a scalar $\lambda \in \mathbf{k}$ results in an endomorphism $\lambda f \in \operatorname{End}_{\mathbf{k}} V$ defined by
$$(\lambda f)(v) = \lambda f(v) \qquad \text{for all } v \in V.$$

**Definition 4.3.29.** Let $R$ be a $\mathbf{k}$-algebra. A left $R$-module $V$ is said to be *faithful* if it has the following property: If $r, s \in R$ are two elements that satisfy

$$(rv = sv \qquad \text{for all } v \in V),$$

then $r = s$.

In other words, a left $R$-module $V$ is said to be *faithful* if the curried form $\rho : R \to \operatorname{End}_{\mathbf{k}} V$ of its action is injective.

Roughly speaking, a left $R$-module $V$ is faithful if and only if you can tell elements of $R$ apart by their action on $V$ (that is, two different elements of $R$ never act the same way on $V$).

**Warning 4.3.30.** Let $G$ be a group. A faithful left $\mathbf{k}[G]$-module is not the same as a faithful representation of $G$. Indeed, for a representation of $G$ to be faithful, we only need different elements of $G$ to act differently. However, for a left $\mathbf{k}[G]$-module to be faithful, we need different **linear combinations** of elements of $G$ to act differently. This is a much stronger requirement! Here is an example for this difference:

**Example 4.3.31.** Consider the natural representation $\mathbf{k}^n$ of the symmetric group $S_n$.

This representation is faithful as a representation of $S_n$ (if $\mathbf{k}$ is nontrivial), since different permutations $g \in S_n$ can be told apart by their action on the standard basis vectors $e_1, e_2, \ldots, e_n$ of $\mathbf{k}^n$.

However, the natural representation $\mathbf{k}^n$ is **not** faithful as a $\mathbf{k}[S_n]$-module, unless $n \leq 2$. For example, let $n \geq 3$, and consider the element

$$\nabla^-_{[3]} = 1 - s_1 - s_2 - t_{1,3} + \operatorname{cyc}_{1,2,3} + \operatorname{cyc}_{1,3,2} \in \mathbf{k}[S_n].$$

We claim that this element $\nabla^-_{[3]}$ (which is nonzero unless $\mathbf{k}$ is trivial) acts as 0 on $\mathbf{k}^n$, meaning that $\nabla^-_{[3]} v = 0$ for all $v \in \mathbf{k}^n$. In order to prove this, it suffices to show that $\nabla^-_{[3]} e_k = 0$ for each $k \in [n]$ (since the standard basis vectors $e_1, e_2, \ldots, e_n$ span $\mathbf{k}^n$). We can check this directly: We have

$$\begin{aligned}\nabla^-_{[3]} e_1 &= \underbrace{1e_1}_{=e_1} - \underbrace{s_1 e_1}_{=e_2} - \underbrace{s_2 e_1}_{=e_1} - \underbrace{t_{1,3} e_1}_{=e_3} + \underbrace{\operatorname{cyc}_{1,2,3} e_1}_{=e_2} + \underbrace{\operatorname{cyc}_{1,3,2} e_1}_{=e_3} \\ &= e_1 - e_2 - e_1 - e_3 + e_2 + e_3 = 0\end{aligned}$$

and similarly $\nabla^-_{[3]} e_2 = 0$ and $\nabla^-_{[3]} e_3 = 0$, and it is even easier to see that $\nabla^-_{[3]} e_k = 0$ for all $k \geq 4$.

More generally, if $X$ is any subset of $[n]$ satisfying $|X| > 2$, then

$$\nabla^-_X v = 0 \qquad \text{for any } v \in \mathbf{k}^n.$$

The simplest example of a faithful $R$-module (for any ring $R$) is $R$ itself:

**Example 4.3.32.** Let $R$ be any ring. Then, the left regular $R$-module $R$ is faithful. Indeed, if $r, s \in R$ are two elements that satisfy

$$(rv = sv \qquad \text{for all } v \in R),$$

then $r = s$ (because applying $rv = sv$ to $v = 1$ yields $r1 = s1$, that is, $r = s$).

The following exercise shows that a faithful left $\mathbf{k}[G]$-module (for a group $G$) has to be "not too small" (whereas a faithful representation of a group $G$ can often be 1-dimensional[43]):

**Exercise 4.3.3.** $\boxed{1}$ Let $\mathbf{k}$ be a field. Let $G$ be a group. Let $V$ be a faithful left $\mathbf{k}[G]$-module. Prove that $\dim V \geq \sqrt{|G|}$, where $\dim V$ denotes the dimension of $V$ as a $\mathbf{k}$-vector space. (If $G$ is infinite, then this is saying that $V$ is infinite-dimensional.)

## 4.4. Direct addends, projections, filtrations, Maschke, Jordan–Hölder

Direct addends of an $R$-module are a particularly well-behaved kind of submodules. We will first briefly discuss them for arbitrary rings $R$, then focus on the case when $R$ is the group algebra $\mathbf{k}[G]$ of a group $G$.

### 4.4.1. Direct addends and projections

First, let us do some elementary linear algebra, albeit with $R$-modules instead of vector spaces.

Recall the following: If $R$ is a ring, and if $W_1, W_2, \ldots, W_k$ are finitely many left $R$-modules, then the direct sum $W_1 \oplus W_2 \oplus \cdots \oplus W_k$ of these left $R$-modules is the left $R$-module

$$\begin{aligned} W_1 \oplus W_2 \oplus \cdots \oplus W_k &= W_1 \times W_2 \times \cdots \times W_k \\ &= \{(w_1, w_2, \ldots, w_k) \mid w_i \in W_i \text{ for each } i\} \end{aligned}$$

with entrywise addition and entrywise action of $R$. (Direct sums of infinitely many addends are a bit more complicated, since – unlike direct products – they only allow essentially finite families of vectors. See [23wa, Definition 3.3.5] for the proper definition of a direct sum in this case.)

This kind of direct sum is called an *external direct sum*, since it constructs a new $R$-module $(W_1 \oplus W_2 \oplus \cdots \oplus W_k)$ from existing ones $(W_1, W_2, \ldots, W_k)$. Sometimes, this feels unnecessary, since $W_1, W_2, \ldots, W_k$ are already contained in a

[43] More precisely: Any cyclic group has a 1-dimensional faithful representation over $\mathbb{C}$ (why?). Non-cyclic groups don't.

single $R$-module $V$ and are "sufficiently spread out" that their direct sum can be identified with an $R$-submodule of $V$ as well. In this situation, the latter $R$-submodule is called the *internal direct sum* of $W_1, W_2, \ldots, W_k$. Here is the actual definition:

**Definition 4.4.1.** Let $R$ be a ring. Let $V$ be a left $R$-module. Let $W_1, W_2, \ldots, W_k$ be $R$-submodules of $V$.

**(a)** The *sum* $W_1 + W_2 + \cdots + W_k$ of these $R$-submodules is defined to be the set

$$\{w_1 + w_2 + \cdots + w_k \mid w_i \in W_i \text{ for each } i \in [k]\}.$$

This is an $R$-submodule of $V$ that contains $W_1, W_2, \ldots, W_k$ as $R$-submodules. It is furthermore the smallest $R$-submodule of $V$ with this property.

**(b)** There is a canonical $R$-module morphism

$$\begin{aligned} \operatorname{sum} : W_1 \oplus W_2 \oplus \cdots \oplus W_k &\to W_1 + W_2 + \cdots + W_k, \\ (w_1, w_2, \ldots, w_k) &\mapsto w_1 + w_2 + \cdots + w_k. \end{aligned}$$

This morphism is always surjective.

**(c)** If this $R$-module morphism sum is also injective (and thus an isomorphism), then we call the sum $W_1 + W_2 + \cdots + W_k$ an *internal direct sum*, and we denote it by $W_1 \oplus W_2 \oplus \cdots \oplus W_k$ as well (although, strictly speaking, it is merely isomorphic to $W_1 \oplus W_2 \oplus \cdots \oplus W_k$), thus implicitly identifying each element $(w_1, w_2, \ldots, w_k)$ of the external direct sum $W_1 \oplus W_2 \oplus \cdots \oplus W_k$ with the element $\operatorname{sum}(w_1, w_2, \ldots, w_k) = w_1 + w_2 + \cdots + w_k$ of the internal direct sum $W_1 + W_2 + \cdots + W_k$. In this situation, we also say that *the sum* $W_1 + W_2 + \cdots + W_k$ *is direct*, or that the $R$-submodules $W_1, W_2, \ldots, W_k$ are *linearly disjoint*.

This definition is frequently used in abstract linear algebra, e.g., when saying that the eigenspaces of a linear map have a direct sum (which almost always means the internal direct sum). For example, the two $\mathbb{Q}$-vector subspaces $\operatorname{span}_{\mathbb{Q}}\{e_1, e_2\}$ and $\operatorname{span}_{\mathbb{Q}}\{e_3 + e_4\}$ of $\mathbb{Q}^4$ are linearly disjoint and have the internal direct sum $\operatorname{span}_{\mathbb{Q}}\{e_1, e_2, e_3 + e_4\}$, whereas the two $\mathbb{Q}$-vector subspaces $\operatorname{span}_{\mathbb{Q}}\{e_1, e_2\}$ and $\operatorname{span}_{\mathbb{Q}}\{e_1 + e_2, e_3 + e_4\}$ of $\mathbb{Q}^4$ are not linearly disjoint and thus don't have an internal direct sum (their sum $\operatorname{span}_{\mathbb{Q}}\{e_1, e_2, e_1 + e_2, e_3 + e_4\}$ is only 3-dimensional, since the vector $e_1 + e_2$ is redundant).

You can think of linear disjointness as a version of linear dependence for submodules instead of single vectors. It is different in one aspect, however: It does not depend on $R$. In fact, the definition of linear disjointness involves only the addition on $V$, not the action of $R$ on $V$. Thus, whether or not some $R$-submodules $W_1, W_2, \ldots, W_k$ are linearly disjoint does not change if we replace $R$ by a subring of $R$. We could just as well have replaced "$R$-module" by "additive abelian group" throughout Definition 4.4.1 (though this would not

have increased its generality, since additive abelian groups are just $\mathbb{Z}$-modules).

Here is a well-known alternative characterization of linear disjointness:

**Proposition 4.4.2.** Let $R$ be a ring. Let $V$ be a left $R$-module. Let $W_1, W_2, \ldots, W_k$ be $R$-submodules of $V$. Then, the following statements are equivalent:

- The sum $W_1 + W_2 + \cdots + W_k$ is direct (i.e., the $R$-submodules $W_1, W_2, \ldots, W_k$ are linearly disjoint).
- The only $k$-tuple $(w_1, w_2, \ldots, w_k) \in W_1 \times W_2 \times \cdots \times W_k$ that satisfies $w_1 + w_2 + \cdots + w_k = 0$ is the $k$-tuple $(0, 0, \ldots, 0)$.
- For each $v \in V$, there exists **at most one** $k$-tuple $(w_1, w_2, \ldots, w_k) \in W_1 \times W_2 \times \cdots \times W_k$ that satisfies $w_1 + w_2 + \cdots + w_k = v$.

Linear disjointness is particularly simple for two submodules:

**Proposition 4.4.3.** Let $R$ be a ring. Let $V$ be a left $R$-module. Let $W_1$ and $W_2$ be two $R$-submodules of $V$. Then, the sum $W_1 + W_2$ is direct if and only if $W_1 \cap W_2 = \{0\}$.

However, the case of several submodules is less simple:

**Warning 4.4.4.** Let $R$ be a ring. Let $V$ be a left $R$-module. Let $W_1, W_2, \ldots, W_k$ be $R$-submodules of $V$. In order for the sum $W_1 + W_2 + \cdots + W_k$ to be direct, it is necessary **but not sufficient** that $W_i \cap W_j = \{0\}$ for all $i \neq j$. In other words, linear disjointness of $W_1, W_2, \ldots, W_k$ is **not** a "pairwise" property.

For an example, the three $\mathbb{Q}$-vector subspaces $\operatorname{span}_{\mathbb{Q}} \{e_1\}$ and $\operatorname{span}_{\mathbb{Q}} \{e_2\}$ and $\operatorname{span}_{\mathbb{Q}} \{e_1 + e_2\}$ of $\mathbb{Q}^2$ are not linearly disjoint, even though any two of them are.

**Exercise 4.4.1.** $\boxed{2}$ Let $R$ be a ring. Let $V$ be a left $R$-module. Let $X_1, X_2, \ldots, X_i, Y_1, Y_2, \ldots, Y_j$ be $R$-submodules of $V$. Prove that the sum $X_1 + X_2 + \cdots + X_i + Y_1 + Y_2 + \cdots + Y_j$ is direct if and only if the three sums

$$\begin{aligned} &X_1 + X_2 + \cdots + X_i, \\ &Y_1 + Y_2 + \cdots + Y_j, \qquad \text{and} \\ &(X_1 + X_2 + \cdots + X_i) + (Y_1 + Y_2 + \cdots + Y_j) \end{aligned}$$

are direct. (Note that the last of these sums is a sum of two addends only! And note that this is not obvious.)

Let us now introduce a notation for the sort of $R$-submodules that are part of a direct sum:

**Definition 4.4.5.** Let $R$ be a ring. Let $W$ be an $R$-submodule of a left $R$-module $V$. We say that $W$ is a *direct addend* of $V$ if there exists a further $R$-submodule $U$ of $V$ such that $V = W \oplus U$ (internal direct sum). In this case, the latter $R$-submodule $U$ is called a *complement* of $W$ in $V$. (Not "the complement" because $U$ is not uniquely determined by $V$ and $W$.)

**Example 4.4.6.** The $\mathbb{Z}$-submodule $\operatorname{span}_{\mathbb{Z}}\{(1,1)\}$ of $\mathbb{Z}^2$ is a direct addend of $\mathbb{Z}^2$, since $\operatorname{span}_{\mathbb{Z}}\{(0,1)\}$ is a complement of it. It has many other complements as well. In fact, it is not hard to see that two "one-dimensional" $\mathbb{Z}$-submodules $\operatorname{span}_{\mathbb{Z}}\{u\}$ and $\operatorname{span}_{\mathbb{Z}}\{v\}$ of $\mathbb{Z}^2$ are complements of each other if and only if $(u, v)$ is a basis of $\mathbb{Z}^2$.

**Example 4.4.7.** The $\mathbb{Z}$-submodule $2\mathbb{Z} = \{\text{even integers}\}$ of $\mathbb{Z}$ is not a direct addend, since it has no complement. (The easiest way to see this is to show that no nonzero submodule of $\mathbb{Z}$ is linearly disjoint from $2\mathbb{Z}$, but $2\mathbb{Z} \oplus 0$ is not $\mathbb{Z}$.) That is, we cannot write $\mathbb{Z}$ as an internal direct sum $2\mathbb{Z} \oplus (\text{another submodule of } \mathbb{Z})$.

**Remark 4.4.8.** As we have seen above, the linear disjointness of two given $R$-submodules of a given $R$-module $V$ does not depend on $R$. However, whether a given $R$-submodule is a direct addend does depend on $R$: It can have a complement for some choices of $R$ but not for others. In Example 4.4.12, we will see an instance of this.

As we see from Example 4.4.7, not every submodule is a direct addend. But over a field, it is:

**Proposition 4.4.9.** Let $\mathbf{k}$ be a field. Then, any vector subspace of a $\mathbf{k}$-vector space $V$ is a direct addend of $V$.

*Proof idea.* Let $W$ be a vector subspace of $V$. We must show that $W$ is a direct addend of $V$.

Pick a basis $(e_i)_{i \in I}$ of $W$, and extend it to a basis $(e_i)_{i \in J}$ of $V$. (Both of these can be done, since $\mathbf{k}$ is a field.)

Let $U$ be the span of the "new" basis elements – i.e., let $U = \operatorname{span}_{\mathbf{k}}\{e_i \mid i \in J \setminus I\}$. Then, $U$ is a $\mathbf{k}$-submodule of $V$. Moreover, it is easy to see that $W \cap U = \{0\}$ (since $W$ and $U$ are spanned by two disjoint subsets of the basis $(e_i)_{i \in J}$). Hence, the sum $W + U$ is direct (by Proposition 4.4.3). Thus, $W + U = W \oplus U$. However, $W = \operatorname{span}_{\mathbf{k}}\{e_i \mid i \in I\}$ (since $(e_i)_{i \in I}$ is a basis of $W$) and $U = \operatorname{span}_{\mathbf{k}}\{e_i \mid i \in J \setminus I\}$, so that

$$\begin{aligned} W + U &= \operatorname{span}_{\mathbf{k}}\{e_i \mid i \in I\} + \operatorname{span}_{\mathbf{k}}\{e_i \mid i \in J \setminus I\} \\ &= \operatorname{span}_{\mathbf{k}} \underbrace{\left(\{e_i \mid i \in I\} \cup \{e_i \mid i \in J \setminus I\}\right)}_{=\{e_i \mid i \in J\}} \\ &= \operatorname{span}_{\mathbf{k}}\{e_i \mid i \in J\} = V \end{aligned}$$

(since $(e_i)_{i\in J}$ is a basis of $V$). Thus, $V = W + U = W \oplus U$. This shows that $W$ is a direct addend of $V$. Proposition 4.4.9 is thus proved. □

The direct addends of an $R$-module can also be characterized in terms of projections. We recall what a projection is:

**Definition 4.4.10.** Let $W$ be a subset of a set $V$. A *projection* from $V$ to $W$ means a map $\pi : V \to W$ such that $\pi|_W = \mathrm{id}$ (that is, $\pi(w) = w$ for each $w \in W$).

Clearly, any projection is surjective. An example of a projection is the "floor function" – i.e., the map $\mathbb{R} \to \mathbb{Z}$, $x \mapsto \lfloor x \rfloor$ which sends each real number $x$ to its integer part $\lfloor x \rfloor$ (that is, the largest integer $\leq x$). But we are more interested in $R$-linear projections.

**Proposition 4.4.11.** Let $R$ be a ring. Let $V$ be a left $R$-module. Let $W$ be an $R$-submodule of $V$. Then, $W$ is a direct addend of $V$ if and only if there exists a left $R$-linear projection $\pi : V \to W$.

*Proof.* $\Longrightarrow$: Assume that $W$ is a direct addend. Thus, $V = W \oplus U$ (internal direct sum) for some $R$-submodule $U$ of $V$. Consider this $U$.

Let us agree to use the symbol "$\oplus$" only for internal direct sums in this proof, while using the symbol "$\times$" for external direct sums. (This is legitimate, since we are only dealing with direct sums of finitely many $R$-modules, and these direct sums are the same as the corresponding direct products.)

We have $V = W \oplus U = W + U$, and moreover, the sum $W + U$ is direct (since we are writing it as $W \oplus U$). The latter fact means that the map

$$\begin{aligned} \mathrm{sum} : W \times U &\to W + U, \\ (w, u) &\mapsto w + u \end{aligned}$$

(recall that $W \times U$ is the external direct sum of $W$ and $U$) is injective. Hence, this map sum is bijective (since it is clearly surjective), and therefore an $R$-module isomorphism (since it is obviously left $R$-linear). Thus, its inverse map $\mathrm{sum}^{-1} : W + U \to W \times U$ is an $R$-module morphism as well. In view of $W + U = V$, we can rewrite the latter map as $\mathrm{sum}^{-1} : V \to W \times U$. Explicitly, this map $\mathrm{sum}^{-1}$ sends each $v \in V$ to the unique pair $(w, u) \in W \times U$ that satisfies $v = w + u$.

Let $\rho : W \times U \to W$ be the map that sends each pair $(w, u)$ to its first entry $w$. This is also a left $R$-linear map.

Consider the composition $\rho \circ \mathrm{sum}^{-1} : V \to W$ of these two maps $\rho : W \times U \to W$ and $\mathrm{sum}^{-1} : V \to W \times U$. This composition is left $R$-linear (since both $\rho$ and $\mathrm{sum}^{-1}$ are left $R$-linear). Moreover, each $w \in W$ satisfies $\mathrm{sum}(w, 0) = w + 0 = w$ and thus $\mathrm{sum}^{-1}(w) = (w, 0)$, so that

$$\left(\rho \circ \mathrm{sum}^{-1}\right)(w) = \rho\left(\underbrace{\mathrm{sum}^{-1}(w)}_{=(w,0)}\right) = \rho(w, 0) = w$$

(by the definition of $\rho$). In other words, $\left(\rho \circ \operatorname{sum}^{-1}\right)|_W = \operatorname{id}$. Hence, the map $\rho \circ \operatorname{sum}^{-1} : V \to W$ is a projection from $V$ to $W$. Thus, there exists a left $R$-linear projection $\pi : V \to W$ (namely, $\pi = \rho \circ \operatorname{sum}^{-1}$). This proves the "$\Longrightarrow$" direction of Proposition 4.4.11.

$\Longleftarrow$: Assume that there is a left $R$-linear projection $\pi : V \to W$. Consider this $\pi$.

Now define $U = \operatorname{Ker} \pi$. This is clearly an $R$-submodule of $W$. We shall now show that $V = W \oplus U$ (internal direct sum). Indeed:

- The sum $W + U$ is direct. [*Proof:* Let $x \in W \cap U$. Thus, $x \in W \cap U \subseteq W$, so that $\pi(x) = x$ (since $\pi : V \to W$ is a projection and thus satisfies $\pi(w) = w$ for all $w \in W$). But we also have $x \in W \cap U \subseteq U = \operatorname{Ker} \pi$ and thus $\pi(x) = 0$. Comparing this with $\pi(x) = x$, we find $x = 0$. Forget that we fixed $x$. We thus have shown that $x = 0$ for each $x \in W \cap U$. In other words, $W \cap U \subseteq \{0\}$. Hence, $W \cap U = \{0\}$ (since $W \cap U$ is an $R$-submodule of $V$). Thus, the sum $W + U$ is direct (by Proposition 4.4.3).]

- We have $V = W + U$. [*Proof:* Let $v \in V$. Then, $\pi(v) \in W$ (since $\pi$ is a map from $V$ to $W$). But $\pi : V \to W$ is a projection and thus satisfies $\pi(w) = w$ for all $w \in W$. Applying this to $w = \pi(v)$, we obtain $\pi(\pi(v)) = \pi(v)$ (since $\pi(v) \in W$). Since $\pi$ is a linear map, we have $\pi(v - \pi(v)) = \pi(v) - \underbrace{\pi(\pi(v))}_{=\pi(v)} = \pi(v) - \pi(v) = 0$. Thus, $v - \pi(v) \in \operatorname{Ker} \pi = U$. Hence, $v = \underbrace{\pi(v)}_{\in W} + \underbrace{v - \pi(v)}_{\in U} \in W + U$.

  Forget that we fixed $v$. We thus have shown that $v \in W + U$ for each $v \in V$. In other words, $V \subseteq W + U$. Hence, $V = W + U$ (since we clearly have $W + U \subseteq V$).]

Now, $V = W + U = W \oplus U$ (since the sum $W + U$ is direct). Hence, $W$ is a direct addend of $V$. This proves the "$\Longleftarrow$" direction of Proposition 4.4.11. $\square$

**Example 4.4.12.** Consider again the natural representation $\mathbf{k}^n$ of $S_n$. This is a left $\mathbf{k}[S_n]$-module. Define its $\mathbf{k}[S_n]$-submodule $D(\mathbf{k}^n)$ as in Subsection 4.2.6.

For any $n$-tuple $(a_1, a_2, \ldots, a_n) \in \mathbf{k}^n$, set $a_\Sigma := a_1 + a_2 + \cdots + a_n$.

**(a)** Consider the $\mathbf{k}$-linear map

$$f_1 : \mathbf{k}^n \to D(\mathbf{k}^n),$$
$$(a_1, a_2, \ldots, a_n) \mapsto (a_\Sigma, a_\Sigma, \ldots, a_\Sigma).$$

This map $f_1$ is easily seen to be $S_n$-equivariant, and thus is a morphism of representations of $S_n$ (since it is $\mathbf{k}$-linear). Hence, by Proposition 4.3.18, this map $f_1$ is left $\mathbf{k}[S_n]$-linear.

However, the map $f_1$ is not a projection (unless $n = 1$ in $\mathbf{k}$), since it sends a vector $(a, a, \ldots, a) \in D(\mathbf{k}^n)$ to $(na, na, \ldots, na)$, which is generally not the same vector.

**(b)** Consider the $\mathbf{k}$-linear map

$$f_2 : \mathbf{k}^n \to D(\mathbf{k}^n), \\ (a_1, a_2, \ldots, a_n) \mapsto (a_1, a_1, \ldots, a_1).$$

This map $f_2$ is a projection (since it sends each vector $(a, a, \ldots, a) \in D(\mathbf{k}^n)$ to itself). But it is not $S_n$-equivariant, and thus not left $\mathbf{k}[S_n]$-linear.

**(c)** Let us now assume that $n$ is invertible in $\mathbf{k}$. Consider the $\mathbf{k}$-linear map

$$f_3 : \mathbf{k}^n \to D(\mathbf{k}^n), \\ (a_1, a_2, \ldots, a_n) \mapsto \left( \frac{a_\Sigma}{n}, \frac{a_\Sigma}{n}, \ldots, \frac{a_\Sigma}{n} \right).$$

(In other words, this map $f_3$ replaces all entries of an $n$-tuple by their average.)

It is easy to see that this map $f_3$ is both left $\mathbf{k}[S_n]$-linear and a projection. It is thus a left $\mathbf{k}[S_n]$-linear projection.

Had we not required $n$ to be invertible, such a map would not exist (unless $n = 0$). Thus, the existence of an $S_n$-equivariant map and the existence of a $\mathbf{k}$-linear projection do not guarantee the existence of an $S_n$-equivariant projection.

We note one last general property of direct addends before we resume studying modules over group algebras:

**Proposition 4.4.13.** Let $R$ be a ring. Let $V$ be a left $R$-module. Let $W$ be a direct addend of $V$. Then,

$$V \cong W \oplus (V/W) \qquad \text{as left } R\text{-modules.}$$

(Here, $W \oplus (V/W)$ means an external direct sum, since $V/W$ is not a submodule of $V$.)

*Proof.* Proceed as in the proof of the "$\Longrightarrow$" direction of Proposition 4.4.11. Thus, write $V = W \oplus U$ (internal direct sum) for some $R$-submodule $U$ of $V$, and observe that the map

$$\text{sum} : W \times U \to W + U, \\ (w, u) \mapsto w + u$$

is an $R$-module isomorphism. Thus, this map sum has an inverse map $\text{sum}^{-1} : W + U \to W \times U$, which is also an $R$-module isomorphism. Since $V = W \oplus$

$U = W + U$, we can rewrite this as follows: The map sum has an inverse map $\operatorname{sum}^{-1} : V \to W \times U$, which is also an $R$-module isomorphism. Hence, in particular, the map $\operatorname{sum}^{-1}$ is $R$-linear and bijective, hence surjective.

Furthermore, consider the $R$-linear map $\tau : W \times U \to U$ that sends each pair $(w, u)$ to its second entry $u$. This map $\tau$ is clearly surjective (since each $u \in U$ is the image of the pair $(0, u)$ under $\tau$).

Let $\omega : V \to U$ be the composition $\tau \circ \operatorname{sum}^{-1}$ of the two maps $\tau : W \times U \to U$ and $\operatorname{sum}^{-1} : V \to W \times U$. This composition $\omega$ is surjective (since both $\tau$ and $\operatorname{sum}^{-1}$ are surjective) and left $R$-linear (since both $\tau$ and $\operatorname{sum}^{-1}$ are $R$-linear).

Moreover, it is easy to see that the kernel $\operatorname{Ker}\omega$ of this composition is $W$. [*Proof:* Let $v \in V$. Let $(w, u) = \operatorname{sum}^{-1}(v)$. Thus, $(w, u) \in W \times U$ and $\operatorname{sum}(w, u) = v$, so that $\operatorname{sum}^{-1}(v) = (w, u)$. Now, from $\omega = \tau \circ \operatorname{sum}^{-1}$, we obtain

$$\omega(v) = \left(\tau \circ \operatorname{sum}^{-1}\right)(v) = \tau\left(\underbrace{\operatorname{sum}^{-1}(v)}_{=(w,u)}\right) = \tau(w, u)$$
$$= u \qquad \text{(by the definition of } \tau\text{)}.$$

Now, if $u = 0$, then $v = \operatorname{sum}(w, u) = w + \underbrace{u}_{=0} = w \in W$. Conversely, if $v \in W$, then $(v, 0) \in W \times U$ and therefore $\operatorname{sum}(v, 0)$ is well-defined and satisfies $\operatorname{sum}(v, 0) = v + 0 = v$, so that $(v, 0) = \operatorname{sum}^{-1}(v) = (w, u)$ and therefore $0 = u$, so that $u = 0$. Thus, each of the two statements $u = 0$ and $v \in W$ implies the other. In other words, we have the equivalence $(u = 0) \iff (v \in W)$. Now, we have the chain of equivalences

$$(v \in \operatorname{Ker}\omega) \iff \left(\underbrace{\omega(v)}_{=u} = 0\right) \iff (u = 0) \iff (v \in W).$$

Forget that we fixed $v$. We thus have shown that the equivalence $(v \in \operatorname{Ker}\omega) \iff (v \in W)$ holds for each $v \in V$. Therefore, $\operatorname{Ker}\omega = W$ (since both $\operatorname{Ker}\omega$ and $W$ are subsets of $V$).]

Now, the first isomorphism theorem (applied to the $R$-linear map $\omega : V \to U$) yields

$$V/\operatorname{Ker}\omega \cong \omega(V) = U \qquad \text{(since } \omega \text{ is surjective)}.$$

In view of $\operatorname{Ker}\omega = W$, we can rewrite this as

$$V/W \cong U.$$

Thus, $W \times (V/W) \cong W \times U$ (because of a general fact that if $X$, $Y$ and $Z$ are three $R$-modules satisfying $Y \cong Z$, then $X \times Y \cong X \times Z$). However, we also have $V \cong W \times U$ (since $\operatorname{sum}^{-1} : V \to W \times U$ is an $R$-module isomorphism). Altogether, we thus find $V \cong W \times U \cong W \times (V/W) = W \oplus (V/W)$ (since a direct product of finitely many $R$-modules is the same as their direct sum). This proves Proposition 4.4.13. $\square$

### 4.4.2. Maschke's theorem

Recall that if $G$ is a group, then the left $\mathbf{k}[G]$-modules are just the representations of $G$ over $\mathbf{k}$. Thus, we obtain a notion of a direct addend for representations of $G$.

So when is a subrepresentation of a representation $V$ of $G$ a direct addend? i.e., when is there a $\mathbf{k}[G]$-linear projection from $V$ to the subrepresentation? The following theorem (known as *Maschke's theorem*) yields a surprisingly general answer to this question, which applies not quite always but in many situations:

**Theorem 4.4.14** (Maschke's theorem)**.** Let $G$ be a finite group. Let $V$ be a left $\mathbf{k}[G]$-module (i.e., a representation of $G$ over $\mathbf{k}$). Let $W$ be a left $\mathbf{k}[G]$-submodule of $V$ (that is, a subrepresentation of $V$). Assume that

1. the number $|G|$ is invertible in $\mathbf{k}$;
2. the $\mathbf{k}$-submodule $W$ is a direct addend of $V$ as a $\mathbf{k}$-module.

Then, the left $\mathbf{k}[G]$-submodule $W$ is also a direct addend of $V$ as a left $\mathbf{k}[G]$-module.

**Remark 4.4.15.** If $\mathbf{k}$ is a field of characteristic 0, or (more generally) a commutative $\mathbb{Q}$-algebra, then the first assumption of Theorem 4.4.14 is automatically satisfied.

If $\mathbf{k}$ is a field, then the second assumption of Theorem 4.4.14 is automatically satisfied (by Proposition 4.4.9).

**Example 4.4.16.** Recall the natural representation $\mathbf{k}^n$ of $S_n$ and its subrepresentation $D(\mathbf{k}^n)$. Then, $D(\mathbf{k}^n)$ is always a direct addend of $\mathbf{k}^n$ as a $\mathbf{k}$-module (e.g., with complement $\operatorname{span}_{\mathbf{k}}\{e_1, e_2, \ldots, e_{n-1}\}$), but not always a direct addend of $\mathbf{k}^n$ as a left $\mathbf{k}[S_n]$-module. Maschke's theorem (Theorem 4.4.14) guarantees that $D(\mathbf{k}^n)$ is a direct addend of $\mathbf{k}^n$ as a left $\mathbf{k}[S_n]$-module whenever $|S_n| = n!$ is invertible in $\mathbf{k}$, but we already know from Proposition 4.2.28 **(a)** that this is true under a weaker assumption (namely, whenever $n$ is invertible in $\mathbf{k}$). Thus, Maschke's theorem does not tell us anything new here. But due to its generality, Maschke's theorem can be used for many other representations, which are much harder to analyze "by hand" than $\mathbf{k}^n$.

*Proof of Theorem 4.4.14.* We have assumed that $W$ is a direct addend of the $\mathbf{k}$-module $V$. Hence, there exists a $\mathbf{k}$-linear projection $\pi : V \to W$ (by Proposition 4.4.11, applied to $R = \mathbf{k}$). Consider this $\pi$. Note that $\pi$ is only $\mathbf{k}$-linear, but usually not left $\mathbf{k}[G]$-linear.

Define a new map $\pi' : V \to W$ by setting

$$\pi'(v) := \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(gv) \qquad \text{for each } v \in V. \tag{113}$$

We claim that this is a left $\mathbf{k}[G]$-linear projection from $V$ to $W$. Indeed, let us prove the $\mathbf{k}[G]$-linearity and the "projection-ness" separately:

- *Proof of $\mathbf{k}[G]$-linearity:* It is clear that the map $\pi'$ is $\mathbf{k}$-linear (since $\pi$ is $\mathbf{k}$-linear, whereas the actions of each $g \in G$ and each $g^{-1}$ are $\mathbf{k}$-linear). Now, we shall show that $\pi'$ is $G$-equivariant, i.e., that we have $\pi'(hv) = h\pi'(v)$ for all $h \in G$ and $v \in V$.

  Indeed, let $h \in G$ and $v \in V$. Recall that $G$ is a group. Thus, the map $G \to G,\ g \mapsto gh$ is a bijection (with inverse $g \mapsto gh^{-1}$). The definition of $\pi'$ yields

  $$\pi'(hv) = \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(ghv).$$

  Comparing this with

  $$\begin{aligned} h\pi'(v) &= h \cdot \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(gv) \qquad \text{(by (113))} \\ &= \frac{1}{|G|} \sum_{g \in G} hg^{-1} \pi(gv) = \frac{1}{|G|} \sum_{g \in G} \underbrace{h(gh)^{-1}}_{=hh^{-1}g^{-1}=g^{-1}} \pi(ghv) \\ &\qquad \left( \begin{array}{c} \text{here, we have substituted } gh \text{ for } g \text{ in the sum,} \\ \text{since the map } G \to G,\ g \mapsto gh \text{ is a bijection} \end{array} \right) \\ &= \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(ghv), \end{aligned}$$

  we find $\pi'(hv) = h\pi'(v)$.

  Forget that we fixed $h$ and $v$. We thus have shown that $\pi'(hv) = h\pi'(v)$ for all $h \in G$ and $v \in V$. In other words, the map $\pi'$ is $G$-equivariant. Hence, $\pi'$ is a morphism of $G$-representations (since $\pi'$ is $\mathbf{k}$-linear). By Proposition 4.3.18, this entails that $\pi'$ is a $\mathbf{k}[G]$-module morphism, i.e., a left $\mathbf{k}[G]$-linear map.

- *Proof of "projection-ness":* We must now show that $\pi' : V \to W$ is a projection. In other words, we must show that $\pi'(w) = w$ for each $w \in W$.

  Let $w \in W$. Then, each $g \in G$ satisfies $gw \in W$ (since $W$ is a subrepresentation of $V$ and thus a left $\mathbf{k}[G]$-submodule of $V$) and thus

  $$\pi(gw) = (\pi|_W)(gw) = gw \tag{114}$$

  (since $\pi : V \to W$ is a projection and thus satisfies $\pi|_W = \mathrm{id}$). Now, the

definition of $\pi'$ yields

$$\pi'(w) = \frac{1}{|G|} \sum_{g \in G} g^{-1} \underbrace{\pi(gw)}_{\substack{=gw \\ \text{(by (114))}}} = \frac{1}{|G|} \sum_{g \in G} \underbrace{g^{-1} g}_{=1} w = \frac{1}{|G|} \underbrace{\sum_{g \in G} w}_{=|G| w}$$

$$= \frac{1}{|G|} |G| w = w.$$

Forget that we fixed $w$. We thus have shown that $\pi'(w) = w$ for each $w \in W$. Hence, the map $\pi' : V \to W$ is a projection.

Altogether, we have now shown that $\pi'$ is a left $\mathbf{k}[G]$-linear projection from $V$ to $W$. Hence, such a projection exists. It thus follows that $W$ is a direct addend of $V$ as a left $\mathbf{k}[G]$-module (by Proposition 4.4.11, applied to $R = \mathbf{k}[G]$). This proves Theorem 4.4.14. □

**Remark 4.4.17.** What is the motivation behind the definition of the map $\pi'$ in the above proof?

In order to apply Proposition 4.4.11, we need a left $\mathbf{k}[G]$-linear (that is, $\mathbf{k}$-linear and $G$-equivariant) projection from $V$ to $W$. We are given a $\mathbf{k}$-linear projection $\pi : V \to W$, but it is (usually) not $G$-equivariant. This lack of $G$-equivariance can be used to define a family of "doppelgängers" of $\pi$: Namely, for each $g \in G$, we can consider the "$g$-tilted version" $\pi_g : V \to W$ of $\pi$ given by

$$\pi_g(v) := g^{-1} \pi(gv) \qquad \text{for each } v \in V.$$

It is easy to see that each such $\pi_g$ is again a $\mathbf{k}$-linear projection $V \to W$ (indeed, $\pi_g$ is the conjugate of $\pi$ by the action of $g^{-1}$, in the appropriate sense).

Thus, we now have not one but $|G|$ many $\mathbf{k}$-linear projections from $V$ to $W$. None of them is (usually) $G$-equivariant, but something similar holds: For any $g \in G$, $h \in G$ and $v \in V$, we have

$$\pi_g(hv) = h \pi_{gh}(v). \tag{115}$$

If the maps $\pi_g$ and $\pi_{gh}$ were the same, then this would show that this map is $G$-equivariant. But they are different. How do we salvage them? Simply: We take the average of the $\pi_g$ over all group elements $g \in G$. Upon averaging over all $g \in G$, the $\pi_g$ and the $\pi_{gh}$ in the equality (115) become the same, since $\sum_{g \in G} \pi_g = \sum_{g \in G} \pi_{gh}$. Thus, the average of the $\pi_g$ over all $g \in G$ is $G$-equivariant. Moreover, this average is still a projection (since any average of $\mathbf{k}$-linear projections from $V$ to $W$ is again a projection from $V$ to $W$). This average is precisely the map $\pi'$ in the above proof.

Maschke's theorem can be generalized from the "single group $G$" setting to the "group $G$ and subgroup $H$" setting:

**Exercise 4.4.2.** [3] Let $G$ be a group, and $H$ a subgroup of $G$. Assume that the index $[G : H]$ (which is defined as the size of the set $G/H$) is finite. Let $V$ be a left $\mathbf{k}[G]$-module (i.e., a representation of $G$ over $\mathbf{k}$). Let $W$ be a left $\mathbf{k}[G]$-submodule of $V$ (that is, a subrepresentation of $V$). Note that every left $\mathbf{k}[G]$-module automatically becomes a left $\mathbf{k}[H]$-module (by restriction of scalars, since the inclusion map $H \to G$ gives rise to a $\mathbf{k}$-algebra morphism $\mathbf{k}[H] \to \mathbf{k}[G]$). Assume that

1. the number $[G : H]$ is invertible in $\mathbf{k}$;
2. the $\mathbf{k}[H]$-submodule $W$ is a direct addend of $V$ as a $\mathbf{k}[H]$-module.

Prove that the left $\mathbf{k}[G]$-submodule $W$ is also a direct addend of $V$ as a left $\mathbf{k}[G]$-module.

[**Hint:** Cosets.]

Maschke's theorem also has the following consequence:

**Corollary 4.4.18.** Let $G$, $V$ and $W$ be as in Theorem 4.4.14. Then,

$$V \cong W \oplus (V/W) \qquad \text{as left } \mathbf{k}[G]\text{-modules.}$$

(Here, $W \oplus (V/W)$ is of course an external direct sum, since $V/W$ is not a subset of $V$.)

*Proof.* Theorem 4.4.14 yields that the left $\mathbf{k}[G]$-submodule $W$ is a direct addend of $V$ as a left $\mathbf{k}[G]$-module. Hence, Proposition 4.4.13 (applied to $R = \mathbf{k}[G]$) yields that

$$V \cong W \oplus (V/W) \qquad \text{as left } \mathbf{k}[G]\text{-modules.}$$

This proves Corollary 4.4.18. □

**Exercise 4.4.3.** Let $n$ be a positive integer. Let $G = C_n$ (the cyclic group with $n$ elements), and let $\mathbf{k} = \mathbb{Z}/n$. Let $V$ be the representation $\mathbf{k}^2$ of $G$, where the generator $g \in C_n$ acts as the $\mathbf{k}$-linear map

$$w : \mathbf{k}^2 \to \mathbf{k}^2,$$
$$(x, y) \mapsto (x,\ x + y).$$

**(a)** [1] Prove that such a representation really exists (i.e., that the above $\mathbf{k}$-linear map $w$ satisfies $w^n = \mathrm{id}$, so that Example 4.2.9 constructs a representation of $C_n$ from it).

**(b)** $\boxed{2}$ Assume that $n$ is prime (so that $\mathbf{k} = \mathbb{Z}/n$ is a finite field). Prove that $\operatorname{span}_{\mathbf{k}} \{(0,1)\}$ is a subrepresentation of $V$, but is not a direct addend. More generally, prove that $V$ cannot be written as a direct sum of two nonzero subrepresentations. (Unsurprisingly, Theorem 4.4.14 does not apply here, since $|G| = |C_n| = n$ is not invertible in $\mathbf{k}$.)

### 4.4.3. Filtrations

Throughout mathematics, a nested chain of subobjects of a given object is called a *filtration* of said object if it starts with the "trivial" subobject and ends with the original object itself. Let us be more specific and define this notion for left $R$-modules in particular:

**Definition 4.4.19.** Let $R$ be a ring. Let $V$ be a left $R$-module.

**(a)** A *filtration* (aka *complete filtration*) of $V$ means a tuple $(W_0, W_1, \ldots, W_k)$ of left $R$-submodules of $V$ satisfying

$$0 = W_0 \subseteq W_1 \subseteq W_2 \subseteq \cdots \subseteq W_k = V,$$

where the 0 at the front means the $R$-submodule $\{0\}$. Note that we allow the "$\subseteq$" signs to be equalities.

When the ring $R$ might not be clear from the context, we can disambiguate it by saying "*left $R$-module filtration*" or "*filtration by left $R$-submodules*" instead of just "filtration".

**(b)** If we drop the $W_k = V$ requirement in Definition 4.4.19 **(a)**, then $(W_0, W_1, \ldots, W_k)$ is called an *incomplete filtration* of $V$.

**(c)** If $(W_0, W_1, \ldots, W_k)$ is a filtration of $V$, then the quotient $R$-modules $W_i/W_{i-1}$ for all $i \in [k]$ are called the *subquotients* of this filtration.

**(d)** If $(W_0, W_1, \ldots, W_k)$ is a filtration of $V$, then the number $k$ is called the *length* of this filtration.

Some other variants of filtrations can be found in the literature: e.g., decreasing filtrations (where the "$\subseteq$" signs are replaced by "$\supseteq$" signs), infinite filtrations (where the tuple is replaced by an infinite sequence), poset filtrations (where the tuple is replaced by a partially ordered family) and others.

Intuitively, a filtration of $V$ can be regarded as a "staircase" going from 0 to $V$, with each stair being a left $R$-submodule of $V$. Its subquotients are the "steps" from a stair to the next.

**Example 4.4.20.** Proposition 4.2.28 **(b)** shows that if $n = 0$ in $\mathbf{k}$, then $(0,\ D(\mathbf{k}^n),\ R(\mathbf{k}^n),\ \mathbf{k}^n)$ is a filtration of the left $\mathbf{k}[S_n]$-module $\mathbf{k}^n$.

**Example 4.4.21.** Let $R$ be a ring. If $A$ and $B$ are two $R$-submodules of a left $R$-module $V$, then $(0,\ A \cap B,\ A,\ A+B,\ V)$ is a filtration of $V$.

Maschke's theorem (Corollary 4.4.18 specifically) yields:

**Corollary 4.4.22.** Let $\mathbf{k}$ be a field of characteristic 0. Let $G$ be a finite group. Let $V$ be a left $\mathbf{k}[G]$-module. If $(W_0, W_1, \ldots, W_k)$ is a filtration of $V$, then

$$V \cong (W_1/W_0) \oplus (W_2/W_1) \oplus \cdots \oplus (W_k/W_{k-1}).$$

*Proof of Corollary 4.4.22 (sketched).* Induct on $k$.

The *base case* $k = 0$ is obvious (here, $V$ must be $\{0\}$, since the definition of a filtration yields $0 = W_0 = W_k = V$).

For the *induction step* (from $k-1$ to $k$), we consider a filtration $(W_0, W_1, \ldots, W_k)$ of $V$. Then, $0 = W_0 \subseteq W_1 \subseteq W_2 \subseteq \cdots \subseteq W_{k-1} \subseteq W_k = V$. Now, $W_{k-1}$ is a left $\mathbf{k}[G]$-submodule of $V$, and thus we can apply Corollary 4.4.18 to $W = W_{k-1}$ (indeed, both assumptions of Theorem 4.4.14 are satisfied by Remark 4.4.15). Thus, we obtain

$$V \cong W_{k-1} \oplus (V/W_{k-1}) = W_{k-1} \oplus (W_k/W_{k-1}) \tag{116}$$

(since $V = W_k$). But the tuple $(W_0, W_1, \ldots, W_{k-1})$ is a filtration of $W_{k-1}$ (since $0 = W_0 \subseteq W_1 \subseteq \cdots \subseteq W_{k-1} = W_{k-1}$), and thus our induction hypothesis (applied to $W_{k-1}$ and $(W_0, W_1, \ldots, W_{k-1})$ instead of $V$ and $(W_0, W_1, \ldots, W_k)$) yields that

$$W_{k-1} \cong (W_1/W_0) \oplus (W_2/W_1) \oplus \cdots \oplus (W_{k-1}/W_{k-2}).$$

Thus, we can rewrite (116) as

$$\begin{aligned} V &\cong ((W_1/W_0) \oplus (W_2/W_1) \oplus \cdots \oplus (W_{k-1}/W_{k-2})) \oplus (W_k/W_{k-1}) \\ &\cong (W_1/W_0) \oplus (W_2/W_1) \oplus \cdots \oplus (W_k/W_{k-1}). \end{aligned}$$

This proves Corollary 4.4.22 for our $k$, and thus completes the induction step. □

Recall that representations of a group $G$ over $\mathbf{k}$ are just left $\mathbf{k}[G]$-modules. Thus, the concept of a filtration can be transferred automatically from left $\mathbf{k}[G]$-modules to representations of $G$.

If $\mathbf{k}$ is a field of characteristic 0, and if $G$ is a finite group, then Corollary 4.4.22 shows that any filtration of a representation of $G$ can be used to decompose this representation into a direct sum. But in general (when $\mathbf{k}$ is not a field of characteristic 0), filtrations are more frequent (and more easily found) than direct sum decompositions.

What is the longest filtration that we can find for a given left $R$-module $V$ ? Taken literally, this question is stupid (we can take an arbitrarily long tuple

of the form $(0, V, V, \ldots, V)$), but we can make it more sensible by restricting ourselves to "nondegenerate" filtrations $(W_0, W_1, \ldots, W_k)$ (that is, ones in which each $W_{i-1}$ is a **proper** subset of $W_i$). Even then, the question does not always have an answer, since (e.g.) the $\mathbb{Z}$-module $\mathbb{Z}$ has arbitrarily long nondegenerate filtrations:

$$\left(0\mathbb{Z},\ 2^k\mathbb{Z},\ 2^{k-1}\mathbb{Z},\ \ldots,\ 2^0\mathbb{Z}\right) \text{ is a filtration of } \mathbb{Z} \text{ for each } k \in \mathbb{N}.$$

However, if $\mathbf{k}$ is a field and $R$ is a $\mathbf{k}$-algebra, and if $V$ is a left $R$-module that is finite-dimensional as a $\mathbf{k}$-vector space, then a nondegenerate filtration of $V$ will always have length $\leq \dim V$, where $\dim V$ is the dimension of $V$ as a $\mathbf{k}$-vector space (why?). Thus, longest nondegenerate filtrations exist in this case. Such filtrations are known as *composition series* of $V$, and can be equivalently described as those filtrations whose subquotients are irreducible $R$-modules (why?). A classical result known as the *Jordan–Hölder theorem* (see, e.g., [EGHetc11, Theorem 3.7.1]) says that the subquotients of such filtrations are uniquely determined (up to reordering and isomorphism):

**Theorem 4.4.23** (Jordan–Hölder theorem)**.** Let $\mathbf{k}$ be a field. Let $R$ be a $\mathbf{k}$-algebra. Let $V$ be a left $R$-module that is finite-dimensional as a $\mathbf{k}$-vector space. Then:

**(a)** There exists a filtration $(W_0, W_1, \ldots, W_k)$ of $V$ such that all its subquotients $W_i/W_{i-1}$ are irreducible left $R$-modules. Such a filtration is called a *composition series* (or *Jordan–Hölder series*) of $V$.

**(b)** Any two composition series of $V$ have the same subquotients (up to reordering and isomorphism).

*Proof idea.* **(a)** As we said above, any nondegenerate filtration of $V$ must have length $\leq \dim V$, so that there exists a longest nondegenerate filtration of $V$, and any such longest filtration must be a composition series.

Alternatively, we can construct a composition series by hand: Start by setting $W_0 = 0$; then let $W_1$ be a minimum-dimensional $R$-submodule of $V$ that properly contains $W_0$; then let $W_2$ be a minimum-dimensional $R$-submodule of $V$ that properly contains $W_1$; and so on. The sequence must come to an end, since $\dim W_0 < \dim W_1 < \dim W_2 < \cdots \leq \dim V$.

**(b)** This is not that easy. See [EGHetc11, Theorem 3.7.1] or any text on representation theory. (There is also an analogous "Jordan–Hölder theorem" in group theory – see [Conrad, "Subgroup Series I", Theorem 2.8] – whose proof is rather analogous to the proof of Theorem 4.4.23 **(b)**.) □

Composition series can be useful in some obvious and some less obvious ways:

- Their most obvious use is to help break an $R$-module $V$ into irreducible $R$-modules (the subquotients of a composition series). This works particularly well when $R$ is a group algebra over a field of characteristic 0, since in this case $V$ is really decomposed into a direct sum (by Corollary 4.4.22). In other cases, the "pieces" might not determine $V$ completely, but they still encode a lot of information about $V$.

- A less obvious use: They impose a "block-triangularity" on the action of any element of $\mathbf{k}[G]$. Indeed, if $R$ is a $\mathbf{k}$-algebra, and if $(W_0, W_1, \ldots, W_k)$ is a filtration of a left $R$-module $V$, then the action $\tau_r$ of an element $r \in R$ on $V$ is not just a $\mathbf{k}$-module endomorphism of $V$, but a $\mathbf{k}$-module endomorphism that preserves each of the submodules $W_0, W_1, \ldots, W_k$. Translated into the language of matrices (where we assume that $\mathbf{k}$ is a field and $V$ is a finite-dimensional $\mathbf{k}$-vector space), such an endomorphism becomes a block-upper-triangular matrix in an appropriate basis of $V$. We might say more about this later.

- Even less obviously: They can be used to find all the irreps (i.e., irreducible representations) of a finite group $G$, or more generally, all irreducible modules over a $\mathbf{k}$-algebra $R$. This relies on the following theorem:

**Theorem 4.4.24.** Let $\mathbf{k}$ be a field. Let $R$ be a $\mathbf{k}$-algebra that is finite-dimensional as a $\mathbf{k}$-vector space. Let $(W_0, W_1, \ldots, W_k)$ be a composition series of the left regular $R$-module $R$. Then, each irreducible left $R$-module is isomorphic to $W_i/W_{i-1}$ for some $i \in [k]$. (Note that this $i$ does not have to be unique.) In other words, each irreducible left $R$-module appears (up to isomorphism) as a subquotient of our composition series.

*Proof of Theorem 4.4.24 (sketched).* Let $J$ be an irreducible left $R$-module. We must show that $J \cong W_i/W_{i-1}$ for some $i \in [k]$.

The $R$-module $J$ is nonzero (since $J$ is irreducible). Hence, there exists a nonzero vector $v \in J$. Consider this $v$.

For each left $R$-submodule $M$ of $R$, we let

$$Mv := \{mv \mid m \in M\}.$$

This $Mv$ is a left $R$-submodule of $J$, since $M$ is a left $R$-submodule of $R$ (thus fixed under addition and left scaling by $R$).

The tuple $(W_0, W_1, \ldots, W_k)$ is a composition series and thus a filtration of the left $R$-module $R$. Hence, $0 = W_0 \subseteq W_1 \subseteq W_2 \subseteq \cdots \subseteq W_k = R$. In particular, $W_k = R$. But $1 \in R$ and thus $1v \in Rv$, so that $v = 1v \in Rv = W_k v$ (since $R = W_k$). Hence, $W_k v \neq 0$ (since $v$ is nonzero). On the other hand, from $W_0 = 0$, we obtain $W_0 v = 0v = 0$.

We have $W_k v \neq 0$. Hence, there exists some $i \in \{0, 1, \ldots, k\}$ such that $W_i v \neq 0$ (namely, $i = k$). Pick the **smallest** such $i$. Thus, $i \neq 0$ (since $W_0 v = 0$), so that $i - 1 \in \{0, 1, \ldots, k\}$ as well. Moreover, $W_{i-1} v = 0$ (since $i$ is the **smallest** number to satisfy $W_i v \neq 0$).

Consider the map

$$f : W_i/W_{i-1} \to J,$$
$$\overline{r} \mapsto rv,$$

where $\overline{r}$ denotes the residue class of a vector $r \in W_i$ modulo $W_{i-1}$. This map $f$ is well-defined, since two vectors $r, s \in W_i$ with equal residue classes $\overline{r} = \overline{s} \in W_i/W_{i-1}$ will necessarily satisfy $rv = sv$ (because $\overline{r} = \overline{s}$ yields $r - s \in W_{i-1}$ and thus $(r-s)\,v \in W_{i-1}v = 0$, so that $(r-s)\,v = 0$, so that $0 = (r-s)\,v = rv - sv$ and therefore $rv = sv$). Moreover, the map $f$ is left $R$-linear (by a trivial computation).

Therefore, the image $f(W_i/W_{i-1})$ of $f$ is a left $R$-submodule of $J$. Since $J$ is irreducible, this image must therefore be either $J$ or 0. Since this image cannot be 0 (because the definition of $f$ yields that $f(W_i/W_{i-1}) = \{rv \mid r \in W_i\} = W_i v \neq 0$), we thus conclude that it is $J$. In other words, the map $f$ is surjective.

The kernel $\operatorname{Ker} f$ of $f$ is a left $R$-submodule of $W_i/W_{i-1}$ (since $f$ is left $R$-linear). But the left $R$-module $W_i/W_{i-1}$ is irreducible (since $(W_0, W_1, \ldots, W_k)$ is a composition series). Hence, its only $R$-submodules are $W_i/W_{i-1}$ and 0. Consequently, $\operatorname{Ker} f$ (being its left $R$-submodule) must be either $W_i/W_{i-1}$ or 0. Since $\operatorname{Ker} f$ cannot be $W_i/W_{i-1}$ (because this would entail $f = 0$, and thus the image of $f$ would be 0, but we have already shown that this image cannot be 0), we thus conclude that $\operatorname{Ker} f$ must be 0. In other words, the map $f$ is injective (since $f$ is $R$-linear).

Now we know that the map $f$ is injective and surjective, hence bijective. Since $f$ is furthermore left $R$-linear, we thus conclude that $f$ is an $R$-module isomorphism. Therefore, $J \cong W_i/W_{i-1}$ as left $R$-modules.

Thus we have found an $i \in [k]$ such that $J \cong W_i/W_{i-1}$. Hence, such an $i$ exists. This proves Theorem 4.4.24. □

Since we are talking about filtrations of the left regular $R$-module $R$, it is worth mentioning that the left $R$-submodules of this $R$-module are better known under a different name:

**Remark 4.4.25.** Let $R$ be a ring. A left $R$-submodule of the left regular $R$-module $R$ is the same thing as a left ideal of $R$, that is, a subgroup $J$ of the additive group $(R, +, 0)$ with the property that

$$\text{all } r \in R \text{ and } j \in J \text{ satisfy } rj \in J$$

(but not necessarily $jr \in J$).

*Proof.* Just compare the definitions of "left ideal" and "left $R$-submodule". (Note that existence of negative inverses follows from scaling by $-1_R$.) □

We note the following simple consequence of Theorem 4.4.24:

**Corollary 4.4.26.** Let $\mathbf{k}$ be a field. Let $R$ be a $\mathbf{k}$-algebra that is finite-dimensional as a $\mathbf{k}$-vector space. Then, there are only finitely many irreducible left $R$-modules (up to isomorphism).

In particular, if $\mathbf{k}$ is a field and $G$ is a finite group, then there are (up to isomorphism) only finitely many irreducible left $\mathbf{k}[G]$-modules, i.e., only finitely many irreps[44] of $G$. If we want to find these irreps, then (by Theorem 4.4.24) we only need to look at a composition series of its left regular representation $\mathbf{k}[G]$. The subquotients of this series will then be the irreps (though some might appear multiple times). If $\mathbf{k}$ is a field of characteristic 0, then we furthermore conclude (by Corollary 4.4.22) that the left regular representation $\mathbf{k}[G]$ can be decomposed as a direct sum of these irreps (with some irreps appearing multiple times as addends), and therefore we see that every irrep of $G$ is isomorphic to a subrepresentation of $\mathbf{k}[G]$. The latter part of this claim is true even when $\mathbf{k}$ does not have characteristic 0:

**Theorem 4.4.27.** Let $\mathbf{k}$ be any field, and let $G$ be a finite group. Then, any irreducible representation of $G$ is isomorphic to a subrepresentation of the left regular representation $\mathbf{k}[G]$.

*Proof.* See `https://mathoverflow.net/questions/72844` or Exercise 4.4.4 **(c)**. □

**Exercise 4.4.4.** Let $\mathbf{k}$ be any field, and let $G$ be a finite group. Let $V$ be an irreducible representation of $G$. Prove the following:

**(a)** [2] If $v \in V$ is a nonzero vector, then $V = \operatorname{span}_{\mathbf{k}} \{gv \mid g \in G\}$.

**(b)** [2] If $f : V \to \mathbf{k}$ is a nonzero $\mathbf{k}$-linear map, and if we define a map $\alpha : V \to \mathbf{k}[G]$ by setting

$$\alpha(v) = \sum_{g \in G} f(gv)\, g^{-1} \qquad \text{for all } v \in V,$$

then $\alpha$ is an injective morphism of $G$-representations (from $V$ to the left regular representation of $G$).

**(c)** [1] The representation $V$ is isomorphic to a subrepresentation of the left regular representation $\mathbf{k}[G]$.

**Exercise 4.4.5.** [2] Prove the following variant of Theorem 4.4.24:

Let $\mathbf{k}$ be a field. Let $R$ be a $\mathbf{k}$-algebra. Let $V$ be a left $R$-module that is finite-dimensional as a $\mathbf{k}$-vector space. Let $(W_0, W_1, \ldots, W_k)$ be a composition series of the left $R$-module $V$. Then, each irreducible left $R$-submodule of $V$ is isomorphic to $W_i / W_{i-1}$ for some $i \in [k]$.

### 4.4.4. An example: $\mathbf{k}[S_3]$

Let us now get our hands dirty and find a composition series in the wild.

[44] Recall that "irrep" is short for "irreducible representation".

We let $\mathbf{k}$ be a field of characteristic $\neq 3$ (from here on until the end of Subsection 4.4.4). We want to find a composition series of the left regular representation of $S_3$.

The left regular representation of $S_3$ is the left regular $\mathbf{k}[S_3]$-module $\mathbf{k}[S_3]$. As a $\mathbf{k}$-vector space, it is 6-dimensional, with basis $(1, s_1, s_2, s_1s_2, s_2s_1, s_1s_2s_1)$ (or any other way to list the six permutations in $S_3$). We define the following left $\mathbf{k}[S_3]$-submodules of the left regular representation $\mathbf{k}[S_3]$:

$$\langle\nabla\rangle := \operatorname{span}_{\mathbf{k}}\{\nabla\} \qquad \left(\begin{array}{c}\text{this is a left } \mathbf{k}[S_3]\text{-submodule,}\\ \text{by Exercise 3.12.1}\end{array}\right);$$

$$A_1 := \{\mathbf{a} \in \mathbf{k}[S_3] \mid \mathbf{a}s_1 = \mathbf{a}\} \qquad \left(\begin{array}{c}\text{this is a left } \mathbf{k}[S_3]\text{-submodule,}\\ \text{since } \mathbf{a}s_1 = \mathbf{a} \text{ implies } \mathbf{b}\mathbf{a}s_1 = \mathbf{b}\mathbf{a}\\ \text{for any } \mathbf{b} \in \mathbf{k}[S_3]\end{array}\right);$$

$$A_2 := \{\mathbf{a} \in \mathbf{k}[S_3] \mid \mathbf{a}s_2 = \mathbf{a}\} \qquad \left(\begin{array}{c}\text{this is a left } \mathbf{k}[S_3]\text{-submodule,}\\ \text{for similar reasons as for } A_1\end{array}\right).$$

It is easy to see that $A_1 \cap A_2 = \langle\nabla\rangle$ (in fact, this is just Exercise 3.3.1 **(b)** for $n = 3$). Thus, the tuple

$$(0,\ \langle\nabla\rangle,\ A_1,\ A_1 + A_2,\ \mathbf{k}[S_3]) \tag{117}$$

(where 0 denotes $\{0\}$ as usual in linear algebra) is a filtration of the left $\mathbf{k}[S_3]$-module $\mathbf{k}[S_3]$. Now we shall show that it is a composition series of $\mathbf{k}[S_3]$.

The easiest way to prove this is to find bases of all the submodules and the subquotients. This is best achieved by finding a basis of $\mathbf{k}[S_3]$ that is "adapted" to this filtration, meaning that each of the submodules $0$, $\langle\nabla\rangle$, $A_1$, $A_1 + A_2$, $\mathbf{k}[S_3]$ in the filtration is spanned by some part of this basis. Indeed, such a basis will automatically provide bases for all subquotients, because of the following simple fact from linear algebra:

**Lemma 4.4.28.** Let $\mathbf{k}$ be any commutative ring. Let $V$ be a $\mathbf{k}$-module, and let $W$ be a $\mathbf{k}$-submodule of $V$. For each $v \in V$, we let $\overline{v}$ denote the residue class of $v$ in $V/W$ (that is, the coset $v + W$).

Assume that we have a basis $(b_1, b_2, \ldots, b_k)$ of $V$ and a number $i \in \{0, 1, \ldots, k\}$ such that the $i$-tuple $(b_1, b_2, \ldots, b_i)$ is a basis of $W$. Then, the $(k-i)$-tuple $\left(\overline{b_{i+1}},\ \overline{b_{i+2}},\ \ldots,\ \overline{b_k}\right)$ is a basis of the quotient $\mathbf{k}$-module $V/W$.

Of course, a basis $(b_1, b_2, \ldots, b_k)$ and a number $i$ as in Lemma 4.4.28 do not always exist for general $\mathbf{k}$ and $V$. But we are dealing with a rather nice situation, in which $\mathbf{k}$ is a field and $V$ is a finite-dimensional vector space. In this situation, $(b_1, b_2, \ldots, b_k)$ and $i$ always exist, since we can choose any basis $(b_1, b_2, \ldots, b_i)$ of $W$ and extend it to a basis $(b_1, b_2, \ldots, b_k)$ of $V$.

This is also how we can find a basis of $\mathbf{k}[S_3]$ that is "adapted" to our filtration (117): We start with the (empty) basis of the zero subspace 0; then we extend it

to a basis of $\langle \nabla \rangle$; then we extend the latter to a basis of $A_1$; and so on. There are many possible options for these steps, but let me choose a particularly nice one.

To define this basis, recall the elements

$$\nabla_{b,a} := \sum_{\substack{w \in S_n; \\ w(a)=b}} w \qquad \text{for all } a, b \in [n]$$

of $\mathbf{k}[S_n]$ (these were defined in Exercise 3.16.1). Now, I claim that the six vectors

$$\begin{aligned}
\nabla &= 1 + s_1 + s_2 + s_1 s_2 + s_2 s_1 + s_1 s_2 s_1, \\
\nabla_{3,3} &= 1 + s_1, \\
\nabla_{2,3} &= s_2 (1 + s_1) = s_2 + s_2 s_1, \\
\nabla_{3,2} &= (1 + s_1) s_2 = s_2 + s_1 s_2, \\
\nabla_{2,2} &= 1 + t_{1,3} = s_2 (1 + s_1) s_2 = 1 + s_2 s_1 s_2 = 1 + s_1 s_2 s_1, \\
1 &= 1
\end{aligned}$$

form a basis of $\mathbf{k}[S_3]$ that is "adapted" to our filtration (117). Indeed, it is not hard to see that

$$0 = \operatorname{span}_{\mathbf{k}} \{\}; \tag{118}$$

$$\langle \nabla \rangle = \operatorname{span}_{\mathbf{k}} \{\nabla\}; \tag{119}$$

$$A_1 = \operatorname{span}_{\mathbf{k}} \{\nabla,\ \nabla_{3,3},\ \nabla_{2,3}\}; \tag{120}$$

$$A_1 + A_2 = \operatorname{span}_{\mathbf{k}} \{\nabla,\ \nabla_{3,3},\ \nabla_{2,3},\ \nabla_{3,2},\ \nabla_{2,2}\}, \tag{121}$$

$$\mathbf{k}[S_3] = \operatorname{span}_{\mathbf{k}} \{\nabla,\ \nabla_{3,3},\ \nabla_{2,3},\ \nabla_{3,2},\ \nabla_{2,2},\ 1\}. \tag{122}$$

(The first and two of these equalities here are obvious. The third is easiest to prove in the form $A_1 = \operatorname{span}_{\mathbf{k}} \{\nabla_{3,3},\ \nabla_{2,3},\ \nabla_{1,3}\}$, which is equivalent because $\nabla = \nabla_{3,3} + \nabla_{2,3} + \nabla_{1,3}$. This can be shown by writing an arbitrary element $\mathbf{a} \in \mathbf{k}[S_3]$ as $\mathbf{a} = \alpha + \beta s_1 + \gamma s_2 + \delta t_{1,3} + \varphi s_1 s_2 + \zeta s_2 s_1$ for $\alpha, \beta, \gamma, \delta, \varphi, \zeta \in \mathbf{k}$, and observing that the equality $\mathbf{a} s_1 = \mathbf{a}$ is equivalent to ($\alpha = \beta$ and $\gamma = \zeta$ and $\delta = \varphi$), which allows us to write $\mathbf{a}$ as $\mathbf{a} = \alpha (1 + s_1) + \gamma s_2 (1 + s_1) + \delta t_{1,3} (1 + s_1) = \alpha \nabla_{3,3} + \beta \nabla_{2,3} + \gamma \nabla_{1,3}$. The fourth equation follows by combining the third with the analogous equation $A_2 = \operatorname{span}_{\mathbf{k}} \{\nabla,\ \nabla_{3,2},\ \nabla_{2,2}\}$. The least obvious of these equalities is perhaps the fifth and last one, but even that is easy[45].) This

[45] In order to see that our list $(\nabla,\ \nabla_{3,3},\ \nabla_{2,3},\ \nabla_{3,2},\ \nabla_{2,2},\ 1)$ spans $\mathbf{k}[S_3]$, we just observe that

$$\begin{aligned}
1 &= 1, \qquad & s_1 &= \nabla_{3,3} - 1, \qquad & s_2 &= -\nabla + \nabla_{3,3} + \nabla_{2,3} + \nabla_{2,3} + \nabla_{2,2} - 1, \\
s_1 s_2 &= \nabla - \nabla_{2,2} - \nabla_{3,3} - \nabla_{2,3} + 1, \qquad & s_2 s_1 &= \nabla - \nabla_{2,2} - \nabla_{3,3} - \nabla_{3,2} + 1, \\
s_1 s_2 s_1 &= \nabla_{2,2} - 1.
\end{aligned}$$

Its linear independence can be proved by solving a system of linear equations, or by recalling the fact (see, e.g., [GriRei14, Exercise 2.5.18 **(a)**]) that any list of $d$ vectors that spans a free $\mathbf{k}$-module of rank $d$ must be a basis and therefore linearly independent.

basis is an instance of what we will later call the *(row) Murphy basis* of $\mathbf{k}[S_n]$ (see Definition 6.8.15).

Now let us use this basis to compute the subquotients of our filtration (117):

- The subquotient $\langle\nabla\rangle/0$ is a free $\mathbf{k}$-module with basis $\left(\overline{1}\right)$ (by (118) and (119) and Lemma 4.4.28). As a representation of $S_3$, it is isomorphic to the trivial representation $\mathbf{k}_{\mathrm{triv}}$ of $S_3$ (as defined in Example 4.2.16). Indeed, the $\mathbf{k}$-linear map

$$\begin{aligned} \mathbf{k}_{\mathrm{triv}} &\to \langle\nabla\rangle, \\ 1 &\mapsto \nabla \end{aligned}$$

  is $S_3$-equivariant (since $g\nabla = \nabla$ for any $g \in S_3$) and bijective, thus an isomorphism of representations of $S_3$. Hence, as representations of $S_3$, we have $\mathbf{k}_{\mathrm{triv}} \cong \langle\nabla\rangle \cong \langle\nabla\rangle/0$. Thus,

$$\langle\nabla\rangle/0 \cong \mathbf{k}_{\mathrm{triv}} \qquad \text{as } S_3\text{-representations.}$$

  Since $\mathbf{k}_{\mathrm{triv}}$ is irreducible, we conclude that $\langle\nabla\rangle/0$ is irreducible.

- The subquotient $A_1/\langle\nabla\rangle$ is a free $\mathbf{k}$-module with basis $\left(\overline{\nabla_{3,3}}, \overline{\nabla_{2,3}}\right)$ (by (119) and (120) and Lemma 4.4.28). I claim that, as a representation of $S_3$, it is isomorphic to the quotient representation $\mathbf{k}^3/D\left(\mathbf{k}^3\right)$ (where $\mathbf{k}^3$ is the natural representation of $S_3$), or, equivalently, to the zero-sum subrepresentation $R\left(\mathbf{k}^3\right)$ (where we are using the notations of Subsection 4.2.6). The best way to see this is the following: Consider the standard basis $(e_1, e_2, e_3)$ of the natural representation $\mathbf{k}^3$, and define a $\mathbf{k}$-linear map

$$\begin{aligned} \mathbf{k}^3 &\to A_1, \\ e_i &\mapsto \nabla_{i,3} \qquad \text{for each } i \in [3]. \end{aligned}$$

  This map is bijective (since $(\nabla_{3,3}, \nabla_{2,3}, \nabla_{1,3})$ is a basis of $A_1$) and $S_3$-equivariant (since all $i \in [3]$ and $g \in S_3$ satisfy $g \rightharpoonup e_i = e_{g(i)}$ and $g\nabla_{i,3} = \nabla_{g(i),3}$). Thus, it is an isomorphism of $S_3$-representations. Moreover, this isomorphism sends the submodule $D\left(\mathbf{k}^3\right)$ of $\mathbf{k}^3$ to the submodule $\langle\nabla\rangle$ of $A_1$ (since it sends $(1,1,1) \in \mathbf{k}^3$ to $\nabla_{3,3} + \nabla_{2,3} + \nabla_{1,3} = \nabla$), and thus it gives rise to a quotient isomorphism $\mathbf{k}^3/D\left(\mathbf{k}^3\right) \to A_1/\langle\nabla\rangle$ of $S_3$-representations. Hence,

$$A_1/\langle\nabla\rangle \cong \mathbf{k}^3/D\left(\mathbf{k}^3\right) \qquad \text{as } S_3\text{-representations.}$$

  Furthermore, Theorem 4.2.30 **(a)** yields that $\mathbf{k}^3/D\left(\mathbf{k}^3\right) \cong R\left(\mathbf{k}^3\right)$ as $S_3$-representations (since $\operatorname{char}\mathbf{k} \neq 3$ shows that 3 is invertible in $\mathbf{k}$). Thus,

$$A_1/\langle\nabla\rangle \cong \mathbf{k}^3/D\left(\mathbf{k}^3\right) \cong R\left(\mathbf{k}^3\right) \qquad \text{as } S_3\text{-representations.}$$

  Finally, Theorem 4.2.27 easily yields that $R\left(\mathbf{k}^3\right)$ is irreducible (since $\operatorname{char}\mathbf{k} \neq 3$ and thus $D\left(\mathbf{k}^3\right) \not\subseteq R\left(\mathbf{k}^3\right)$).

- What about the subquotient $(A_1 + A_2)/A_1$ ? We could analyze this by hand, but it is easier to use a trick:

  A general fact about modules (known as the *second isomorphism theorem*, true for left $R$-modules over an arbitrary ring $R$) says that

  $$(A_1 + A_2)/A_1 \cong A_2/(A_1 \cap A_2).$$

  But recall that $A_1 \cap A_2 = \langle \nabla \rangle$. Thus, we find

  $$(A_1 + A_2)/A_1 \cong A_2/\underbrace{(A_1 \cap A_2)}_{=\langle \nabla \rangle} = A_2/\langle \nabla \rangle.$$

  Just as we showed $A_1/\langle \nabla \rangle \cong R\left(\mathbf{k}^3\right)$ above, we can show that $A_2/\langle \nabla \rangle \cong R\left(\mathbf{k}^3\right)$ (as $S_3$-representations). Thus, we find that

  $$(A_1 + A_2)/A_1 \cong A_2/\langle \nabla \rangle \cong R\left(\mathbf{k}^3\right) \qquad \text{as } S_3\text{-representations.}$$

- Finally, what is the last subquotient $\mathbf{k}[S_3]/(A_1 + A_2)$ ? As a $\mathbf{k}$-module, it has basis $\left(\overline{1}\right)$ (again by a use of Lemma 4.4.28). To see what $S_3$-action it bears, we probe its basis element $\overline{1}$ with the two group elements $s_1$ and $t_{1,3}$ (which generate the group $S_3$). To wit, we have

  $$s_1 \rightharpoonup \overline{1} = \overline{s_1 \rightharpoonup 1} = \overline{s_1} = -\overline{1} \qquad (\text{since } 1 + s_1 \in A_1 + A_2)$$

  and

  $$t_{1,3} \rightharpoonup \overline{1} = \overline{t_{1,3} \rightharpoonup 1} = \overline{t_{1,3}} = -\overline{1} \qquad (\text{since } 1 + t_{1,3} \in A_1 + A_2).$$

  From these two equalities, it easily follows that

  $$g \rightharpoonup \overline{1} = (-1)^g \overline{1} \qquad \text{for each } g \in S_3$$

  (since each $g \in S_3$ is a product of several $s_1$'s and $t_{1,3}$'s). But this is precisely the rule for the $S_3$-action on the sign representation $\mathbf{k}_{\text{sign}}$. Hence, the $\mathbf{k}$-linear map

  $$\begin{aligned} \mathbf{k}_{\text{sign}} &\to \mathbf{k}[S_3]/(A_1 + A_2), \\ 1 &\mapsto \overline{1} \end{aligned}$$

  is $S_3$-equivariant. Since this map is also bijective[46], we conclude that it is an isomorphism of $S_3$-modules. Thus,

  $$\mathbf{k}[S_3]/(A_1 + A_2) \cong \mathbf{k}_{\text{sign}} \qquad \text{as } S_3\text{-representations.}$$

  Of course, this shows that $\mathbf{k}[S_3]/(A_1 + A_2)$ is irreducible.

Thus, all four subquotients of our filtration (117) are irreducible, so that this filtration is a composition series. Let us combine this and our other results in a proposition:

[46] because $\mathbf{k}[S_3]/(A_1 + A_2)$ is a free $\mathbf{k}$-module with basis $\left(\overline{1}\right)$

**Proposition 4.4.29.** Let $\mathbf{k}$ be a field of characteristic $\neq 3$. Then, the filtration

$$(0,\ \langle\nabla\rangle,\ A_1,\ A_1+A_2,\ \mathbf{k}\left[S_3\right])$$

is a composition series of the left regular $\mathbf{k}\left[S_3\right]$-module $\mathbf{k}\left[S_3\right]$. Its subquotients (from left to right) are isomorphic to $\mathbf{k}_{\mathrm{triv}}$, $R\left(\mathbf{k}^3\right)$, $R\left(\mathbf{k}^3\right)$ and $\mathbf{k}_{\mathrm{sign}}$. Thus, $\mathbf{k}_{\mathrm{triv}}$, $R\left(\mathbf{k}^3\right)$ and $\mathbf{k}_{\mathrm{sign}}$ are all the irreducible representations of $S_3$.

Note that $\mathbf{k}_{\mathrm{triv}} = \mathbf{k}_{\mathrm{sign}}$ if $\operatorname{char}\mathbf{k} = 2$.

**Exercise 4.4.6.** [2] Find a composition series of $\mathbf{k}\left[S_3\right]$ when $\mathbf{k}$ is a field of characteristic 3.

[**Hint:** In this case, all irreps of $S_3$ are 1-dimensional, so the composition series will be a filtration of length 6.]

So much for $\mathbf{k}\left[S_3\right]$. Analyzing $\mathbf{k}\left[S_n\right]$ with the same primitive tools is a lot harder. But I hope it is clear that finding a composition series of $\mathbf{k}\left[S_n\right]$ is a worthwhile endeavor. Thus, we pose the following goal on ourselves:

**Question 4.4.30.** Find a composition series of $\mathbf{k}\left[S_n\right]$, at least when $\mathbf{k}$ is a field of characteristic 0.

We will answer this question in Subsection 6.8.9, by constructing such a composition series (actually, we will construct two, but we will focus on just one, which generalizes (117)). This will also yield a complete description of the irreps of $S_n$, although we will have already obtained such a characterization by then (Corollary 5.13.8).

**Exercise 4.4.7. (a)** [3] Set $A_i := \{\mathbf{a} \in \mathbf{k}\left[S_n\right] \mid \mathbf{a}s_i = \mathbf{a}\}$ for each $i \in [n-1]$ (where $n \in \mathbb{N}$ and $\mathbf{k}$ are again arbitrary). Set $A := A_1 + A_2 + \cdots + A_{n-1}$. Show that

$$A = \operatorname{Ker}\left(\varepsilon \circ T_{\mathrm{sign}}\right) \qquad \text{and} \qquad \mathbf{k}\left[S_n\right] = A \oplus \operatorname{span}_{\mathbf{k}}\left\{1\right\}.$$

(See Definition 3.11.2 and Definition 3.12.1 for the meanings of $T_{\mathrm{sign}}$ and $\varepsilon$.)

**(b)** [2] Assume that $\mathbf{k}$ is a field, and that $n \geq 2$. Prove that the left regular $\mathbf{k}\left[S_n\right]$-module $\mathbf{k}\left[S_n\right]$ has a composition series of the form

$$(0,\ \langle\nabla\rangle,\ \ldots,\ A,\ \mathbf{k}\left[S_n\right]),$$

where $\langle\nabla\rangle = \operatorname{span}_{\mathbf{k}}\left\{\nabla\right\}$ and where $A$ is as in part **(a)**. (The "…" part is far less clear at this point. For $n = 2$, the "$\langle\nabla\rangle,\ \ldots,\ A$" segment should be read as a single "$\langle\nabla\rangle$".) Moreover, prove that the first and the last subquotients of this composition series are $\langle\nabla\rangle\,/0 \cong \mathbf{k}_{\mathrm{triv}}$ and $\mathbf{k}\left[S_n\right]/A \cong \mathbf{k}_{\mathrm{sign}}$ as representations of $S_n$.

(Note that applying the automorphism $T_{\mathrm{sign}}$ to the above composition series yields another composition series, which starts with $0,\ \langle\nabla^-\rangle$ instead of $0,\ \langle\nabla\rangle$. Composition series are rarely unique!)

### 4.4.5. Artin–Wedderburn

The theorems of Maschke (Theorem 4.4.14) and Jordan–Hölder (Theorem 4.4.23) are two of the most fundamental results in the representation theory of finite groups. They are among the best results of their kind that hold at their level of generality.

Stronger results exist when we assume some things about $\mathbf{k}$ (such as $\mathbf{k}$ being algebraically closed) or about $G$ (such as $G$ being a $p$-group for a prime $p$). These results are commonly found in textbooks on representation theory, such as [EGHetc11] or [Webb16]. We shall abstain from discussing them here, except for briefly presenting one known as the *Artin–Wedderburn theorem*. We will not prove nor use this theorem, but it helps us put Theorem 3.15.1 **(a)** in context.

**Theorem 4.4.31** (Artin–Wedderburn theorem)**.** Let $\mathbf{k}$ be an algebraically closed field of characteristic 0. Let $G$ be a finite group. Let $V_1, V_2, \ldots, V_k$ be all irreducible representations of $G$ over $\mathbf{k}$, each listed exactly once (up to isomorphism). Consider the curried forms $\rho_i : \mathbf{k}[G] \to \operatorname{End}_{\mathbf{k}}(V_i)$ of these representations $V_i$. Then, we can combine these curried forms into a single $\mathbf{k}$-algebra morphism

$$\begin{aligned} \rho : \mathbf{k}[G] &\to \prod_{i=1}^{k} \operatorname{End}_{\mathbf{k}}(V_i), \\ \mathbf{a} &\mapsto (\rho_1(\mathbf{a}),\ \rho_2(\mathbf{a}),\ \ldots,\ \rho_k(\mathbf{a})). \end{aligned}$$

This morphism is a $\mathbf{k}$-algebra isomorphism! Thus, in particular,

$$\mathbf{k}[G] \cong \prod_{i=1}^{k} \operatorname{End}_{\mathbf{k}}(V_i) \qquad \text{as } \mathbf{k}\text{-algebras.}$$

Since each $V_i$ is a finite-dimensional $\mathbf{k}$-vector space, we can identify its endomorphism ring $\operatorname{End}_{\mathbf{k}}(V_i)$ with a matrix ring $\mathbf{k}^{f_i \times f_i}$, where $f_i = \dim V_i$. Thus,

$$\mathbf{k}[G] \cong \prod_{i=1}^{k} \operatorname{End}_{\mathbf{k}}(V_i) \cong \prod_{i=1}^{k} \mathbf{k}^{f_i \times f_i} \qquad \text{as } \mathbf{k}\text{-algebras}$$

(a direct product of matrix rings). By comparing dimensions, this entails

$$|G| = \sum_{i=1}^{k} f_i^2. \tag{123}$$

This generalizes the $\mathbf{k}[S_3] \cong \mathbf{k} \times \mathbf{k} \times \mathbf{k}^{2\times 2}$ isomorphism that we found in Subsection 3.15.2, and to some extent generalizes Theorem 3.15.1 **(a)**. However, the generality comes at the cost of stronger requirements (since the Artin–Wedderburn theorem requires $\mathbf{k}$ to be algebraically closed) and weaker claims

(the Artin–Wedderburn theorem says nothing about what the $f_i$ are for any given $G$). We will thus have to prove Theorem 3.15.1 **(a)** independently of the Artin–Wedderburn theorem.

Theorem 3.15.1 **(a)** shows that the Artin–Wedderburn theorem holds for $G = S_n$ even if $\mathbf{k}$ is not algebraically closed.[47] For other finite groups $G$, the "algebraically closed" requirement in the Artin–Wedderburn theorem can be necessary. For example, the cyclic group $C_n$ has $n$ mutually non-isomorphic 1-dimensional representations over the field $\mathbb{C}$ (all irreducible, of course), and no other irreps; thus, the Artin–Wedderburn theorem yields a $\mathbb{C}$-algebra isomorphism

$$\mathbb{C}\left[C_n\right] \cong \underbrace{\mathbb{C}^{1\times 1} \times \mathbb{C}^{1\times 1} \times \cdots \times \mathbb{C}^{1\times 1}}_{n \text{ times}} = \left(\mathbb{C}^{1\times 1}\right)^n \cong \mathbb{C}^n.$$

This is precisely the isomorphism we saw in Exercise 3.16.5 (and called the discrete Fourier transform). But no such isomorphism exists for $\mathbf{k} = \mathbb{R}$.

The Artin–Wedderburn theorem is very useful for computations:

- It allows you to work in the $\mathbf{k}$-algebra $\prod_{i=1}^{k} \mathbf{k}^{f_i \times f_i}$ instead of $\mathbf{k}\left[G\right]$. In other words, instead of calculating with formal linear combinations of elements of $G$, you can calculate with $k$-tuples of matrices. The latter are better supported by computer algebra software (and the entrywise nature of operations in $\prod_{i=1}^{k} \mathbf{k}^{f_i \times f_i}$ is conducive to parallelization).
- If you have found a bunch of mutually non-isomorphic irreps $V_1, V_2, \ldots, V_m$ of a finite group $G$ over an algebraically closed field $\mathbf{k}$, then the formula (123) helps you determine whether you have found all irreps of $G$ or there are more to be found: Indeed, if $|G| = \sum_{i=1}^{m} f_i^2$, then you have found them all. If $|G| > \sum_{i=1}^{m} f_i^2$, then there are more. (Why?)
- The isomorphism $\rho$ in Theorem 4.4.31 is how the magic elements $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}$ were found in Subsection 3.15.2. Indeed, they should be the preimages of $\left(0, 0, E_{i,j}\right)$ under $\rho$, so we can compute them using linear algebra (viewing $\rho$ as a matrix, and inverting this matrix).

**Exercise 4.4.8.** $\boxed{1}$ Show that the Artin–Wedderburn theorem fails for $G = C_3$ and $\mathbf{k} = \mathbb{R}$. In other words, show that $\mathbb{R}\left[C_3\right]$ is not isomorphic (as an $\mathbb{R}$-algebra) to the direct product of the endomorphism rings of the irreps of $C_3$ over $\mathbb{R}$.

[47] Caveat: Theorem 3.15.1 **(a)**, as stated, says that $\mathbf{k}\left[S_n\right]$ is isomorphic to a direct product of matrix rings, but it does not give any specific isomorphism. Thus, it does not properly yield the claim of the Artin–Wedderburn theorem for $G = S_n$. But the proof does.

## 4.5. Noncommutative polynomials

In this section, we will study yet another way to obtain representations of symmetric groups, or two such ways to be precise, since two different symmetric groups will be acting on the same space here.

The space in question is a space of *noncommutative polynomials*. We begin by defining them first. Then we will define the two actions of symmetric groups on them, and we will use them to half-prove Theorem 3.10.10 **(b)** ("half-prove" because we will only reduce it to a result from noncommutative algebra that we won't prove). Note that noncommutative polynomials are useful far beyond representation theory.

### 4.5.1. Definition and symmetric group actions

Let $k \in \mathbb{N}$. Recall that the usual (i.e., commutative) polynomial ring $\mathbf{k}[x_1, x_2, \ldots, x_k]$ was defined as the monoid algebra of the monoid of monomials over $\mathbf{k}$. Monomials are, in essence, expressions of the form $x_1^{a_1} x_2^{a_2} \cdots x_k^{a_k}$ with $a_1, a_2, \ldots, a_k \in \mathbb{N}$. They are multiplied by adding the respective exponents:

$$\left(x_1^{a_1} x_2^{a_2} \cdots x_k^{a_k}\right) \left(x_1^{b_1} x_2^{b_2} \cdots x_k^{b_k}\right) = x_1^{a_1+b_1} x_2^{a_2+b_2} \cdots x_k^{a_k+b_k}.$$

Let us try to define a noncommutative polynomial ring $\mathbf{k}\langle x_1, x_2, \ldots, x_k\rangle$ similarly: It shall be the monoid algebra of the monoid of "noncommutative monomials" over $\mathbf{k}$. Noncommutative monomials are, in essence, expressions of the form $x_{i_1} x_{i_2} \cdots x_{i_d}$ with $i_1, i_2, \ldots, i_d \in [k]$. (We want noncommutativity, so we must not equate $x_1 x_2 x_1$ with $x_1 x_1 x_2$. Thus, we cannot rewrite such monomials as $x_1^{a_1} x_2^{a_2} \cdots x_k^{a_k}$ in general!) Such noncommutative monomials are multiplied by concatenation:

$$\left(x_{i_1} x_{i_2} \cdots x_{i_d}\right) \left(x_{j_1} x_{j_2} \cdots x_{j_e}\right) = x_{i_1} x_{i_2} \cdots x_{i_d} x_{j_1} x_{j_2} \cdots x_{j_e}.$$

To make this rigorous, we encode these (so far undefined) noncommutative monomials $x_{i_1} x_{i_2} \cdots x_{i_d}$ as the tuples $(i_1, i_2, \ldots, i_d)$ consisting of their subscripts. Thus, we arrive at the following rigorous definition:

**Definition 4.5.1.** Let $k \in \mathbb{N}$.

**(a)** A *noncommutative monomial* in $k$ indeterminates $x_1, x_2, \ldots, x_k$ is just a tuple (i.e., finite list) of elements of $[k]$. The *concatenation* $ij$ of two such tuples $i = (i_1, i_2, \ldots, i_d)$ and $j = (j_1, j_2, \ldots, j_e)$ is defined to be the tuple

$$ij := (i_1, i_2, \ldots, i_d, j_1, j_2, \ldots, j_e).$$

We define $\mathfrak{NM}_k$ to be the monoid of all noncommutative monomials (in $k$ indeterminates $x_1, x_2, \ldots, x_k$), where the multiplication is defined to be concatenation. The neutral element of $\mathfrak{NM}_k$ is the empty tuple $()$.

**(b)** For each $i \in [k]$, we let $x_i$ denote the 1-tuple $(i) \in \mathfrak{NM}_k$. Thus, each noncommutative monomial $(i_1, i_2, \ldots, i_d) \in \mathfrak{NM}_k$ can be rewritten as $x_{i_1} x_{i_2} \cdots x_{i_d}$.

**(c)** The *noncommutative polynomial ring* $\mathbf{k}\langle x_1, x_2, \ldots, x_k \rangle$ (also denoted $\mathbf{NP}_k$ in this section) is the monoid algebra of the monoid $\mathfrak{NM}_k$ over $\mathbf{k}$. Its elements are called *noncommutative polynomials* in $k$ indeterminates $x_1, x_2, \ldots, x_k$ over $\mathbf{k}$.

For example, $x_3 x_1 - 5 x_1 x_2^4 x_3 x_2 + 7$ is a noncommutative polynomial over $\mathbb{Z}$. (Note that $x_2^4 = x_2 x_2 x_2 x_2$.)

Next, we define some basic features of noncommutative polynomials, analogous to corresponding notions for usual (commutative) polynomials:

**Definition 4.5.2. (a)** The *degree* of a noncommutative monomial $\mathfrak{m} = x_{i_1} x_{i_2} \cdots x_{i_d}$ is defined to be $d$. (This is just the length of the tuple $\mathfrak{m}$.) This degree is denoted $\deg \mathfrak{m}$.

**(b)** Let $d \in \mathbb{N}$. A noncommutative polynomial $p \in \mathbf{NP}_k$ is said to be *homogeneous of degree $d$* if it is a $\mathbf{k}$-linear combination of noncommutative monomials of degree $d$. (Note that the zero of $\mathbf{NP}_k$ is homogeneous of any degree.)

**(c)** Let $d \in \mathbb{N}$. We let $\mathbf{NP}_{k,d}$ be the $\mathbf{k}$-submodule of $\mathbf{NP}_k$ consisting of all noncommutative polynomials homogeneous of degree $d$. Equivalently, this is the $\mathbf{k}$-linear span of all noncommutative monomials of degree $d$. We call this the *degree-$d$ component* of $\mathbf{NP}_k$.

Here are some properties of noncommutative polynomials:

- The degree of a product is the sum of the degrees: If $p \in \mathbf{NP}_{k,d}$ and $q \in \mathbf{NP}_{k,e}$, then $pq \in \mathbf{NP}_{k,d+e}$.
- Each noncommutative polynomial can be written as the sum of its homogeneous components (almost all of which are zero). Thus, $\mathbf{NP}_k = \bigoplus_{d \in \mathbb{N}} \mathbf{NP}_{k,d}$ as a $\mathbf{k}$-module.
- We can *evaluate* a noncommutative polynomial $p \in \mathbf{NP}_k$ at any $k$-tuple $(a_1, a_2, \ldots, a_k)$ of elements of any $\mathbf{k}$-algebra $A$. These elements $a_1, a_2, \ldots, a_k$ don't even have to commute, unlike for usual (commutative) polynomials. This evaluation is defined in the obvious way: We "plug" $a_i$ for each $x_i$ (that is, we replace each noncommutative monomial $x_{i_1} x_{i_2} \cdots x_{i_d}$ by the product $a_{i_1} a_{i_2} \cdots a_{i_d}$). The result of this evaluation is denoted by $p(a_1, a_2, \ldots, a_k)$.

If $a_1, a_2, \ldots, a_k$ are $k$ elements of a $\mathbf{k}$-algebra $A$, then the map

$$\begin{aligned} \mathbf{NP}_k &\to A, \\ p &\mapsto p\left(a_1, a_2, \ldots, a_k\right) \end{aligned}$$

is a $\mathbf{k}$-algebra morphism (called the *evaluation morphism* at $a_1, a_2, \ldots, a_k$). This yields a universal property of the noncommutative polynomial ring $\mathbf{NP}_k$ (similar to the one of a usual polynomial ring, but no longer requiring commutativity).

- If $\mathbf{k}$ is an integral domain, then $\mathbf{NP}_k$ is a "noncommutative integral domain" (i.e., a ring in which a product of nonzero elements is always nonzero). This can be proved using leading terms with respect to an appropriately defined lexicographic order on monomials.
- For $k = 1$, there is no difference between noncommutative polynomials and usual polynomials. That is, we have $\mathbf{NP}_1 \cong \mathbf{k}\left[x\right]$ as $\mathbf{k}$-algebras. (The isomorphism sends each noncommutative monomial $\underbrace{x_1 x_1 \cdots x_1}_{i \text{ times}}$ to the commutative monomial $x^i$.)

**Remark 4.5.3.** If you know tensor algebra, then you will recognize the noncommutative polynomial ring $\mathbf{NP}_k$ as (an isomorphic copy of) the tensor algebra of the free $\mathbf{k}$-module $\mathbf{k}^k$.

More concretely: For any $d \in \mathbb{N}$, we have a $\mathbf{k}$-module isomorphism

$$\mathbf{NP}_{k,d} \to \underbrace{\mathbf{k}^k \otimes \mathbf{k}^k \otimes \cdots \otimes \mathbf{k}^k}_{d \text{ times}},$$

which sends each noncommutative monomial $x_{i_1} x_{i_2} \cdots x_{i_d}$ to the pure tensor $e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_d}$. Taking the direct sum of these isomorphisms over all $d \in \mathbb{N}$, we obtain a $\mathbf{k}$-algebra isomorphism from $\mathbf{NP}_k$ to the tensor algebra of $\mathbf{k}^k$.

Now we define actions of the symmetric groups $S_k$ and $S_d$ on $\mathbf{NP}_{k,d}$:

**Definition 4.5.4.** Let $k, d \in \mathbb{N}$. Let $\mathfrak{MM}_{k,d}$ be the set of all noncommutative monomials of degree $d$ in the $k$ indeterminates $x_1, x_2, \ldots, x_k$. These monomials form a basis of the $\mathbf{k}$-module $\mathbf{NP}_{k,d}$, and thus we can view $\mathbf{NP}_{k,d}$ as the free $\mathbf{k}$-module $\mathbf{k}^{\left(\mathfrak{MM}_{k,d}\right)}$.

**(a)** We define a left action of the symmetric group $S_k$ on $\mathfrak{MM}_{k,d}$ by setting

$$\begin{aligned} g \rightharpoonup \left(x_{i_1} x_{i_2} \cdots x_{i_d}\right) &= x_{g(i_1)} x_{g(i_2)} \cdots x_{g(i_d)} \\ &\text{for all } g \in S_k \text{ and } x_{i_1} x_{i_2} \cdots x_{i_d} \in \mathfrak{MM}_{k,d}. \end{aligned}$$

This is called the *action on entries*. The permutation module corresponding to this action (as explained in Example 4.3.9) is thus a left $\mathbf{k}[S_k]$-module $\mathbf{k}^{(\mathfrak{NM}_{k,d})} = \mathbf{NP}_{k,d}$. Hence, $\mathbf{NP}_{k,d}$ becomes a representation of $S_k$.

**(b)** We define a left action of the symmetric group $S_d$ on $\mathfrak{NM}_{k,d}$ by setting

$$g \rightharpoonup \left(x_{i_1} x_{i_2} \cdots x_{i_d}\right) = x_{i_{g^{-1}(1)}} x_{i_{g^{-1}(2)}} \cdots x_{i_{g^{-1}(d)}}$$
$$\text{for all } g \in S_d \text{ and } x_{i_1} x_{i_2} \cdots x_{i_d} \in \mathfrak{NM}_{k,d}.$$

This is called the *action on places*. The permutation module corresponding to this action (as explained in Example 4.3.9) is thus a left $\mathbf{k}[S_d]$-module $\mathbf{k}^{(\mathfrak{NM}_{k,d})} = \mathbf{NP}_{k,d}$. Hence, $\mathbf{NP}_{k,d}$ becomes a representation of $S_d$.

(These two representations are not the same even when $k = d$.)

Strictly speaking, much of this definition is redundant: After all, our noncommutative monomials $x_{i_1} x_{i_2} \cdots x_{i_d}$ of degree $d$ are just the $d$-tuples $(i_1, i_2, \ldots, i_d) \in [k]^d$ disguised by a fancy notation, and thus the set $\mathfrak{NM}_{k,d}$ is just $[k]^d$. The two actions of $S_k$ and $S_d$ that we have just defined on this set are precisely the action on entries defined in Example 4.1.16 (for $A = [k]$) and the action on places defined in Example 4.1.17 (again for $A = [k]$). The only new thing we are doing now is extending these actions (by linearity) to the noncommutative **poly**nomials, i.e., to the $\mathbf{k}$-linear combinations of the noncommutative monomials. For instance, for $k = 4$ and $d = 4$, we have

$$\mathrm{cyc}_{1,2,3} \rightharpoonup \left(x_2 x_4 x_1 x_2 + 5 x_1^3 x_2\right) = x_3 x_4 x_2 x_3 + 5 x_2^3 x_3 \qquad \text{(using the action on entries)}$$

and

$$\mathrm{cyc}_{1,2,3} \rightharpoonup \left(x_2 x_4 x_1 x_2 + 5 x_1^3 x_2\right) = x_1 x_2 x_4 x_2 + 5 x_1^3 x_2 \qquad \text{(using the action on places)}.$$

The action on entries can be extended to the whole $\mathbf{k}$-algebra $\mathbf{NP}_k$ (not just the degree-$d$ component $\mathbf{NP}_{k,d}$), but the action on places cannot (since different groups would be acting on different degrees).

**Remark 4.5.5.** The two actions commute, i.e., we have $g \rightharpoonup (h \rightharpoonup p) = h \rightharpoonup (g \rightharpoonup p)$ for any $g \in S_k$, any $h \in S_d$ and any $p \in \mathfrak{NM}_{k,d}$. Thus, they can be combined into a single action of $S_k \times S_d$ on $\mathfrak{NM}_{k,d}$ given by

$$(g, h) \rightharpoonup \left(x_{i_1} x_{i_2} \cdots x_{i_d}\right) = x_{g\left(i_{h^{-1}(1)}\right)} x_{g\left(i_{h^{-1}(2)}\right)} \cdots x_{g\left(i_{h^{-1}(d)}\right)}.$$

This will not be used in what follows.

Let us now explore the faithfulness of above-defined module structures on $\mathbf{NP}_{k,d}$:

**Proposition 4.5.6.** **(a)** The left $\mathbf{k}[S_d]$-module $\mathbf{NP}_{k,d}$ (coming from the action on places) is faithful if $k \geq d$.

**(b)** If $k < d$, then the element $\nabla^- \in \mathbf{k}[S_d]$ acts as 0 on the left $\mathbf{k}[S_d]$-module $\mathbf{NP}_{k,d}$ (coming from the action on places).

**(c)** The left $\mathbf{k}[S_k]$-module $\mathbf{NP}_{k,d}$ (coming from the action on entries) is faithful if $k \geq d-1$.

**(d)** If $k < d-1$, then the element $\nabla^- \in \mathbf{k}[S_k]$ acts as 0 on the left $\mathbf{k}[S_k]$-module $\mathbf{NP}_{k,d}$ (coming from the action on entries).

*Proof of Proposition 4.5.6* ***(a)****.* Assume that $k \geq d$. Consider the noncommutative monomial $\mathfrak{m} = x_1 x_2 \cdots x_d \in \mathfrak{MN}_{k,d}$ (this is well-defined, since $k \geq d$; otherwise there would be no $x_d$). Then, any $g \in S_d$ satisfies

$$g \rightharpoonup \mathfrak{m} = x_{g^{-1}(1)} x_{g^{-1}(2)} \cdots x_{g^{-1}(d)} \tag{124}$$

(by the definition of the action on places).

Now, let $\mathbf{a}$ and $\mathbf{b}$ be two elements of $\mathbf{k}[S_d]$ such that $\mathbf{a}\mathfrak{m} = \mathbf{b}\mathfrak{m}$. We shall show that $\mathbf{a} = \mathbf{b}$.

Indeed, write the difference $\mathbf{a} - \mathbf{b} \in \mathbf{k}[S_d]$ in the form

$$\mathbf{a} - \mathbf{b} = \sum_{g \in S_d} \lambda_g g \tag{125}$$

for some scalars $\lambda_g \in \mathbf{k}$. Then,

$$\begin{aligned}(\mathbf{a}-\mathbf{b})\,\mathfrak{m} &= \left(\sum_{g\in S_d} \lambda_g g\right)\mathfrak{m} = \sum_{g\in S_d} \lambda_g \underbrace{g \rightharpoonup \mathfrak{m}}_{\substack{= x_{g^{-1}(1)} x_{g^{-1}(2)} \cdots x_{g^{-1}(d)} \\ \text{(by (124))}}} \\ &= \sum_{g\in S_d} \lambda_g x_{g^{-1}(1)} x_{g^{-1}(2)} \cdots x_{g^{-1}(d)}.\end{aligned}$$

Hence,

$$\begin{aligned}\sum_{g\in S_d} \lambda_g x_{g^{-1}(1)} x_{g^{-1}(2)} \cdots x_{g^{-1}(d)} &= (\mathbf{a}-\mathbf{b})\,\mathfrak{m} = \mathbf{a}\mathfrak{m} - \mathbf{b}\mathfrak{m} \\ &= 0\end{aligned} \tag{126}$$

(since we assumed that $\mathbf{a}\mathfrak{m} = \mathbf{b}\mathfrak{m}$).

But the noncommutative monomials $x_{g^{-1}(1)} x_{g^{-1}(2)} \cdots x_{g^{-1}(d)}$ for all $g \in S_d$ are distinct (since they are just the $d$-tuples $\left(g^{-1}(1), g^{-1}(2), \ldots, g^{-1}(d)\right)$, and clearly you can reconstruct a permutation $g \in S_d$ from this $d$-tuple), and thus are $\mathbf{k}$-linearly independent in $\mathbf{NP}_{k,d}$. In other words, a $\mathbf{k}$-linear combination of these monomials $x_{g^{-1}(1)} x_{g^{-1}(2)} \cdots x_{g^{-1}(d)}$ can only be 0 if all coefficients in it are 0.

Hence, the equality (126) shows that all coefficients $\lambda_g$ are 0. Thus, we can rewrite (125) as $\mathbf{a} - \mathbf{b} = \sum_{g \in S_d} 0g = 0$, so that $\mathbf{a} = \mathbf{b}$.

Forget that we fixed $\mathbf{a}$ and $\mathbf{b}$. We thus have shown that if $\mathbf{a}$ and $\mathbf{b}$ are two elements of $\mathbf{k}[S_d]$ such that $\mathbf{a}\mathfrak{m} = \mathbf{b}\mathfrak{m}$, then $\mathbf{a} = \mathbf{b}$. In other words, if two elements of $\mathbf{k}[S_d]$ act in the same way on $\mathfrak{m}$, then they are equal. Thus, we can tell distinct elements of $\mathbf{k}[S_d]$ apart by looking at how they act on $\mathfrak{m}$. Therefore, the action of $\mathbf{k}[S_d]$ is faithful (and moreover, just looking at the action on the single element $\mathfrak{m}$ is already enough for this faithfulness). This proves Proposition 4.5.6 **(a)**. □

**Exercise 4.5.1.** [5 (2, 1, 2)] Prove the other three parts of Proposition 4.5.6.

More generally, the elements of $\mathbf{k}[S_d]$ or of $\mathbf{k}[S_k]$ that act as 0 on $\mathbf{NP}_{k,d}$ can be characterized – see [Grinbe25, Theorems 2.8.1 and 2.8.3] (where $\mathbf{NP}_{k,d}$ appears in the equivalent guise of the tensor power $\left(\mathbf{k}^k\right)^{\otimes d}$).

### 4.5.2. Application: the Dynkin element $\mathbf{V}_n$

Using Proposition 4.5.6 **(a)**, we can prove the equality of two elements of $\mathbf{k}[S_d]$ by showing that they act equally on $\mathbf{NP}_{k,d}$ for some $k \geq d$. This technique can be quite useful; let me show an example of its use. Sadly, the proof will be incomplete, since it boils down to an algebraic result that cannot be shown very easily.

Assume that $n \geq 1$ for the rest of Subsection 4.5.2.

Recall from Theorem 3.10.10 that

$$\begin{aligned}\mathbf{V}_n &:= \left(1 - \mathrm{cyc}_{2,1}\right)\left(1 - \mathrm{cyc}_{3,2,1}\right)\left(1 - \mathrm{cyc}_{4,3,2,1}\right)\cdots\left(1 - \mathrm{cyc}_{n,n-1,\ldots,1}\right) \\ &= \prod_{i=2}^{n}\left(1 - \mathrm{cyc}_{i,i-1,\ldots,1}\right) \in \mathbf{k}[S_n].\end{aligned}$$

Theorem 3.10.10 **(b)** claims that $\mathbf{V}_n^2 = n\mathbf{V}_n$. We have not proved this so far.

How can we prove this? The first step will be a simple transformation inside $\mathbf{k}[S_n]$. Indeed, recall the antipode $S$ of $\mathbf{k}[S_n]$ (see Definition 3.11.12). We know from Theorem 3.11.14 **(a)** that the antipode $S$ is an algebra anti-automorphism. In particular, $S$ is injective. Thus, in order to prove $\mathbf{V}_n^2 = n\mathbf{V}_n$, it will suffice to show that $S\left(\mathbf{V}_n^2\right) = S\left(n\mathbf{V}_n\right)$. But we have $S\left(\mathbf{V}_n^2\right) = S\left(\mathbf{V}_n\right)^2$ (since $S$ is an algebra anti-automorphism) and $S\left(n\mathbf{V}_n\right) = nS\left(\mathbf{V}_n\right)$ (likewise). Hence, it will suffice to show that

$$S\left(\mathbf{V}_n\right)^2 = nS\left(\mathbf{V}_n\right). \tag{127}$$

Note that the definition of $\mathbf{V}_n$ yields

$$\begin{aligned} S\left(\mathbf{V}_n\right) &= S\left(\left(1-\mathrm{cyc}_{2,1}\right)\left(1-\mathrm{cyc}_{3,2,1}\right)\left(1-\mathrm{cyc}_{4,3,2,1}\right)\cdots\left(1-\mathrm{cyc}_{n,n-1,\ldots,1}\right)\right) \\ &= S\left(1-\mathrm{cyc}_{n,n-1,\ldots,1}\right)\cdot S\left(1-\mathrm{cyc}_{n-1,n-2,\ldots,1}\right)\cdot\cdots\cdot S\left(1-\mathrm{cyc}_{2,1}\right) \\ &\qquad \left(\text{since } S \text{ is an algebra anti-automorphism}\right) \\ &= \left(1-\mathrm{cyc}_{1,2,\ldots,n}\right)\left(1-\mathrm{cyc}_{1,2,\ldots,n-1}\right)\cdots\left(1-\mathrm{cyc}_{1,2}\right) \end{aligned} \tag{128}$$

(since each $k \in \{2,3,\ldots,n\}$ satisfies

$$\begin{aligned} S\left(1-\mathrm{cyc}_{k,k-1,\ldots,1}\right) &= \underbrace{1^{-1}}_{=1} - \underbrace{\mathrm{cyc}_{k,k-1,\ldots,1}^{-1}}_{=\mathrm{cyc}_{1,2,\ldots,k}} \qquad \left(\text{by the definition of } S\right) \\ &= 1-\mathrm{cyc}_{1,2,\ldots,k} \end{aligned}$$

).

Now is the time to use faithful modules. Recall that the left $\mathbf{k}\left[S_n\right]$-module $\mathbf{NP}_{k,n}$ (coming from the action on places) is faithful for any integer $k \geq n$ (by Proposition 4.5.6 **(a)**, applied to $d = n$). Hence, in order to show that $S\left(\mathbf{V}_n\right)^2 = nS\left(\mathbf{V}_n\right)$, it will suffice to prove that

$$S\left(\mathbf{V}_n\right)^2 \cdot p = nS\left(\mathbf{V}_n\right)\cdot p \tag{129}$$

for all $k \in \mathbb{N}$ and all noncommutative polynomials $p \in \mathbf{NP}_{k,n}$, where we are using the action on places. (Actually, it will even suffice to prove (129) for any one integer $k \geq n$; but this is no easier than proving (129) for all $k \in \mathbb{N}$.)

We fix an integer $k \in \mathbb{N}$ for the rest of Subsection 4.5.2.

In order to prove (129), we must figure out how $S\left(\mathbf{V}_n\right)$ acts on $\mathbf{NP}_{k,n}$. There is a really nice formula for this, which is in fact the reason why $\mathbf{V}_n$ was first introduced.

Before we state it, we recall that the set $\mathbf{NP}_{k,1}$ consists of homogeneous noncommutative polynomials of degree 1; these are $\mathbf{k}$-linear combinations $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k$ of indeterminates. Any product $a_1 a_2 \cdots a_n$ of $n$ such combinations $a_1, a_2, \ldots, a_n \in \mathbf{NP}_{k,1}$ is a homogeneous polynomial of degree $n$, that is, an element of $\mathbf{NP}_{k,n}$. Moreover, all noncommutative monomials $x_{i_1} x_{i_2} \cdots x_{i_n}$ of degree $n$ are such products (since the indeterminates $x_{i_1}, x_{i_2}, \ldots, x_{i_n}$ themselves belong to $\mathbf{NP}_{k,1}$). Hence, such products $a_1 a_2 \cdots a_n$ span the $\mathbf{k}$-module $\mathbf{NP}_{k,n}$. Thus, a $\mathbf{k}$-linear map on $\mathbf{NP}_{k,n}$ is uniquely determined by its values on such products $a_1 a_2 \cdots a_n$.

We shall describe the action of $S\left(\mathbf{V}_n\right)$ on $\mathbf{NP}_{k,n}$ by how it acts on such products. To do so, we introduce another notation:

Recall that the commutator $ab - ba$ of two elements $a$ and $b$ of a ring is denoted by $[a,b]$. More generally, if $a_1, a_2, \ldots, a_m$ are $m$ elements of a ring (where $m \geq 1$), then their *left-nested commutator* $\left[\cdots\left[\left[a_1, a_2\right], a_3\right], \ldots, a_m\right]$ is defined recursively as follows:

- For $m = 1$, it is defined to be $a_1$.
- For $m > 1$, it is defined to be the (usual) commutator $[u, a_m]$, where $u$ is the left-nested commutator $[\cdots [[a_1, a_2], a_3], \ldots, a_{m-1}]$ of the first $m-1$ elements $a_1, a_2, \ldots, a_{m-1}$.

Thus, the left-nested commutator $[\cdots [[a_1, a_2], a_3], \ldots, a_m]$ is the usual commutator $[a_1, a_2]$ when $m = 2$, the nested commutator $[[a_1, a_2], a_3]$ when $m = 3$, and the nested commutator $[[[a_1, a_2], a_3], a_4]$ when $m = 4$.

Now we can finally describe the action of $S(\mathbf{V}_n)$ on $\mathbf{NP}_{k,n}$:

**Proposition 4.5.7.** Let $n \geq 1$. Let $a_1, a_2, \ldots, a_n \in \mathbf{NP}_{k,1}$. Then,

$$S(\mathbf{V}_n) \cdot (a_1 a_2 \cdots a_n) = [\cdots [[a_1, a_2], a_3], \ldots, a_n].$$

For instance, for $n = 3$, this is saying that

$$S(\mathbf{V}_3) \cdot (abc) = [[a, b], c] \qquad \text{for all } a, b, c \in \mathbf{NP}_{k,1}.$$

Let us check this: By (128), we have

$$S(\mathbf{V}_3) = \left(1 - \mathrm{cyc}_{1,2,3}\right)\left(1 - \mathrm{cyc}_{1,2}\right) = 1 - \mathrm{cyc}_{1,2,3} - s_1 + t_{1,3}$$

and thus

$$\begin{aligned} S(\mathbf{V}_3) \cdot (abc) &= \left(1 - \mathrm{cyc}_{1,2,3} - s_1 + t_{1,3}\right) \cdot abc \\ &= abc - cab - bac + cba \qquad \text{(action on places!)}, \end{aligned}$$

which indeed agrees with

$$\begin{aligned} [[a, b], c] &= [ab - ba,\ c] = (ab - ba)\, c - c\,(ab - ba) \\ &= abc - bac - cab + cba. \end{aligned}$$

The proof of Proposition 4.5.7 is not hard. The main step is the following lemma:

**Lemma 4.5.8.** Let $j \in [n]$. Let $u, g, v \in \mathbf{NP}_k$ be homogeneous noncommutative polynomials of respective degrees $j-1$, $1$ and $n-j$. Then,

$$\mathrm{cyc}_{1,2,\ldots,j} \cdot (ugv) = guv \tag{130}$$

and

$$\left(1 - \mathrm{cyc}_{1,2,\ldots,j}\right) \cdot (ugv) = [u, g]\, v. \tag{131}$$

*Proof.* Both equalities (130) and (131) depend $\mathbf{k}$-linearly on each of $u$, $g$ and $v$. Thus, we can WLOG assume that the homogeneous noncommutative polynomials $u$, $g$ and $v$ are actually noncommutative monomials of the respective degrees (since the noncommutative monomials of a given degree $d$ span $\mathbf{NP}_{k,d}$). In other words, we can WLOG assume that

$$u = x_{i_1} x_{i_2} \cdots x_{i_{j-1}}, \qquad g = x_{i_j}, \qquad \text{and} \qquad v = x_{i_{j+1}} x_{i_{j+2}} \cdots x_{i_n}$$

for some $i_1, i_2, \ldots, i_n \in [k]$. Assume this, and consider these $i_1, i_2, \ldots, i_n$. Multiplying the three equalities

$$u = x_{i_1} x_{i_2} \cdots x_{i_{j-1}}, \qquad g = x_{i_j}, \qquad \text{and} \qquad v = x_{i_{j+1}} x_{i_{j+2}} \cdots x_{i_n}$$

together, we obtain

$$ugv = \left(x_{i_1} x_{i_2} \cdots x_{i_{j-1}}\right) x_{i_j} \left(x_{i_{j+1}} x_{i_{j+2}} \cdots x_{i_n}\right) = x_{i_1} x_{i_2} \cdots x_{i_n}.$$

Set $\sigma := \mathrm{cyc}_{1,2,\ldots,j}$. Then, $\sigma^{-1} = \mathrm{cyc}_{1,2,\ldots,j}^{-1} = \mathrm{cyc}_{j,j-1,\ldots,1}$ and therefore

$$\begin{aligned} &\left(\sigma^{-1}(1), \sigma^{-1}(2), \ldots, \sigma^{-1}(n)\right) \\ &= (j, 1, 2, \ldots, j-1, j+1, j+2, \ldots, n). \end{aligned} \tag{132}$$

But

$$\begin{aligned} \sigma \cdot \underbrace{(ugv)}_{= x_{i_1} x_{i_2} \cdots x_{i_n}} &= \sigma \cdot \left(x_{i_1} x_{i_2} \cdots x_{i_n}\right) \\ &= x_{i_{\sigma^{-1}(1)}} x_{i_{\sigma^{-1}(2)}} \cdots x_{i_{\sigma^{-1}(n)}} \qquad \left(\begin{array}{c}\text{by the definition of}\\ \text{the action on places}\end{array}\right) \\ &= \underbrace{x_{i_j}}_{=g} \underbrace{x_{i_1} x_{i_2} \cdots x_{i_{j-1}}}_{=u} \underbrace{x_{i_{j+1}} x_{i_{j+2}} \cdots x_{i_n}}_{=v} \qquad \text{(by (132))} \\ &= guv \end{aligned}$$

and thus

$$(1-\sigma) \cdot (ugv) = ugv - \underbrace{\sigma \cdot (ugv)}_{=guv} = ugv - guv = \underbrace{(ug - gu)}_{=[u,g]} v = [u, g]\, v.$$

In view of $\sigma = \mathrm{cyc}_{1,2,\ldots,j}$, we can rewrite these two equalities as

$$\mathrm{cyc}_{1,2,\ldots,j} \cdot (ugv) = guv$$

and

$$\left(1 - \mathrm{cyc}_{1,2,\ldots,j}\right) \cdot (ugv) = [u, g]\, v.$$

Thus, Lemma 4.5.8 is proved. □

*Proof of Proposition 4.5.7.* We start with an example: For $n = 4$, we have

$$S(\mathbf{V}_4) = \left(1 - \mathrm{cyc}_{1,2,3,4}\right)\left(1 - \mathrm{cyc}_{1,2,3}\right)\left(1 - \mathrm{cyc}_{1,2}\right) \qquad \text{(by (128))}$$

and thus

$$\begin{aligned}
&S(\mathbf{V}_4) \cdot (abcd) \\
&= \left(1 - \mathrm{cyc}_{1,2,3,4}\right)\left(1 - \mathrm{cyc}_{1,2,3}\right) \underbrace{\left(1 - \mathrm{cyc}_{1,2}\right) \cdot (abcd)}_{\substack{=[a,b]cd \\ \text{(by (131), applied to } j=2,\ u=a,\ g=b \text{ and } v=cd)}} \\
&= \left(1 - \mathrm{cyc}_{1,2,3,4}\right) \underbrace{\left(1 - \mathrm{cyc}_{1,2,3}\right) \cdot ([a,b]\, cd)}_{\substack{=[[a,b],c]d \\ \text{(by (131), applied to } j=3,\ u=ab,\ g=c \text{ and } v=d)}} \\
&= \left(1 - \mathrm{cyc}_{1,2,3,4}\right) \cdot ([[a,b]\,,c]\, d) \\
&= [[[a,b]\,,c]\,,d] \qquad \text{(by (131), applied to } j=4,\ u=abc,\ g=d \text{ and } v=1).
\end{aligned}$$

The same type of computation works for any $n$: Repeated use of (131) gradually turns the product $a_1 a_2 \cdots a_n$ into the left-nested commutator $[\cdots[[a_1, a_2]\,, a_3]\,, \ldots, a_n]$ by progressively "enclosing" more and more factors into commutator brackets. This can be made rigorous as an induction argument as follows:

We shall show that

$$\begin{aligned}
&\left(1 - \mathrm{cyc}_{1,2,\ldots,j}\right)\left(1 - \mathrm{cyc}_{1,2,\ldots,j-1}\right)\cdots\left(1 - \mathrm{cyc}_{1,2}\right) \cdot (a_1 a_2 \cdots a_n) \\
&= \left[\cdots[[a_1, a_2]\,, a_3]\,, \ldots, a_j\right] a_{j+1} a_{j+2} \cdots a_n
\end{aligned} \tag{133}$$

for each $j \in [n]$.

Indeed, we shall prove this by induction on $j$:

*Base case:* For $j = 1$, both sides of (133) equal $a_1 a_2 \cdots a_n$. (For the left hand side, this is because the product $\left(1 - \mathrm{cyc}_{1,2,\ldots,j}\right)\left(1 - \mathrm{cyc}_{1,2,\ldots,j-1}\right)\cdots\left(1 - \mathrm{cyc}_{1,2}\right)$ is empty and thus equals 1. For the right hand side, this is because the left-nested commutator $[\cdots[[a_1, a_2]\,, a_3]\,, \ldots, a_j]$ is defined to be $a_1$ in this case, and therefore the right hand side is $\underbrace{[\cdots[[a_1, a_2]\,, a_3]\,, \ldots, a_j]}_{=a_1}\ \underbrace{a_{j+1} a_{j+2} \cdots a_n}_{=a_2 a_3 \cdots a_n} = a_1 (a_2 a_3 \cdots a_n) = a_1 a_2 \cdots a_n$.) Hence, the equality (133) holds for $j = 1$.

*Induction step:* Let $j \in [n]$ be such that $j > 1$. Assume as the induction hypothesis that (133) holds for $j - 1$ instead of $j$. In other words, assume that

$$\begin{aligned}
&\left(1 - \mathrm{cyc}_{1,2,\ldots,j-1}\right)\left(1 - \mathrm{cyc}_{1,2,\ldots,j-2}\right)\cdots\left(1 - \mathrm{cyc}_{1,2}\right) \cdot (a_1 a_2 \cdots a_n) \\
&= \left[\cdots[[a_1, a_2]\,, a_3]\,, \ldots, a_{j-1}\right] a_j a_{j+1} \cdots a_n.
\end{aligned} \tag{134}$$

We must now prove that (133) holds for $j$ as well. But this follows from

$$\begin{aligned}
&\left(1-\mathrm{cyc}_{1,2,\ldots,j}\right)\left(1-\mathrm{cyc}_{1,2,\ldots,j-1}\right)\cdots\left(1-\mathrm{cyc}_{1,2}\right)\cdot\left(a_1a_2\cdots a_n\right)\\
&=\left(1-\mathrm{cyc}_{1,2,\ldots,j}\right)\cdot\underbrace{\left(1-\mathrm{cyc}_{1,2,\ldots,j-1}\right)\left(1-\mathrm{cyc}_{1,2,\ldots,j-2}\right)\cdots\left(1-\mathrm{cyc}_{1,2}\right)\cdot\left(a_1a_2\cdots a_n\right)}_{=\left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_{j-1}\right]a_ja_{j+1}\cdots a_n\ \text{(by (134))}}\\
&=\left(1-\mathrm{cyc}_{1,2,\ldots,j}\right)\cdot\left(\left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_{j-1}\right]\ \underbrace{a_ja_{j+1}\cdots a_n}_{=a_j\left(a_{j+1}a_{j+2}\cdots a_n\right)}\right)\\
&=\left(1-\mathrm{cyc}_{1,2,\ldots,j}\right)\cdot\left(\left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_{j-1}\right]\ a_j\left(a_{j+1}a_{j+2}\cdots a_n\right)\right)\\
&=\underbrace{\left[\left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_{j-1}\right],a_j\right]}_{=\left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_j\right]}a_{j+1}a_{j+2}\cdots a_n\\
&\qquad\left(\begin{array}{c}\text{by (131), applied to } u=\left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_{j-1}\right],\ g=a_j\\ \text{and } v=a_{j+1}a_{j+2}\cdots a_n \text{ (since a straightforward induction shows}\\ \text{that } \left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_{j-1}\right] \text{ is homogeneous of degree } j-1\end{array}\right)\\
&=\left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_j\right]a_{j+1}a_{j+2}\cdots a_n.
\end{aligned}$$

Thus, the induction step is complete. Hence, (133) is proved by induction.

Now, (128) yields

$$\begin{aligned}
&S\left(\mathbf{V}_n\right)\cdot\left(a_1a_2\cdots a_n\right)\\
&=\left(1-\mathrm{cyc}_{1,2,\ldots,n}\right)\left(1-\mathrm{cyc}_{1,2,\ldots,n-1}\right)\cdots\left(1-\mathrm{cyc}_{1,2}\right)\cdot\left(a_1a_2\cdots a_n\right)\\
&=\left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_n\right]\ \underbrace{a_{n+1}a_{n+2}\cdots a_n}_{=(\text{empty product})=1}\qquad\text{(by (133), applied to } j=n)\\
&=\left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_n\right].
\end{aligned}$$

This proves Proposition 4.5.7. □

Let us denote the action of $S\left(\mathbf{V}_n\right)$ on $\mathbf{NP}_{k,n}$ by $\mathcal{L}_n$. Thus, $\mathcal{L}_n$ is the $\mathbf{k}$-linear map

$$\begin{aligned}\mathbf{NP}_{k,n}&\to\mathbf{NP}_{k,n},\\ p&\mapsto S\left(\mathbf{V}_n\right)\cdot p.\end{aligned}$$

Proposition 4.5.7 tells us that this map $\mathcal{L}_n$ sends each product $a_1a_2\cdots a_n$ (with $a_1,a_2,\ldots,a_n\in\mathbf{NP}_{k,1}$) to $\left[\cdots\left[\left[a_1,a_2\right],a_3\right],\ldots,a_n\right]$. Thus, this map is sometimes called the *left-bracketing operator*. And the claimed equality (129) – which we are trying to prove – becomes

$$\mathcal{L}_n^2=n\mathcal{L}_n. \tag{135}$$

This equality was first proved by Franz Wever in 1947 ([Wever47]) by a rather messy semi-combinatorial argument. Nowadays, the slickest proof of (135) uses Hopf algebras, specifically [GriRei14, Exercise 1.5.14]:

*Proof of (135).* Let $V$ be the free $\mathbf{k}$-module $\mathbf{k}^k$. Let $A$ be its tensor algebra $T(V) = \bigoplus_{d \in \mathbb{N}} V^{\otimes d}$. As we know from Remark 4.5.3, we can identify this tensor algebra $A$ with the noncommutative polynomial ring $\mathbf{NP}_k$. In particular, its degree-1 homogeneous component $V^{\otimes 1} = V$ is thus $\mathbf{NP}_{k,1}$.

But as a tensor algebra, $A$ also has the structure of a cocommutative graded Hopf algebra (see [GriRei14, Examples 1.3.11, 1.3.17 and 1.5.3]). Let us denote its antipode by $S_T$ (note that it is called $S$ in [GriRei14], but is unrelated to our antipode $S$ of $\mathbf{k}[S_n]$). Let $E : A \to A$ be the $\mathbf{k}$-linear map that sends each homogeneous element $a \in A$ to $(\deg a) \cdot a$ (that is, sends each $a \in \mathbf{NP}_{k,d}$ to $da$). Consider the $\mathbf{k}$-linear map $S_T \star E : A \to A$, where the symbol $\star$ denotes convolution (defined in [GriRei14, Definition 1.4.1]).

Now, [GriRei14, Exercise 1.5.14 (e)] shows that

$$(S_T \star E)(a_1 a_2 \cdots a_n) = [\ldots[[a_1, a_2], a_3], \ldots, a_n] \tag{136}$$

for any $a_1, a_2, \ldots, a_n \in V = \mathbf{NP}_{k,1}$. Thus, each $p \in \mathbf{NP}_{k,n}$ satisfies

$$(S_T \star E)(p) = \mathcal{L}_n(p). \tag{137}$$

*Proof of (136).* Both sides of the equality (136) are $\mathbf{k}$-linear in $p$. Thus, in order to prove it, we can WLOG assume that $p$ is a product $a_1 a_2 \cdots a_n$ of $n$ elements $a_1, a_2, \ldots, a_n \in \mathbf{NP}_{k,1}$ (since such products span the $\mathbf{k}$-module $\mathbf{NP}_{k,n}$). Assume this. Then, $p = a_1 a_2 \cdots a_n$, so that

$$\begin{aligned}
(S_T \star E)(p) &= (S_T \star E)(a_1 a_2 \cdots a_n) \\
&= [\ldots[[a_1, a_2], a_3], \ldots, a_n] && \text{(by (136))} \\
&= S(\mathbf{V}_n) \cdot \underbrace{(a_1 a_2 \cdots a_n)}_{=p} && \text{(by Proposition 4.5.7)} \\
&= S(\mathbf{V}_n) \cdot p = \mathcal{L}_n(p) && \text{(by the definition of } \mathcal{L}_n\text{)}.
\end{aligned}$$

This proves (136). □

Now, we recall some properties of the map $S_T \star E$ proved in [GriRei14]. From [GriRei14, Exercise 1.5.14 (a)], it follows that for every $a \in A$,

$$\text{the element } (S_T \star E)(a) \text{ is primitive.} \tag{138}$$

But [GriRei14, Exercise 1.5.14 (b)] yields that

$$(S_T \star E)(p) = E(p) \tag{139}$$

for each primitive element $p$ of $A$.

Now, let $a \in \mathbf{NP}_{k,n}$. Hence, (137) yields $(S_T \star E)(a) = \mathcal{L}_n(a)$. But (138) shows that the element $(S_T \star E)(a)$ is primitive. In other words, $\mathcal{L}_n(a)$ is primitive (since $(S_T \star E)(a) = \mathcal{L}_n(a)$). Hence, (139) (applied to $p = \mathcal{L}_n(a)$) yields

$(S_T \star E)(\mathcal{L}_n(a)) = E(\mathcal{L}_n(a))$. But $\mathcal{L}_n(a)$ belongs to $\mathbf{NP}_{k,n}$ and thus is a homogeneous element of $A$ of degree $n$. Hence, $E(\mathcal{L}_n(a)) = n\mathcal{L}_n(a)$. Furthermore, (137) (applied to $p = \mathcal{L}_n(a)$) yields $(S_T \star E)(\mathcal{L}_n(a)) = \mathcal{L}_n(\mathcal{L}_n(a)) = \mathcal{L}_n^2(a)$. Thus,

$$\mathcal{L}_n^2(a) = (S_T \star E)(\mathcal{L}_n(a)) = E(\mathcal{L}_n(a)) = n\mathcal{L}_n(a).$$

Forget that we fixed $a$. We thus have proved that $\mathcal{L}_n^2(a) = n\mathcal{L}_n(a)$ for each $a \in \mathbf{NP}_{k,n}$. Thus, $\mathcal{L}_n^2 = n\mathcal{L}_n$. This proves (135). □

As we said, we now easily obtain (129), since each $p \in \mathbf{NP}_{k,n}$ satisfies

$$\begin{aligned} S(\mathbf{V}_n)^2 \cdot p &= S(\mathbf{V}_n) \cdot \underbrace{S(\mathbf{V}_n) \cdot p}_{\substack{=\mathcal{L}_n(p) \\ \text{(by the definition of } \mathcal{L}_n\text{)}}} = S(\mathbf{V}_n) \cdot \mathcal{L}_n(p) \\ &= \mathcal{L}_n(\mathcal{L}_n(p)) \qquad \text{(by the definition of } \mathcal{L}_n\text{)} \\ &= \underbrace{\mathcal{L}_n^2}_{\substack{=n\mathcal{L}_n \\ \text{(by (135))}}}(p) = n \underbrace{\mathcal{L}_n(p)}_{\substack{=S(\mathbf{V}_n)\cdot p \\ \text{(by the definition of } \mathcal{L}_n\text{)}}} = nS(\mathbf{V}_n) \cdot p. \end{aligned}$$

So we have proved that (129) holds for each $k \in \mathbb{N}$. By picking $k$ to be $\geq n$, we can thus use the faithfulness of the left $\mathbf{k}[S_n]$-module $\mathbf{NP}_{k,n}$ (Proposition 4.5.6 **(a)**, applied to $d = n$) to conclude that $S(\mathbf{V}_n)^2 = nS(\mathbf{V}_n)$. That is, (127) holds. As we explained, this proves $\mathbf{V}_n^2 = n\mathbf{V}_n$, and thus finishes our proof of Theorem 3.10.10 **(b)**.

# 5. Young and Specht modules

Now we aim to introduce and study a wider class of representations of $S_n$, which will include the irreducible representations over fields of characteristic 0. These are the *Specht modules*, introduced by Wilhelm Specht in 1935 after Alfred Young planted the seeds in his "substitutional analysis" (1901–1931)[48]. There are several equivalent (up to isomorphism) ways to define Specht modules ("avatars" as I will call them), and several levels of generality at which they can be studied (Young diagrams, skew Young diagrams and arbitrary diagrams). I aim to give a maximally broad and general approach, although most of the substantial results hold only for more specialized situations.

[48] Young published his work in a sequence of nine papers, titled "*On Quantitative Substitutional Analysis*" (all reprinted in [Young77]). These papers contain a phenomenal amount of ideas, often concealed behind antiquated notations and a terse Victorian writing style. Many important concepts in this field that are commonly associated with more recent names (Specht modules, von Neumann sandwich lemma, Garnir relations) actually appear in some inchoate form in Young's papers.

## 5.1. Partitions, compositions, diagrams

### 5.1.1. Definitions and examples

To define Specht modules, we will have to define Young modules beforehand. This, in turn, relies on some basic combinatorics that I will now partly recall, partly introduce.

We begin by recalling some combinatorial concepts (see, e.g., [21s, §3.9 and Chapter 4] for more details):

**Definition 5.1.1.** **(a)** An *integer composition* (or, for short, *composition*) means a tuple $(a_1, a_2, \ldots, a_k)$ of positive integers.

**(b)** The *size* $|a|$ of a composition $a = (a_1, a_2, \ldots, a_k)$ is defined to be $a_1 + a_2 + \cdots + a_k \in \mathbb{N}$.

**(c)** The *length* $\ell(a)$ of a composition $a = (a_1, a_2, \ldots, a_k)$ is defined to be $k$.

**(d)** Given $n \in \mathbb{Z}$, a *composition of* $n$ means a composition with size $n$.

**(e)** An *integer partition* (or, for short, *partition*) means an integer composition $(a_1, a_2, \ldots, a_k)$ that is weakly decreasing (i.e., satisfies $a_1 \geq a_2 \geq \cdots \geq a_k$).

**(f)** Given $n \in \mathbb{Z}$, a *partition of* $n$ means a partition with size $n$.

**(g)** We will use the notation $a_i$ for the $i$-th entry of a composition $a$. (This applies, in particular, to all partitions $a$.)

**(h)** If $a$ is a partition, then we furthermore set $a_i := 0$ for all $i > \ell(a)$. That is, we pretend that a partition consists not only of its original entries, but also contains an infinite row of zeroes following them. Thus, a partition $a$ satisfies not only the finite chain of inequalities $a_1 \geq a_2 \geq \cdots \geq a_{\ell(a)}$, but also the infinite chain $a_1 \geq a_2 \geq a_3 \geq \cdots$.

We thus identify each partition $a$ with the infinite sequence $(a_1, a_2, a_3, \ldots)$. From this point of view, a partition can also be defined as an infinite weakly decreasing sequence of nonnegative integers that contains only finitely many positive integers.

**(i)** Two partitions $a$ and $b$ are said to satisfy $a \supseteq b$ (or, equivalently, $b \subseteq a$) if every positive integer $i$ satisfies $a_i \geq b_i$. (It suffices to check this for all $i \leq \ell(b)$, because all $i > \ell(b)$ satisfy $b_i = 0 \leq a_i$ and thus automatically $a_i \geq b_i$. But it does not suffice to check this for all $i \leq \ell(a)$.)

If two partitions $a$ and $b$ satisfy $a \supseteq b$, then the pair $(b, a)$ is called a *skew partition* and denoted by $a/b$.

**(j)** We denote the empty partition $() = (0, 0, 0, \ldots)$ by $\varnothing$.

**(k)** The notation "$\lambda \vdash n$" means "$\lambda$ is a partition of $n$".

For instance:

- The composition $(2,1,2)$ has size 5 and length 3.
- There are four compositions of 3, namely $(3)$, $(2,1)$, $(1,2)$ and $(1,1,1)$. Only three of them are partitions (indeed, $(1,2)$ is not).
- The composition $(2,2,1)$ is a partition of 5. Thus, we can write $(2,2,1) \vdash 5$.
- The partition $(4,2,2,1)$ is identified with the infinite sequence $(4,2,2,1,0,0,0,\ldots)$.
- The partitions $(3,1)$ and $(4,2,1)$ satisfy $(3,1) \subseteq (4,2,1)$ (since $3 \leq 4$ and $1 \leq 2$). Thus, the pair $((3,1),(4,2,1))$ is a skew partition, and is denoted by $(4,2,1) \,/\, (3,1)$.

  The partitions $(2,2,1)$ and $(3,1,1)$ do not satisfy $(2,2,1) \subseteq (3,1,1)$, since their second entries do not satisfy $2 \leq 1$.

  The partitions $(3,1)$ and $(4)$ do not satisfy $(3,1) \subseteq (4)$, since their second entries do not satisfy $1 \leq 0$ (remember that the second entry of $(4)$ is 0).

Next, we define some "two-dimensional" combinatorial objects:

**Definition 5.1.2. (a)** The *English coordinate system* is a Cartesian coordinate system on the plane, in which the x-axis goes top-down and the y-axis goes left-to-right. (This is exactly the way the entries of a matrix are indexed: The rows are counted from the top; the columns are counted from the left.)

(In contrast, the usual Cartesian coordinate system has its x-axis go left-to-right and its y-axis go bottom-to-top.)

We will be using the English system throughout this topic.

**(b)** A *cell* (or *box* or *square*) shall mean a pair $(i,j)$ of integers. Thus, the set of all cells is $\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z}$.

We draw each cell $(i,j)$ as a $1 \times 1$-square in the plane, centered at the point $(i,j)$ in the English coordinate system. For now we leave these cells empty; soon we will put numbers into them.

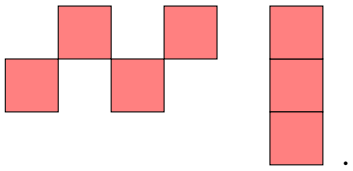
**(c)** We will use the obvious geographical directions (left, right, above, below, north, east, south, west, northwest, etc.) in reference to the English coordinate system. For example, a cell $(i,j)$ has the westward neighbor $(i,j-1)$, the northward neighbor $(i,j-1)$, the eastward neighbor $(i,j+1)$ and the southward neighbor $(i+1,j)$, and we can also refer to these four neighbors as the "left neighbor", the "upward neighbor", the "right neighbor" and the "downward neighbor", respectively.

We say that a cell $(i,j)$ lies *weakly northwest* of a cell $(u,v)$ if $i \leq u$ and $j \leq v$. In this case, we write $(i,j) \leq (u,v)$. Similarly, we say that a cell $(i,j)$ lies *weakly northeast* of a cell $(u,v)$ if $i \leq u$ and $j \geq v$. (But we do not have a symbol for this.) Generally, the words "weakly" and "strictly" will typically refer to weak and strict inequalities (so that a cell lies weakly north of itself but not strictly north of itself, etc.).

**(d)** A *diagram* (or *shape*) will mean a finite set of cells. For instance, the set

$$\{(1,2),\ (2,1),\ (2,3),\ (1,4),\ (1,6),\ (2,6),\ (3,6)\}$$

is a diagram. In the English coordinate system, it is drawn as follows:



(Beware that, if the axes are not shown, such a drawing determines the diagram only up to parallel translation. But this will usually suffice for us.)

We always draw cells and diagrams in the English coordinate system. This is called the *English notation* for diagrams. (There is an alternative notation used in parts of the literature, known as *French notation*, which has the x-axis go bottom-to-top rather than top-to-bottom.)

The *cells* of a diagram $D$ mean the cells that belong to $D$.

**(e)** The *rows* and the *columns* are defined in the way you would expect:

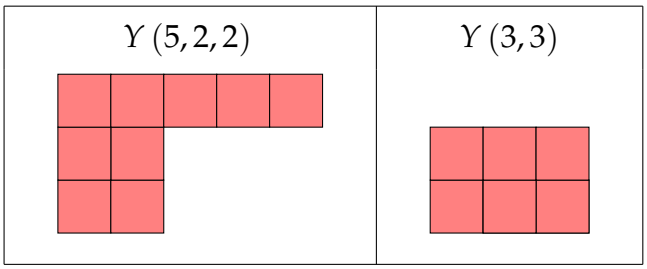
- For any given $i \in \mathbb{Z}$, we define the *$i$-th row* (aka *row $i$*) to be the set of all cells of the form $(i,j)$ for $j \in \mathbb{Z}$.
- For any diagram $D$, the *$i$-th row of $D$* (aka *row $i$ of $D$*) is the set of all such cells in $D$ (that is, the intersection of $D$ with the $i$-th row).
- For any given $j \in \mathbb{Z}$, we define the *$j$-th column* (aka *column $j$*) to be the set of all cells of the form $(i,j)$ for $i \in \mathbb{Z}$.
- For any diagram $D$, the *$j$-th column of $D$* (aka *column $j$ of $D$*) is the set of all such cells in $D$ (that is, the intersection of $D$ with the $j$-th column).

**(f)** The *Young diagram* $Y(a)$ of a partition $a = (a_1, a_2, \ldots, a_k)$ is defined to be the diagram

$$\begin{aligned}
&\left\{(i,j) \in \{1,2,3,\ldots\}^2 \ \mid\ j \leq a_i\right\} \\
&= \{(i,j) \in \{1,2,\ldots,k\} \times \{1,2,3,\ldots\} \ \mid\ j \leq a_i\} \\
&\qquad\qquad \left(\begin{array}{c}\text{since } j \leq a_i \text{ can happen only if } a_i > 0, \\ \text{which entails that } i \in \{1,2,\ldots,k\}\end{array}\right) \\
&= \bigcup_{i=1}^{k} (\{i\} \times [a_i]) \qquad \left(\begin{array}{c}\text{since a positive integer } j \in \{1,2,3,\ldots\} \\ \text{satisfies } j \leq a_i \text{ if and only if } j \in [a_i]\end{array}\right) \\
&= \bigcup_{i \geq 1} (\{i\} \times [a_i]) \qquad (\text{since } \{i\} \times [a_i] \text{ is empty whenever } i > k).
\end{aligned}$$

Visually, this is a collection of cells that forms a "table with $k$ left-aligned rows", where row $i$ has $a_i$ cells for each $i \in [k]$.

For example, the Young diagrams $Y(5,2,2)$ and $Y(3,3)$ of the partitions $(5,2,2)$ and $(3,3)$ are drawn below:

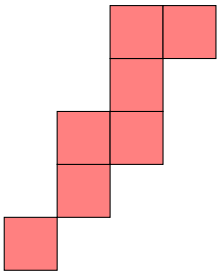



.

**(g)** The *skew Young diagram* $Y(a/b)$ of a skew partition $a/b$ is defined to be the diagram

$$Y(a) \setminus Y(b) = \left\{ (i,j) \in \{1,2,3,\ldots\}^2 \mid b_i < j \leq a_i \right\}.$$

For example, the skew Young diagram $Y((4,3,3,2,1) / (2,2,1,1))$ is drawn below:



.

As the above definitions of $Y(a)$ and $Y(a/b)$ suggest, we will mostly use cells $(i,j)$ whose coordinates $i$ and $j$ are positive, although we do not formally require this.

We note that whenever $a$ is a partition, the pair $a/\varnothing = (\varnothing, a)$ is a skew partition (i.e., we have $\varnothing \subseteq a$), and its skew Young diagram $Y(a/\varnothing)$ is just the usual Young diagram $Y(a)$. Thus, skew Young diagrams generalize Young diagrams.

Here is a little table of Young diagrams:

- The only partition of 0 is the empty partition $\varnothing = ()$, and its Young diagram $Y(\varnothing)$ is empty (i.e., has no cells).
- The only partition of 1 is $(1)$, and its Young diagram $Y(1)$ is a single cell: .

- The partitions of 2 are $(2)$ and $(1,1)$, and their Young diagrams are

$$Y(2) = \ydiagram{2} \qquad \text{and} \qquad Y(1,1) = \ydiagram{1,1}.$$

- Here are the partitions of 3 and their Young diagrams:

| $\lambda$ | $(3)$ | $(2,1)$ | $(1,1,1)$ |
|---|---|---|---|
| $Y(\lambda)$ | $\ydiagram{3}$ | $\ydiagram{2,1}$ | $\ydiagram{1,1,1}$ |

- Here are the partitions of 4 and their Young diagrams (shrunk to fit the page):

| $\lambda$ | $(4)$ | $(3,1)$ | $(2,2)$ | $(2,1,1)$ | $(1,1,1,1)$ |
|---|---|---|---|---|---|
| $Y(\lambda)$ | $\ydiagram{4}$ | $\ydiagram{3,1}$ | $\ydiagram{2,2}$ | $\ydiagram{2,1,1}$ | $\ydiagram{1,1,1,1}$ |

### 5.1.2. Basic properties

Let us prove some basic properties of Young diagrams as a warmup. First, let us convince ourselves that the relation $\subseteq$ on partitions deserves its "subset" symbol:

**Proposition 5.1.3.** Let $a$ and $b$ be two partitions. Then, $b \subseteq a$ holds if and only if $Y(b) \subseteq Y(a)$.

*Proof.* Recall that the definition of $Y(a)$ says that

$$Y(a) = \left\{ (i,j) \in \{1,2,3,\ldots\}^2 \mid j \leq a_i \right\}. \tag{140}$$

Similarly,

$$Y(b) = \left\{ (i,j) \in \{1,2,3,\ldots\}^2 \mid j \leq b_i \right\}. \tag{141}$$

We need to prove the equivalence $(b \subseteq a) \iff (Y(b) \subseteq Y(a))$. We shall prove its "$\Longrightarrow$" and "$\Longleftarrow$" implications separately:

$\Longrightarrow$: Assume that $b \subseteq a$. We must prove that $Y(b) \subseteq Y(a)$.

Let $(i,j) \in Y(b)$ be arbitrary. By (141), this entails $i \geq 1$ and $j \leq b_i$, so that $j \leq b_i \leq a_i$ (since $b \subseteq a$ means that $b_i \leq a_i$ for each $i \geq 1$). Hence, $i \geq 1$ and $j \leq a_i$. This, in turn, shows that $(i,j) \in Y(a)$ (by (140)).

Forget that we fixed $(i,j)$. We thus have shown that $(i,j) \in Y(a)$ for each $(i,j) \in Y(b)$. In other words, $Y(b) \subseteq Y(a)$. This proves the "$\Longrightarrow$" direction of the equivalence $(b \subseteq a) \iff (Y(b) \subseteq Y(a))$.

$\Longleftarrow$: Assume that $Y(b) \subseteq Y(a)$. We must prove that $b \subseteq a$. In other words, we must prove that $b_i \leq a_i$ for each $i \geq 1$.

Let $i \geq 1$. We must prove that $b_i \leq a_i$. Assume the contrary. Thus, $b_i > a_i \geq 0$, so that $b_i \in \{1,2,3,\ldots\}$. By (141), this entails that $(i,b_i) \in Y(b)$ (since $b_i \leq b_i$). Hence, $(i,b_i) \in Y(b) \subseteq Y(a)$. By (140), this shows that $b_i \leq a_i$, which contradicts $b_i > a_i$. This contradiction shows that our assumption was false. Hence, the proof of the "$\Longleftarrow$" direction of the equivalence $(b \subseteq a) \iff (Y(b) \subseteq Y(a))$ is complete.

We have now proved both directions of the equivalence; thus, the equivalence (i.e., Proposition 5.1.3) holds. □

Next, we show that the size (i.e., number of cells) of the Young diagram $Y(\lambda)$ of a partition $\lambda$ is just the size $|\lambda|$ of $\lambda$:

**Proposition 5.1.4.** Let $\lambda$ be any partition. Then, its Young diagram $Y(\lambda)$ has size $|Y(\lambda)| = |\lambda|$.

*Proof.* Write $\lambda$ as $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$. Then, the definition of the size $|\lambda|$ shows that $|\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_k$.

However, the definition of the Young diagram $Y(\lambda)$ yields $Y(\lambda) = \bigcup_{i=1}^{k} (\{i\} \times [\lambda_i])$. Thus,

$$\begin{aligned}
|Y(\lambda)| &= \left| \bigcup_{i=1}^{k} (\{i\} \times [\lambda_i]) \right| \\
&= \sum_{i=1}^{k} \underbrace{|(\{i\} \times [\lambda_i])|}_{=|\{i\}| \cdot |[\lambda_i]|} \qquad \left( \begin{array}{c} \text{by the sum rule, since the sets } \{i\} \times [\lambda_i] \text{ for} \\ \text{all } i \in [k] \text{ are distinct (because any element} \\ \text{of } \{i\} \times [\lambda_i] \text{ is a pair whose first entry is } i) \end{array} \right) \\
&= \sum_{i=1}^{k} \underbrace{|\{i\}|}_{=1} \cdot \underbrace{|[\lambda_i]|}_{=\lambda_i} = \sum_{i=1}^{k} \lambda_i = \lambda_1 + \lambda_2 + \cdots + \lambda_k = |\lambda|
\end{aligned}$$

(since $|\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_k$). This proves Proposition 5.1.4. □

**Proposition 5.1.5.** Let $\lambda$ be any partition. Then, its Young diagram $Y(\lambda)$ has the following property: If a cell $c \in \{1,2,3,\ldots\}^2$ lies weakly northwest of some cell $d \in Y(\lambda)$, then we have $c \in Y(\lambda)$ as well.

*Proof.* Let $c \in \{1,2,3,\ldots\}^2$ be a cell that lies weakly northwest of some cell $d \in Y(\lambda)$. We must show that $c \in Y(\lambda)$.

The definition of $Y(\lambda)$ yields that

$$Y(\lambda) = \left\{(i,j) \in \{1,2,3,\ldots\}^2 \mid j \leq \lambda_i\right\}. \tag{142}$$

Write the cells $c$ and $d$ as $c = (i,j)$ and $d = (u,v)$. Since $c$ lies weakly northwest of $d$, we thus have $i \leq u$ and $j \leq v$. Moreover, $i$ and $j$ are positive integers (since $(i,j) = c \in \{1,2,3,\ldots\}^2$).

But $(u,v) = d \in Y(\lambda)$. Hence, $u, v \in \{1,2,3,\ldots\}$ and $v \leq \lambda_u$ (by (142)).

However, the sequence $(\lambda_1, \lambda_2, \lambda_3, \ldots)$ is weakly decreasing (since $\lambda$ is a partition). Thus, from $i \leq u$, we obtain $\lambda_i \geq \lambda_u$. Now, $j \leq v \leq \lambda_u \leq \lambda_i$ (since $\lambda_i \geq \lambda_u$). Since $i$ and $j$ are positive integers, this entails that $(i,j) \in Y(\lambda)$ (by (142)). In other words, $c \in Y(\lambda)$ (since $c = (i,j)$). This proves Proposition 5.1.5. □

**Proposition 5.1.6.** Let $\lambda/\mu$ be any skew partition. Then, its skew Young diagram $Y(\lambda/\mu)$ has the following property: If $c, d, e \in \mathbb{Z}^2$ are three cells satisfying $c \leq d \leq e$ and $c, e \in Y(\lambda/\mu)$, then $d \in Y(\lambda/\mu)$.

*Proof.* Let $c, d, e \in \mathbb{Z}^2$ be three cells satisfying $c \leq d \leq e$ and $c, e \in Y(\lambda/\mu)$. We must show that $d \in Y(\lambda/\mu)$.

We have $Y(\lambda/\mu) = Y(\lambda) \setminus Y(\mu)$ (by the definition of a skew Young diagram). Now, $c \in Y(\lambda/\mu) = Y(\lambda) \setminus Y(\mu)$, so that $c \notin Y(\mu)$. Furthermore, $e \in Y(\lambda/\mu) = Y(\lambda) \setminus Y(\mu)$, so that $e \in Y(\lambda)$.

Moreover, $c \in Y(\lambda) \setminus Y(\mu) \subseteq Y(\lambda) \subseteq \{1,2,3,\ldots\}^2$. Thus, because of $c \leq d$, it is easy to see that $d \in \{1,2,3,\ldots\}^2$ as well[49].

The cell $c$ lies weakly northwest of $d$ (since $c \leq d$). Hence, if we had $d \in Y(\mu)$, then we would have $c \in Y(\mu)$ (by Proposition 5.1.5, applied to $\mu$ instead of $\lambda$), which would contradict $c \notin Y(\mu)$. Thus, we cannot have $d \in Y(\mu)$. In other words, $d \notin Y(\mu)$.

The cell $d$ lies weakly northwest of $e$ (since $d \leq e$). Hence, from $e \in Y(\lambda)$, we obtain $d \in Y(\lambda)$ (by Proposition 5.1.5, applied to $d$ and $e$ instead of $c$ and $d$).

Combining $d \in Y(\lambda)$ with $d \notin Y(\mu)$, we obtain $d \in Y(\lambda) \setminus Y(\mu) = Y(\lambda/\mu)$. This proves Proposition 5.1.6. □

**Proposition 5.1.7.** Let $D \subseteq \{1,2,3,\ldots\}^2$ be a diagram. Then:

**(a)** The diagram $D$ is the Young diagram $Y(\lambda)$ of some partition $\lambda$ if and only if it has the following property: If a cell $c \in \{1,2,3,\ldots\}^2$ lies weakly northwest of a cell $d \in D$, then we have $c \in D$ as well.

**(b)** The diagram $D$ is the skew Young diagram $Y(\lambda/\mu)$ of some skew partition $\lambda/\mu$ if and only if it has the following property: If $c, d, e \in \{1,2,3,\ldots\}^2$ are three cells satisfying $c \leq d \leq e$ and $c, e \in D$, then $d \in D$.

**(c)** A partition $\lambda$ is uniquely determined by its Young diagram $Y(\lambda)$.

[49] *Proof.* Write the cells $c$ and $d$ as $c = (i,j)$ and $d = (i',j')$. Then, $(i,j) = c \leq d = (i',j')$; in other words, $i \leq i'$ and $j \leq j'$. But $(i,j) = c \in \{1,2,3,\ldots\}^2$, so that $i \geq 1$ and $j \geq 1$. Thus, from $i \leq i'$, we obtain $i' \geq i \geq 1$ and thus $i' \in \{1,2,3,\ldots\}$, and similarly $j' \in \{1,2,3,\ldots\}$. Hence, $(i',j') \in \{1,2,3,\ldots\}^2$, so that $d = (i',j') \in \{1,2,3,\ldots\}^2$.

**(d)** A skew partition $\lambda/\mu$, however, is not always uniquely determined by its skew Young diagram $Y(\lambda/\mu)$.

*Partial proof.* **(a)** $\Longrightarrow$: This follows from Proposition 5.1.5.

$\Longleftarrow$: Assume that $D$ has the following property: If a cell $c \in \{1,2,3,\ldots\}^2$ lies weakly northwest of a cell $d \in D$, then we have $c \in D$ as well. We shall refer to this property as the "northwest-closure property".

We must show that $D$ is the Young diagram $Y(\lambda)$ of some partition $\lambda$.

If $D$ is empty, then this is clear (since the empty diagram is $Y(\varnothing)$ for the empty partition $\varnothing = ()$). Thus, we WLOG assume that $D$ is nonempty.

Let $(p,q)$ be a southernmost cell in $D$ (that is, a cell with maximum $p$). This is well-defined, since $D$ is finite and nonempty. By the northwest-closure property, it follows that all $p$ cells $(1,q)$, $(2,q)$, $\ldots$, $(p,q)$ belong to $D$ as well (since they lie weakly northwest of the cell $(p,q) \in D$). Hence, the diagram $D$ has cells in each of rows $1,2,\ldots,p$ (namely, the $p$ cells we just mentioned). Furthermore, $D$ has no cells south of row $p$ (since $(p,q)$ is a southernmost cell in $D$).

For each $i \in [p]$, let $(i,m_i)$ be the easternmost cell in row $i$ of $D$ (that is, the cell in row $i$ of $D$ having maximum $m_i$). This is well-defined (since $D$ is finite and has at least one cell in row $i$ (because the diagram $D$ has cells in each of rows $1,2,\ldots,p$)). Thus, we have defined $p$ positive integers $m_1, m_2, \ldots, m_p$, which form a composition $(m_1, m_2, \ldots, m_p)$.

We next claim that this composition $(m_1, m_2, \ldots, m_p)$ is a partition.

[*Proof:* Let $i \in \{2,3,\ldots,p\}$. We shall prove that $m_{i-1} \geq m_i$. Indeed, $(i,m_i)$ is a cell in $D$ (by the definition of $m_i$). Hence, its northward neighbor $(i-1,m_i)$ also belongs to $D$ (by the northwest-closure property, since $i-1 \geq 1$). But $(i-1,m_{i-1})$ is defined as the easternmost cell in row $i-1$ of $D$. Thus, each cell in row $i-1$ of $D$ must lie weakly west of this cell $(i-1,m_{i-1})$. Applying this to the cell $(i-1,m_i)$ (which clearly lies in row $i-1$ and belongs to $D$), we conclude that $(i-1,m_i)$ lies weakly west of the cell $(i-1,m_{i-1})$. In other words, $m_i \leq m_{i-1}$. In other words, $m_{i-1} \geq m_i$.

Forget that we fixed $i$. We thus have shown that $m_{i-1} \geq m_i$ for each $i \in \{2,3,\ldots,p\}$. Combining these inequalities, we arrive at $m_1 \geq m_2 \geq \cdots \geq m_p$. In other words, the composition $(m_1, m_2, \ldots, m_p)$ is a partition.]

The definition of the Young diagram $Y(m_1, m_2, \ldots, m_p)$ yields

$$Y(m_1, m_2, \ldots, m_p) = \bigcup_{i=1}^{p} (\{i\} \times [m_i]). \tag{143}$$

Let us compare this with the diagram $D$. Let $i \in [p]$. Then, $(i,m_i)$ is the easternmost cell in row $i$ of $D$. Hence, $(i,m_i) \in D$. Therefore, all $m_i$ cells $(i,1)$, $(i,2)$, $\ldots$, $(i,m_i)$ must belong to $D$ (by the northwest-closure property, since they lie weakly west of $(i,m_i)$). No other cells in the $i$-th row can belong to $D$, since $(i,m_i)$ is the easternmost cell in this row. Thus, in total, the $i$-th row of $D$ consists of the cells $(i,1)$, $(i,2)$, $\ldots$, $(i,m_i)$ and no others. In other words,

$$\begin{aligned}(\text{the } i\text{-th row of } D) &= \{(i,1), (i,2), \ldots, (i,m_i)\} \\ &= \{(i,j) \mid j \in [m_i]\} = \{i\} \times [m_i].\end{aligned}$$

Forget that we fixed $i$. We thus have shown that

$$(\text{the } i\text{-th row of } D) = \{i\} \times [m_i] \tag{144}$$

for each $i \in [p]$. Moreover, recall that $D$ has no cells south of row $p$. Hence,

$$D = \bigcup_{i=1}^{p} \underbrace{(\text{the } i\text{-th row of } D)}_{\substack{=\{i\}\times[m_i] \\ \text{(by (144))}}} = \bigcup_{i=1}^{p} (\{i\} \times [m_i]) = Y\left(m_1, m_2, \ldots, m_p\right)$$

(by (143)). Hence, $D$ is the Young diagram $Y(\lambda)$ of some partition $\lambda$ (namely, of $\lambda = \left(m_1, m_2, \ldots, m_p\right)$). This proves the "$\Longleftarrow$" direction of Proposition 5.1.7 **(a)**.

**(b)** $\Longrightarrow$: This follows from Proposition 5.1.6.
$\Longleftarrow$: See Exercise 5.1.1 below.

**(c)** It is easy to see that for any partition $\lambda$ and any $i \geq 1$, we have

$$\lambda_i = (\#\text{ of cells in the } i\text{-th row of } Y(\lambda))$$

(indeed, the cells in the $i$-th row of $Y(\lambda)$ are $(i,1)$, $(i,2)$, $\ldots$, $(i,\lambda_i)$). Thus, each $\lambda_i$ and therefore the whole partition $\lambda$ can be uniquely reconstructed from $Y(\lambda)$. This proves Proposition 5.1.7 **(c)**.

**(d)** For a counterexample, observe that the skew partitions $(1) / (1)$ and $\varnothing/\varnothing$ have the same skew Young diagram (namely, the empty set). More generally, for any partition $\lambda$, the skew partition $\lambda/\lambda$ is well-defined and its skew Young diagram $Y(\lambda/\lambda)$ is the empty set.

Nonempty counterexamples exist as well: For instance,

$$Y((3,1,1) / (2,1)) = Y((3,2,1) / (2,2)) = \{(1,3), (3,1)\}.$$

□

**Exercise 5.1.1.** [3] Prove the "$\Longleftarrow$" direction of Proposition 5.1.7 **(b)**.

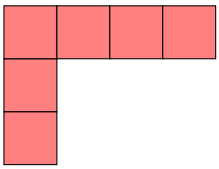
### 5.1.3. Some notations regarding partitions

**Definition 5.1.8.** When writing partitions or compositions, we will often use *exponential notation*, which means that we will abbreviate a sequence $\underbrace{i, i, \ldots, i}_{k \text{ times}}$ of equal entries by writing $i^k$. To avoid confusion between this notation and an actual power, we shall never put powers in a partition or composition without warning.

For example, $\left(4, 2^3, 1\right)$ means the partition $(4,2,2,2,1)$.

**Definition 5.1.9.** A *hook partition* means a partition of the form $\left(i, 1^j\right)$ for some positive integer $i$ and some nonnegative integer $j$. (If $j = 0$, then this just means $(i)$. If $i = 1$, then this just means $\left(1^{j+1}\right)$.)

Hook partitions owe their name to the fact that their Young diagrams look like hooks: For example, the Young diagram $Y\left(4, 1^2\right) = Y(4,1,1)$ is



.

### 5.1.4. Conjugate partitions

We next introduce a fundamental transformation on the set of partitions, called *conjugation* or *transposition*:

**Theorem 5.1.10.** Let $\mathbf{r} : \mathbb{Z}^2 \to \mathbb{Z}^2$ be the map that sends each cell $(i, j) \in \mathbb{Z}^2$ to $(j, i)$. (In the English coordinate system, this map $\mathbf{r}$ is just the reflection across the northwest-to-southeast diagonal.)

Let $\lambda$ be a partition. Then:

**(a)** There exists a unique partition $\mu$ such that $Y(\mu) = \mathbf{r}(Y(\lambda))$.

This partition $\mu$ is called the *conjugate* (or the *transpose*, or the *conjugate partition*) of $\lambda$. It is denoted by $\lambda^t$ (or, by many authors, by $\lambda'$).

**(b)** This partition $\lambda^t$ is explicitly given by the formula

$$\begin{aligned}\lambda_j^t &= (\#\text{ of positive integers } i \text{ such that } \lambda_i \geq j) \\ &= (\#\text{ of cells in the } j\text{-th column of } Y(\lambda)) \qquad \text{for all } j \geq 1.\end{aligned}$$

**(c)** This partition $\lambda^t$ satisfies $\left|\lambda^t\right| = |\lambda|$ and $\lambda_1^t = \ell(\lambda)$ and $\ell\left(\lambda^t\right) = \lambda_1$.

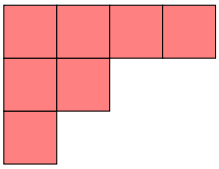
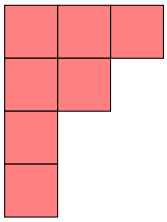
**(d)** The conjugate $\left(\lambda^t\right)^t$ of this partition $\lambda^t$ is $\lambda$.

**(e)** Let $j$ be a positive integer. Then, the cells in the $j$-th column of $Y(\lambda)$ are

$$(1, j), (2, j), \ldots, \left(\lambda_j^t, j\right).$$

**Example 5.1.11.** If $\lambda = (4,2,1)$, then the conjugate of $\lambda$ is $\lambda^t = (3,2,1,1)$.

Here are the Young diagrams of these two partitions:

$Y(\lambda) =$  and $Y\left(\lambda^t\right) =$  .

(In the future, we will often be sloppy and just write $\lambda$ for $Y(\lambda)$.)

*Proof of Theorem 5.1.10 (sketched).* **(a)** The uniqueness of the partition $\mu$ follows from Proposition 5.1.7 **(c)**. It remains to prove its existence. In other words, it remains to prove that $\mathbf{r}(Y(\lambda))$ is the Young diagram of some partition.

By Proposition 5.1.7 **(a)**, this will follow if we can show that $\mathbf{r}(Y(\lambda))$ has the following property: If a cell $c \in \{1,2,3,\ldots\}^2$ lies weakly northwest of a cell $d \in \mathbf{r}(Y(\lambda))$, then we have $c \in \mathbf{r}(Y(\lambda))$ as well.

But this property of $\mathbf{r}(Y(\lambda))$ follows easily from the analogous property of $Y(\lambda)$ (which is true by Proposition 5.1.5), since the reflection $\mathbf{r}$ preserves the "lying weakly northwest of" relation. Thus, the existence of $\mu$ is proved, and the proof of Theorem 5.1.10 **(a)** is complete.

**(b)** The definition of $\lambda^t$ yields $Y\left(\lambda^t\right) = \mathbf{r}(Y(\lambda))$. Thus, each row of $Y\left(\lambda^t\right)$ is the reflection of the corresponding column of $Y(\lambda)$ across the northwest-to-southeast diagonal (since the reflection $\mathbf{r}$ turns columns into rows). In particular, the length of each row of $Y\left(\lambda^t\right)$ equals the length of the corresponding column of $Y(\lambda)$. In other words, for each $j \geq 1$, we have

$$\begin{aligned}&\left(\text{the length of the } j\text{-th row of } Y\left(\lambda^t\right)\right)\\ &= \left(\text{the length of the } j\text{-th column of } Y(\lambda)\right)\\ &= \left(\#\text{ of cells in the } j\text{-th column of } Y(\lambda)\right).\end{aligned}$$

But the left hand side of this equality is clearly $\lambda^t_j$, since the length of the $j$-th row of any Young diagram $Y(\mu)$ is $\mu_j$. Thus, this equality can be rewritten as

$$\begin{aligned}\lambda^t_j &= \left(\#\text{ of cells in the } j\text{-th column of } Y(\lambda)\right)\\ &= \left(\#\text{ of positive integers } i \text{ such that } \lambda_i \geq j\right)\end{aligned}$$

(since the cells in the $j$-th column of $Y(\lambda)$ are the cells $(i,j)$ for all positive integers $i$ satisfying $\lambda_i \geq j$). This proves Theorem 5.1.10 **(b)**.

**(d)** This follows from the fact that the reflection $\mathbf{r}$ is an involution (i.e., satisfies $\mathbf{r} \circ \mathbf{r} = \mathrm{id}$).

**(c)** Proposition 5.1.4 yields $|Y(\lambda)| = |\lambda|$ and $\left|Y\left(\lambda^t\right)\right| = \left|\lambda^t\right|$. However, the definition of $\lambda^t$ yields $Y\left(\lambda^t\right) = \mathbf{r}(Y(\lambda))$, so that $\left|Y\left(\lambda^t\right)\right| = |Y(\lambda)|$ (since the reflection $\mathbf{r}$ is a bijection and thus preserves the size of a set). This rewrites as $\left|\lambda^t\right| = |\lambda|$ (since $|Y(\lambda)| = |\lambda|$ and $\left|Y\left(\lambda^t\right)\right| = \left|\lambda^t\right|$).

Theorem 5.1.10 **(b)** yields

$$\lambda_1^t = (\#\text{ of positive integers } i \text{ such that } \lambda_i \geq 1) = \ell(\lambda)$$

(because if $\ell(\lambda) = k$, then $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$, and the positive integers $i$ such that $\lambda_i \geq 1$ are precisely the numbers $1, 2, \ldots, k$).

Thus we have proved the equality $\lambda_1^t = \ell(\lambda)$. Applying this equality to $\lambda^t$ instead of $\lambda$, we obtain $\left(\lambda^t\right)_1^t = \ell\left(\lambda^t\right)$. Thus, $\ell\left(\lambda^t\right) = \left(\lambda^t\right)_1^t = \lambda_1$ (since Theorem 5.1.10 **(d)** yields $\left(\lambda^t\right)^t = \lambda$). With this equality, the proof of Theorem 5.1.10 **(c)** is complete.

**(e)** The reflection $\mathbf{r}$ is an involution, i.e., satisfies $\mathbf{r} \circ \mathbf{r} = \mathrm{id}$. Hence, $\mathbf{r}$ is invertible, with inverse $\mathbf{r}^{-1} = \mathbf{r}$.

As we have seen in the proof of Theorem 5.1.10 **(b)** above, each row of $Y\left(\lambda^t\right)$ is the reflection of the corresponding column of $Y(\lambda)$ across the northwest-to-southeast diagonal. Thus, the $j$-th row of $Y\left(\lambda^t\right)$ is the reflection of the $j$-th column of $Y(\lambda)$ across this diagonal. In other words,

$$\left(\text{the } j\text{-th row of } Y\left(\lambda^t\right)\right) = \mathbf{r}\left(\text{the } j\text{-th column of } Y(\lambda)\right).$$

Since $\mathbf{r}$ is invertible, this entails that

$$\begin{aligned}
\left(\text{the } j\text{-th column of } Y(\lambda)\right) &= \underbrace{\mathbf{r}^{-1}}_{=\mathbf{r}} \underbrace{\left(\text{the } j\text{-th row of } Y\left(\lambda^t\right)\right)}_{=\left\{(j,1),\ (j,2),\ \ldots,\ \left(j,\lambda_j^t\right)\right\}} \\
&= \mathbf{r}\left(\left\{(j,1),\ (j,2),\ \ldots,\ \left(j,\lambda_j^t\right)\right\}\right) \\
&= \left\{\mathbf{r}(j,1),\ \mathbf{r}(j,2),\ \ldots,\ \mathbf{r}\left(j,\lambda_j^t\right)\right\} \\
&= \left\{(1,j),\ (2,j),\ \ldots,\ \left(\lambda_j^t, j\right)\right\}
\end{aligned}$$

(since $\mathbf{r}(u,v) = (v,u)$ for any cell $(u,v) \in \mathbb{Z}^2$). In other words, the cells in the $j$-th column of $Y(\lambda)$ are $(1,j),\ (2,j),\ \ldots,\ \left(\lambda_j^t, j\right)$. This proves Theorem 5.1.10 **(e)**. □

See [21s, §4.1.5] for a basic application of conjugate partitions.

**Remark 5.1.12.** Let $\lambda/\mu$ be a skew partition. Then, $\lambda^t/\mu^t$ is again a skew partition, and we have $Y\left(\lambda^t/\mu^t\right) = \mathbf{r}\left(Y(\lambda/\mu)\right)$.

*Proof.* We know that $\lambda/\mu$ is a skew partition, i.e., a pair $(\mu, \lambda)$ of two partitions satisfying $\lambda \supseteq \mu$ (by the definition of a skew partition). We have $\lambda \supseteq \mu$; in other words, $\mu \subseteq \lambda$. Hence, Proposition 5.1.3 (applied to $a = \lambda$ and $b = \mu$) shows that $Y(\mu) \subseteq Y(\lambda)$. Hence, $\mathbf{r}(Y(\mu)) \subseteq \mathbf{r}(Y(\lambda))$.

But the definition of $\lambda^t$ shows that $Y\left(\lambda^t\right) = \mathbf{r}(Y(\lambda))$. Similarly, $Y\left(\mu^t\right) = \mathbf{r}(Y(\mu))$. Thus, $Y\left(\mu^t\right) = \mathbf{r}(Y(\mu)) \subseteq \mathbf{r}(Y(\lambda)) = Y\left(\lambda^t\right)$. By Proposition 5.1.3 (applied to $a = \lambda^t$ and $b = \mu^t$), we thus conclude that $\mu^t \subseteq \lambda^t$. Thus, $\lambda^t/\mu^t$ is a skew partition.

We have $Y(\lambda/\mu) = Y(\lambda) \setminus Y(\mu)$ (by the definition of a skew Young diagram). Hence,

$$\mathbf{r}(Y(\lambda/\mu)) = \mathbf{r}(Y(\lambda) \setminus Y(\mu)) = \mathbf{r}(Y(\lambda)) \setminus \mathbf{r}(Y(\mu))$$

(since $\mathbf{r}$ is a bijection, and thus respects set differences). Comparing this with

$$\begin{aligned} Y\left(\lambda^t/\mu^t\right) &= \underbrace{Y\left(\lambda^t\right)}_{=\mathbf{r}(Y(\lambda))} \setminus \underbrace{Y\left(\mu^t\right)}_{=\mathbf{r}(Y(\mu))} \qquad \left(\begin{array}{c}\text{by the definition of}\\ \text{a skew Young diagram}\end{array}\right) \\ &= \mathbf{r}\left(Y\left(\lambda\right)\right) \setminus \mathbf{r}\left(Y\left(\mu\right)\right), \end{aligned}$$

we obtain $Y\left(\lambda^t/\mu^t\right) = \mathbf{r}\left(Y\left(\lambda/\mu\right)\right)$. The proof of Remark 5.1.12 is thus complete. □

## 5.2. Tableaux

### 5.2.1. Definitions and examples

While Young diagrams have some uses visualizing partitions (see, e.g., [21s, Proof of Proposition 4.1.15] for a nice application), to us they are mainly important as tables to be filled with numbers. Once filled with numbers, they are called *tableaux*. (This is a loanword from French; its singular form is "tableau".) Let us define tableaux accurately and rigorously:

**Definition 5.2.1.** Let $D$ be a diagram.

**(a)** A *tableau* of shape $D$ (aka a *filling* of $D$) means a map $T : D \to \{1,2,3,\ldots\}$ that assigns a positive integer $T\left(c\right)$ to each cell $c \in D$. In this case, $T\left(c\right)$ is called the *entry* of $T$ in cell $c$, and we say that the cell $c$ of $T$ *contains* the number $T\left(c\right)$. We visualize a tableau $T$ by placing the entry $T\left(c\right)$ into the cell $c$ for each cell $c$.

For example, here is a tableau of shape $Y\left(\left(5,4,4\right)/\left(2,1\right)\right)$ visualized in this way:

| | | | | |
|---|---|---|---|---|
| | | 2 | 3 | 1 |
| | 1 | 3 | 1 | |
| 5 | 3 | 5 | 1 | |

.

Rigorously speaking, it is a map $T : Y\left(\left(5,4,4\right)/\left(2,1\right)\right) \to \{1,2,3,\ldots\}$ that sends the cell $\left(1,3\right)$ to 2, the cell $\left(1,4\right)$ to 3 and so on.

**(b)** A tableau $T$ is said to be an *$m$-tableau* for some $m \in \mathbb{N}$ if it is injective (i.e., has no two equal entries) and its image is $[m]$. (This means that the entries of $T$ are the numbers $1,2,\ldots,m$, with each of these numbers appearing exactly once. Thus, an $m$-tableau of shape $D$ can be viewed as a bijection from $D$ to $[m]$.)

Obviously, there are $m!$ many $m$-tableaux of shape $D$ if $|D| = m$, and otherwise there are none.

**(c)** If $T$ is a tableau of shape $D$, and $i$ is a positive integer, then the *$i$-th row* of $T$ shall mean the restriction of $T$ to the $i$-th row of $D$. This restriction contains the entries of $T$ that lie in the $i$-th row. Similarly, the *$j$-th column* of $T$

is defined. The *cells* of a tableau $T$ are just defined to be the cells of its shape $D$.

**(d)** A tableau $T$ is said to be

- *row-standard* if its entries strictly increase (left to right) along each row.
- *row-semistandard* if its entries weakly increase (left to right) along each row.
- *column-standard* if its entries strictly increase (top to bottom) down each column.
- *column-semistandard* if its entries weakly increase (top to bottom) down each column.
- *standard* if it is row-standard and column-standard and an $m$-tableau for some $m$.
- *semistandard* if it is row-semistandard and column-standard (not column-semistandard!).
- *straight-shaped* (or of *straight shape*, or just *straight*) if it has shape $Y(\lambda)$ for a partition $\lambda$;
- *skew-shaped* (or of *skew shape*, or just *skew*) if it has shape $Y(\lambda/\mu)$ for a skew partition $\lambda/\mu$ (this includes the straight-shaped case, since $Y(\lambda/\varnothing) = Y(\lambda)$).
- *bad-shaped* if it is not skew-shaped.

We often abbreviate "tableau of shape $Y(\lambda)$" (where $\lambda$ is a partition) as "tableau of shape $\lambda$", and likewise we abbreviate "tableau of shape $Y(\lambda/\mu)$" (where $\lambda/\mu$ is a skew partition) as "tableau of shape $\lambda/\mu$".

**Remark 5.2.2.** Tableaux are also known as *Young tableaux*. Depending on the author, this name can come with extra requirements, such as standard or semistandard. Different books do not agree on what exactly they require! (For instance, Fulton's [Fulton97] defines "tableaux" and "Young tableaux" to mean what we call straight semistandard tableaux, whereas all other tableaux are just called "fillings".)

**Example 5.2.3.** There are only 2 standard tableaux of shape $Y(2,2)$, namely

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array} \quad \text{and} \quad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \hline \end{array} \ .$$

There are 5 standard tableaux of shape $Y(3,2)$, namely

$$\begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5\\\cline{1-2}\end{array}\quad \begin{array}{|c|c|c|}\hline 1&2&4\\\hline 3&5\\\cline{1-2}\end{array}\quad \begin{array}{|c|c|c|}\hline 1&2&5\\\hline 3&4\\\cline{1-2}\end{array}\quad \begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2&5\\\cline{1-2}\end{array}\quad \begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4\\\cline{1-2}\end{array}\ .$$

We cannot list the other types of tableaux, since there are usually infinitely many of them. But here are some examples:

- The tableau $\begin{array}{|c|c|c|}\cline{2-3} &2&3\\\hline 1&3&3\\\cline{1-3}\end{array}$ is row-semistandard, column-semistandard and skew (having shape $Y((3,3)\,/\,(1))$), but neither row-standard (since 3 appears twice in the second row) nor column-standard (since 3 appears twice in the third column) nor straight nor semistandard.

- The tableau $\begin{array}{|c|}\hline 1\\\cline{1-1}\end{array}\quad \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}4&6\\\cline{1-2}\end{array}$ is row-standard, column-standard and bad-shaped. It is nevertheless not standard, since it is not an $m$-tableau for any $m \in \mathbb{N}$ (indeed, it contains the entry 1 twice and the entry 5 never).

- The tableau $\begin{array}{|c|c|c|}\cline{2-3} &1&3\\\hline 2&2&4\\\hline 3&5&5\\\cline{1-3}\end{array}$ is semistandard and skew.

For users of LaTeX: the `ytableau` package provides an easy way to typeset tableaux of various kinds. See its documentation for several ways it can be used.

For users of SageMath: There is a lot of infrastructure for working with straight-shaped and skew-shaped tableaux in SageMath (not so much for bad-shaped tableaux). For instance, typing in `list(StandardTableaux([4,2,1]))` will return all standard tableaux of shape $Y(4,2,1)$. For another instance, `SkewTableau([[None,2,3],[1,4]])` returns the skew (standard) tableau $\begin{array}{|c|c|c|}\cline{2-3} &2&3\\\hline 1&4\\\cline{1-2}\end{array}$.

**Convention 5.2.4.** Sometimes we will use a shorthand notation for tableaux (particularly when tableaux will appear as subscripts), in which we write all rows of a tableau in a single line (from top to bottom), separated by "\\" signs, and with each row being written as the list of its entries (with the "∘" sign signifying a missing cell). For example, the tableau

| | | |
|---|---|---|
| | 1 | 3 |
| 2 | 2 | 4 |
| 3 | 5 | 5 |

will thus be written as

$$\circ, 1, 3 \backslash\backslash 2, 2, 4 \backslash\backslash 3, 5, 5.$$

Often we will also omit the commas between the entries (as long as the entries are single-digit numbers). Thus, the above tableau will be written

$$\circ 13 \backslash\backslash 224 \backslash\backslash 355.$$

### 5.2.2. First properties

Obviously, all straight tableaux are skew as well (since $Y(\lambda) = Y(\lambda/\varnothing)$ for any partition $\lambda$). Thus, the following basic property of skew tableaux applies to straight ones as well:

**Proposition 5.2.5.** Let $T$ be a skew tableau of some shape $D$. Let $c$ and $e$ be two cells of $D$ such that $c$ lies weakly northwest of $e$. Then:

**(a)** If $T$ is row-semistandard and column-semistandard, then $T(c) \leq T(e)$.

**(b)** If $T$ is semistandard, and if $c$ and $e$ lie in different rows, then $T(c) < T(e)$.

*Proof.* The tableau $T$ is skew; thus, its shape $D$ has the form $Y(\lambda/\mu)$ for some skew partition $\lambda/\mu$. Consider this $\lambda/\mu$.

Write the cells $c$ and $e$ in the forms $c = (i, j)$ and $e = (u, v)$. Since $c$ lies weakly northwest of $e$, we thus have $i \leq u$ and $j \leq v$.

Let $d$ be the cell $(i, v)$. This cell $d$ lies in the same row as $c = (i, j)$ but weakly east of $c$ (since $v \geq j$). It furthermore lies in the same column as $e = (u, v)$ but weakly north of $e$ (since $i \leq u$). In particular, we have $c \leq d \leq e$. Hence, Proposition 5.1.6 yields $d \in Y(\lambda/\mu)$ (since $c \in D = Y(\lambda/\mu)$ and $e \in D = Y(\lambda/\mu)$). In other words, $d \in D$ (since $D = Y(\lambda/\mu)$). Hence, the entry $T(d)$ exists.

**(a)** Assume that $T$ is row-semistandard and column-semistandard.

Since $T$ is row-semistandard, the entries in the $i$-th row of $T$ weakly increase left to right. Hence, $T(c) \leq T(d)$ (since the cells $c = (i,j)$ and $d = (i,v)$ both lie in the $i$-th row, and $d$ lies weakly east of $c$).

Since $T$ is column-semistandard, the entries in the $v$-th column of $T$ weakly increase top to bottom. Hence, $T(d) \leq T(e)$ (since the cells $d = (i,v)$ and $e = (u,v)$ both lie in the $v$-th column, and $d$ lies weakly north of $e$).

Altogether, we conclude that $T(c) \leq T(d) \leq T(e)$. This proves Proposition 5.2.5 **(a)**.

**(b)** Assume that $T$ is semistandard. Thus, $T$ is row-semistandard and column-standard. Hence, as in the proof of part **(b)**, we can show that $T(c) \leq T(d)$.

Now, let us additionally assume that $c$ and $e$ lie in different rows. In other words, $i \neq u$ (since $c = (i,j)$ lies in row $i$, whereas $e = (u,v)$ lies in row $u$). Hence, $i < u$ (since $i \leq u$). Thus, the cell $d = (i,v)$ lies strictly (not just weakly) north of $e = (u,v)$.

Since $T$ is column-semistandard, the entries in the $v$-th column of $T$ strictly increase top to bottom. Hence, $T(d) < T(e)$ (since the cells $d = (i,v)$ and $e = (u,v)$ both lie in the $v$-th column, and $d$ lies strictly north of $e$).

Altogether, we conclude that $T(c) \leq T(d) < T(e)$. This proves Proposition 5.2.5 **(b)**. □

Even more obvious is the following property:

**Proposition 5.2.6.** Let $D$ be any diagram with $|D| = n$. Then, the standard tableaux of shape $D$ are precisely the standard $n$-tableaux of shape $D$.

*Proof.* Let $T$ be a standard tableau of shape $D$. We shall show that $T$ is an $n$-tableau.

Indeed, $T$ is standard, and thus is an $m$-tableau for some $m \in \mathbb{N}$. Consider this $m$. Then, $T$ is a bijection from $D$ to $[m]$ (since any $m$-tableau of shape $D$ can be viewed as a bijection from $D$ to $[m]$). Hence, $|D| = |[m]| = m$. Thus, $m = |D| = n$. But $T$ is an $m$-tableau. In other words, $T$ is an $n$-tableau (since $m = n$). Thus, $T$ is a standard $n$-tableau of shape $D$.

Forget that we fixed $T$. We thus have shown that each standard tableau $T$ of shape $D$ is a standard $n$-tableau of shape $D$. Conversely, of course, any standard $n$-tableau of shape $D$ is a standard tableau of shape $D$. Combining these two facts, we conclude that the standard tableaux of shape $D$ are precisely the standard $n$-tableaux of shape $D$. This proves Proposition 5.2.6. □

The following "local" criterion for row- and column-standardness is an important advantage of skew-shaped tableaux (it does not hold for bad-shaped tableaux):

**Proposition 5.2.7.** Let $D$ be a skew Young diagram. Let $T$ be a tableau of shape $D$. Then:

**(a)** If every $(i,j) \in D$ satisfying $(i,j+1) \in D$ satisfies $T(i,j) < T(i,j+1)$, then $T$ is row-standard.

**(b)** If every $(i,j) \in D$ satisfying $(i+1,j) \in D$ satisfies $T(i,j) < T(i+1,j)$, then $T$ is column-standard.

In words, Proposition 5.2.7 **(a)** is saying that in order to check that a skew-shaped tableau is row-standard, we only need to check that its entries in two **horizontally adjacent** cells (i.e., two cells $(i,j)$ and $(i,j+1)$) are always in the right order (i.e., the left one is always smaller than the right one). Intuitively speaking, this is because the rows of $D$ are contiguous blocks of cells, so you can walk from every cell to any other (in the same row) without having to jump over any missing cells. Proposition 5.2.7 **(b)** is analogous, just replacing rows by columns. For the sake of completeness, we give a rigorous proof of both parts of Proposition 5.2.7 in the Appendix (Section A.1).

### 5.2.3. The hook length formula

Now here is something a lot less obvious:

**Theorem 5.2.8** (Hook length formula)**.** Let $\lambda$ be any partition with $|\lambda| = n$. Then, the # of standard tableaux of shape $Y(\lambda)$ is

$$\frac{n!}{\prod_{c \in Y(\lambda)} |H_\lambda(c)|}.$$

Here, for each cell $c \in Y(\lambda)$, we define the *hook* $H_\lambda(c)$ of this cell in $\lambda$ to be

$$H_\lambda(c) := \{c\} \cup \{\text{all cells in } Y(\lambda) \text{ that lie east of } c \text{ (in the same row as } c)\} \cup \{\text{all cells in } Y(\lambda) \text{ that lie south of } c \text{ (in the same column as } c)\}.$$

**Example 5.2.9.** Let $n = 5$ and $\lambda = (3,2)$. This partition $\lambda$ has Young diagram

$$Y(\lambda) = \begin{array}{|c|c|c|}\hline \phantom{x} & \phantom{x} & \phantom{x} \\ \hline \phantom{x} & \phantom{x} \\ \cline{1-2}\end{array} = \{(1,1),\ (1,2),\ (1,3),\ (2,1),\ (2,2)\}.$$

The hooks of its cells are

$$\begin{aligned} H_\lambda(1,1) &= \{(1,1)\} \cup \{(1,2),\ (1,3)\} \cup \{(2,1)\}, \\ H_\lambda(1,2) &= \{(1,2)\} \cup \{(1,3)\} \cup \{(2,2)\}, \\ H_\lambda(1,3) &= \{(1,3)\} \cup \varnothing \cup \varnothing, \\ H_\lambda(2,1) &= \{(2,1)\} \cup \{(2,2)\} \cup \varnothing, \\ H_\lambda(2,2) &= \{(2,2)\} \cup \varnothing \cup \varnothing, \end{aligned}$$

with respective sizes 4, 3, 1, 2 and 1. Thus, the hook length formula (Theorem 5.2.8) yields that the # of standard tableaux of shape $Y(\lambda)$ is

$$\frac{n!}{\prod\limits_{c\in Y(\lambda)} |H_\lambda(c)|} = \frac{5!}{4\cdot 3\cdot 1\cdot 2\cdot 1} = 5,$$

just as we found in Example 5.2.3.

The hook length formula is a surprising and nontrivial result (in general, it is far from obvious that $\dfrac{n!}{\prod\limits_{c\in Y(\lambda)} |H_\lambda(c)|}$ is an integer!), and provides some first evidence that standard tableaux have a deeper meaning. We will prove it in Section 5.24. Other (usually simpler and shorter) proofs of Theorem 5.2.8 can also be found in [Aigner07, Chapter 8, Highlight], [Bona22, Theorem 6.5], [CSScTo10, Theorem 4.2.14], [EGHetc11, Theorem 5.17.1], [Fulton97, §4.3, Exercise 10], [Hopkin22, Remark 19], [James78, §20], [JamKer81, 2.3.21], [Kerber99, 6.3.8], [Knuth98, §5.1.4, Theorem H], [Krishn86, §V.3], [LakBro18, Theorem 5.4.10], [Loehr11, §12.10, Theorem 12.50], [Lorenz18, §4.3.5], [Macdon95, §I.5, Example 2], [Meliot17, Theorem 3.41], [MenRem15, (5.7)], [Prasad15, Theorem 5.8.3], [Proces07, (5.2.1)], [Romik15, Theorem 1.12], [Sagan01, Theorem 3.10.2], [Sagan19, Theorem 7.3.1], [Stanle23, Corollary 7.21.6], [Stanle18, Theorem 8.1], [Zelevi81b, §6.5, Proposition (b)], and myriad other sources. See [Pak22, §11.2] for a taxonomy of different proofs and references to each type.

**Exercise 5.2.1.** $\boxed{2}$ Prove Theorem 5.2.8 by hand in the case when $\lambda$ is a hook partition.

The following exercise applies Theorem 5.2.8 to Young diagrams with two rows (i.e., diagrams of the form $Y(\lambda)$ where $\lambda = (\lambda_1, \lambda_2)$):

**Exercise 5.2.2.** **(a)** $\boxed{1}$ Use Theorem 5.2.8 to show that for each $m \in \mathbb{N}$, the # of standard tableaux of shape $Y(m,m)$ is $\dfrac{1}{m+1}\dbinom{2m}{m}$. This number is known as the $m$-th Catalan number (see, e.g., [21s, §3.1.2] or [22fco, Lecture 29, §5.3]).

**(b)** $\boxed{2}$ Prove the same claim without using Theorem 5.2.8, by instead constructing a bijection between the standard tableaux of shape $Y(m,m)$ and the Catalan lattice paths from $(0,0)$ to $(m,m)$ (see [22fco, Lecture 29, §5.2]).

**(c)** $\boxed{3}$ More generally, compute the # of standard tableaux of shape $Y(a,b)$ for all $a, b \in \mathbb{N}$ with $a \geq b$ (that is, of shape $Y(\lambda)$ for any partition $\lambda$ of length $\leq 2$) both using Theorem 5.2.8 and without Theorem 5.2.8.

[**Hint:** For part **(c)**, use [22fco, Theorem 5.2.4 **(d)**].]

In the form we presented it, Theorem 5.2.8 was first stated by Frame, Robinson and Thrall in 1954 ([FrRoTh54, Theorem 1]), who proved it by deriving it from an older (1902) formula due to Frobenius and Young:

**Theorem 5.2.10** (Frobenius–Young formula)**.** Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ be any partition with $|\lambda| = n$. Let $\ell_i := \lambda_i + k - i$ for each $i \in [k]$. Then, the # of standard tableaux of shape $Y(\lambda)$ is

$$\frac{n!}{\ell_1! \cdot \ell_2! \cdot \cdots \cdot \ell_k!} \prod_{1 \leq i < j \leq k} (\ell_i - \ell_j).$$

**Exercise 5.2.3.** [3] Prove that Theorem 5.2.8 and Theorem 5.2.10 are equivalent, by showing that

$$\frac{1}{\prod\limits_{c \in Y(\lambda)} |H_\lambda(c)|} = \frac{1}{\ell_1! \cdot \ell_2! \cdot \cdots \cdot \ell_k!} \prod_{1 \leq i < j \leq k} (\ell_i - \ell_j)$$

(where all the notations are as in Theorem 5.2.10).

Another nice and useful property of skew tableaux is the *non-messing-up theorem*, which we give in two variants as an exercise:

**Exercise 5.2.4.** **(a)** [2] Let $T$ be a skew row-semistandard tableau.

Let $S$ be the tableau obtained from $T$ by sorting each column in weakly increasing order. (This is supposed to mean that for each $j \in \mathbb{Z}$, the $j$-th column of $S$ contains the same entries as the $j$-th column of $T$ but rearranged in weakly increasing order from top to bottom. The shapes of $S$ and $T$ are the same, and no entries get moved out of their column.)

Prove that $S$ is row-semistandard (in addition to obviously being column-semistandard).

**(b)** [1] Prove that the same is true if we replace "row-semistandard" by "row-standard" (both for $T$ and for $S$).

**(c)** [1] Prove that neither part **(a)** nor part **(b)** remains true if the tableaux are allowed to be bad-shape.

However, don't get too encouraged:

**Exercise 5.2.5.** [1] If $T$ is an $n$-tableau, then $T^{\mathrm{row}}$ shall denote the tableau obtained from $T$ by sorting each row in increasing order, whereas $T^{\mathrm{col}}$ shall denote the tableau obtained from $T$ by sorting each column in increasing order.

Find an example of a straight-shaped $n$-tableau $T$ such that $\left(T^{\mathrm{row}}\right)^{\mathrm{col}} \neq \left(T^{\mathrm{col}}\right)^{\mathrm{row}}$.

## 5.3. Young modules

### 5.3.1. Horizontal and vertical permutations

Let us now get symmetric groups into the action.

**Definition 5.3.1.** Let $T$ be an $n$-tableau (of any shape $D$). Thus, each $i \in [n]$ lies in exactly one cell of $T$, so that we can talk about "the cell of $T$ that contains $i$". (Rigorously speaking, this cell is simply $T^{-1}(i)$.)

**(a)** A permutation $w \in S_n$ is said to be *horizontal* for $T$ if it has the property that for each $i \in [n]$, the number $w(i)$ lies in the same row of $T$ as $i$ does.

The *row group* $\mathcal{R}(T)$ of $T$ is the subgroup of $S_n$ consisting of all permutations $w \in S_n$ that are horizontal for $T$.

**(b)** A permutation $w \in S_n$ is said to be *vertical* for $T$ if it has the property that for each $i \in [n]$, the number $w(i)$ lies in the same column of $T$ as $i$ does.

The *column group* $\mathcal{C}(T)$ of $T$ is the subgroup of $S_n$ consisting of all permutations $w \in S_n$ that are vertical for $T$.

In other words, a permutation $w \in S_n$ is *horizontal* for an $n$-tableau $T$ if it permutes the set of entries of each row[50]. Likewise, a permutation $w \in S_n$ is *vertical* for an $n$-tableau $T$ if it permutes the set of entries of each column.

For example, for $T = \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 & 5 \\ \cline{1-2}\end{array}$, we have

$$\mathcal{R}(T) = \{\text{all permutations in } S_5 \text{ permuting } \{1,3,4\} \text{ and permuting } \{2,5\}\}$$

(a subgroup of $S_5$ having size $3! \cdot 2! = 12$) and

$$\mathcal{C}(T) = \{\text{all permutations in } S_5 \text{ permuting } \{1,2\} \text{ and permuting } \{3,5\} \text{ and permuting } \{4\}\}$$

(a subgroup of $S_5$ having size $2! \cdot 2! \cdot 1! = 4$; its elements are id, $t_{1,2}$, $t_{3,5}$ and $t_{1,2}t_{3,5}$).

### 5.3.2. Row-equivalence and column-equivalence

**Definition 5.3.2. (a)** Two tableaux $S$ and $T$ are said to be *row-equivalent* if they have the same entries in corresponding rows (i.e., if for every $i \in \mathbb{Z}$, the multiset of entries in the $i$-th row of $S$ is the multiset of entries in the $i$-th row of $T$).

**(b)** Two tableaux $S$ and $T$ are said to be *column-equivalent* if they have the same entries in corresponding columns (i.e., if for every $j \in \mathbb{Z}$, the multiset of entries in the $j$-th column of $S$ is the multiset of entries in the $j$-th column of $T$).

[50] A permutation $w \in S_n$ is said to *permute* some subset $I$ of $[n]$ if it sends $I$ to itself (i.e., if $w(i) \in I$ for each $i \in I$).

**Example 5.3.3. (a)** The tableaux $\begin{array}{|c|c|c|}\hline 1&1&4\\\hline 2&3\\\cline{1-2}\end{array}$ and $\begin{array}{|c|c|c|}\hline 1&4&1\\\hline 3&2\\\cline{1-2}\end{array}$ are row-equivalent, but they are not row-equivalent to $\begin{array}{|c|c|c|}\hline 1&4&4\\\hline 2&3\\\cline{1-2}\end{array}$ (since multisets are not equal if they have different multiplicities).

**(b)** The tableaux $\begin{array}{|c|c|c|}\hline 1&1&4\\\hline 2&3\\\cline{1-2}\end{array}$ and $\begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2&1\\\cline{1-2}\end{array}$ are column-equivalent, but they are not column-equivalent to $\begin{array}{|c|c|c|}\hline 1&1&4\\\hline 3&2\\\cline{1-2}\end{array}$.

Note that tableaux of different shapes can be row-equivalent! For instance, the tableaux $\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}$ and $\begin{array}{|c|c|}\hline 1&2\\\hline \multicolumn{1}{c}{}&3\\\cline{2-2}\end{array}$ are row-equivalent. That said, we will usually not compare tableaux of different shapes in what follows.

Obviously, row-equivalence is an equivalence relation, and so is column-equivalence. The following is obvious as well:

**Proposition 5.3.4. (a)** If two $n$-tableaux $S$ and $T$ are row-equivalent, then $\mathcal{R}(S) = \mathcal{R}(T)$.

**(b)** If two $n$-tableaux $S$ and $T$ are column-equivalent, then $\mathcal{C}(S) = \mathcal{C}(T)$.

Note that the converse of either statement is false; for example, the two 4-tableaux $S = \begin{array}{|c|c|}\hline 1&2\\\hline 3&4\\\cline{1-2}\end{array}$ and $T = \begin{array}{|c|c|}\hline 3&4\\\hline 1&2\\\cline{1-2}\end{array}$ are not row-equivalent but satisfy $\mathcal{R}(S) = \mathcal{R}(T)$.

Here is a property of row and column groups that is only slightly less obvious than Proposition 5.3.4:

**Proposition 5.3.5.** Let $T$ be an $n$-tableau of any shape $D$. Then:

**(a)** We have $\mathcal{R}(T) \cap \mathcal{C}(T) = \{\mathrm{id}\}$.

**(b)** The only pair $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$ satisfying $cr = \mathrm{id}$ is the pair $(\mathrm{id}, \mathrm{id})$.

*Proof.* **(a)** Let $w \in \mathcal{R}(T) \cap \mathcal{C}(T)$. Thus, $w \in \mathcal{R}(T)$ and $w \in \mathcal{C}(T)$.

Let $i \in [n]$. The tableau $T$ is an $n$-tableau, and thus contains each of the integers $1, 2, \ldots, n$ exactly once. Hence, there is a unique cell $c$ of $T$ that contains the number $i$, and there is a unique cell $d$ of $T$ that contains the number $w(i)$. Consider these two cells $c$ and $d$. Thus, $T(c) = i$ (since the cell $c$ of $T$ contains the number $i$) and $T(d) = w(i)$ (similarly).

From $w \in \mathcal{R}(T)$, we see that the permutation $w$ is horizontal for $T$ (by the definition of $\mathcal{R}(T)$). Hence, the number $w(i)$ lies in the same row of $T$ as $i$ does (by the definition

of "horizontal"). In other words, the cell of $T$ that contains the number $i$ lies in the same row as the cell of $T$ that contains the number $w(i)$. Since the former cell is $c$, and since the latter cell is $d$, we can rewrite this as follows: The cell $c$ lies in the same row as the cell $d$.

Similarly, from $w \in \mathcal{C}(T)$, we can see that the cell $c$ lies in the same column as the cell $d$.

However, if two cells lie in the same row and in the same column, then they are simply the same cell. Thus, $c$ and $d$ are the same cell (since $c$ lies in the same row and in the same column as $d$). In other words, $c = d$. Hence, $T(c) = T(d) = w(i)$, so that $w(i) = T(c) = i$.

Forget that we fixed $i$. We thus have proved that $w(i) = i$ for each $i \in [n]$. Thus, $w = \mathrm{id}$.

Forget that we fixed $w$. We have now shown that $w = \mathrm{id}$ for each $w \in \mathcal{R}(T) \cap \mathcal{C}(T)$. In other words, $\mathcal{R}(T) \cap \mathcal{C}(T) \subseteq \{\mathrm{id}\}$. Combining this with $\{\mathrm{id}\} \subseteq \mathcal{R}(T) \cap \mathcal{C}(T)$ (which is obvious, since the permutation id is both horizontal and vertical for $T$), we obtain $\mathcal{R}(T) \cap \mathcal{C}(T) = \{\mathrm{id}\}$. Proposition 5.3.5 **(a)** is now proved.

**(b)** Clearly, $(\mathrm{id}, \mathrm{id})$ is a pair $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$ satisfying $cr = \mathrm{id}$. It remains to show that there are no other such pairs. So let $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$ be a pair that satisfies $cr = \mathrm{id}$. We must prove that $(c, r) = (\mathrm{id}, \mathrm{id})$.

From $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$, we obtain $c \in \mathcal{C}(T)$ and $r \in \mathcal{R}(T)$. From $cr = \mathrm{id}$, we obtain $c = r^{-1} \in \mathcal{R}(T)$ (since $r \in \mathcal{R}(T)$, but $\mathcal{R}(T)$ is a group). Combining this with $c \in \mathcal{C}(T)$, we obtain $c \in \mathcal{R}(T) \cap \mathcal{C}(T) = \{\mathrm{id}\}$ (by Proposition 5.3.5 **(a)**), so that $c = \mathrm{id}$. Hence, the equality $cr = \mathrm{id}$ rewrites as $\mathrm{id}\, r = \mathrm{id}$, thus $r = \mathrm{id}$. Hence, $(c, r) = (\mathrm{id}, \mathrm{id})$ (since $c = \mathrm{id}$ and $r = \mathrm{id}$).

Forget that we fixed $(c, r)$. We thus have shown that if $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$ is a pair that satisfies $cr = \mathrm{id}$, then $(c, r) = (\mathrm{id}, \mathrm{id})$. Hence, the only such pair is $(\mathrm{id}, \mathrm{id})$ (since we already know that $(\mathrm{id}, \mathrm{id})$ is such a pair). This proves Proposition 5.3.5 **(b)**. □

### 5.3.3. Tabloids

**Definition 5.3.6. (a)** Let $D$ be a diagram. An *$n$-tabloid* of shape $D$ is an equivalence class of $n$-tableaux of shape $D$ with respect to row-equivalence. For example, for $D = Y(2,1)$ and $n = 3$, the set $\{12\backslash\backslash 3,\ 21\backslash\backslash 3\}$ (where we are using the shorthand notation from Convention 5.2.4) is an $n$-tabloid of shape $D$.

**(b)** We use the notation $\overline{T}$ for the $n$-tabloid corresponding to an $n$-tableau $T$ (that is, the equivalence class containing $T$). For instance, the 3-tabloid $\{12\backslash\backslash 3,\ 21\backslash\backslash 3\}$ can thus be written as $\overline{12\backslash\backslash 3}$ or (equivalently) as $\overline{21\backslash\backslash 3}$.

To draw an $n$-tabloid, we just draw one of the $n$-tableaux in it (often we pick the row-standard one), and we drop the vertical subdividers between the entries of a row. For instance, the 3-tabloid $\overline{12\backslash\backslash 3}$ is thus drawn as $\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}$.

**(c)** If $\overline{T}$ is an $n$-tabloid, and if $i \in \mathbb{Z}$, then the *$i$-th row of $\overline{T}$* shall mean the set of the entries of the $i$-th row of $T$. This is well-defined, since two row-

equivalent tableaux will have the same entries in their $i$-th row. For example, the 2-nd row of the 5-tabloid $\begin{array}{|cc|}\hline 2 & 4 \\ \hline 1 & 5 \\ \hline 3 \\ \cline{1-1}\end{array}$ is the set $\{1,5\}$.

**Example 5.3.7. (a)** There are four 4-tabloids of shape $Y(3,1)$, namely

$$\begin{array}{|ccc|}\hline 1 & 2 & 3 \\ \hline 4 \\ \cline{1-1}\end{array}, \quad \begin{array}{|ccc|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1}\end{array}, \quad \begin{array}{|ccc|}\hline 1 & 3 & 4 \\ \hline 2 \\ \cline{1-1}\end{array}, \quad \begin{array}{|ccc|}\hline 2 & 3 & 4 \\ \hline 1 \\ \cline{1-1}\end{array}.$$

**(b)** More generally: Assume that $n \geq 2$. What are the $n$-tabloids of shape $Y(n-1,1)$ ?

Such an $n$-tabloid is uniquely determined by its only entry in the 2-nd row. Indeed, if its entry in the 2-nd row is $i$, then its 1-st row must contain the remaining numbers $1, 2, \ldots, i-1, i+1, i+2, \ldots, n$, and the order in which they appear is immaterial (since row-equivalent tableaux produce the same tabloid). The entry $i$ in the 2-nd row can be any $i \in [n]$, and we get a different $n$-tabloid for each $i$. Thus, in total, there are $n$ many $n$-tabloids of shape $Y(n-1,1)$.

**Exercise 5.3.1.** Let $k \in [n-1]$.

**(a)** $\boxed{1}$ What are the $n$-tabloids of shape $Y\left(n-k,1^k\right)$ ? How many are there?

**(b)** $\boxed{1}$ Assume that $k \leq n/2$. What are the $n$-tabloids of shape $Y(n-k,k)$ ? How many are there?

While $n$-tabloids are equivalence classes, they also can be represented in a canonical way, namely by row-standard $n$-tableaux:

**Proposition 5.3.8.** Let $D$ be any diagram. Then, there is a bijection

$$\begin{aligned} &\text{from } \{\text{row-standard } n\text{-tableaux of shape } D\} \\ &\text{to } \{n\text{-tabloids of shape } D\} \end{aligned}$$

that sends each row-standard $n$-tableau $T$ to its tabloid $\overline{T}$.

*Proof.* The map from $\{$row-standard $n$-tableaux of shape $D\}$ to $\{n$-tabloids of shape $D\}$ that sends each $T$ to $\overline{T}$ is clearly well-defined. It remains to show that this map is a bijection. In other words, it remains to show that each $n$-tabloid of shape $D$ can be written as $\overline{T}$ for a unique row-standard $n$-tableau $T$ of shape $D$. In other words, we need to show that each $n$-tabloid of shape $D$ has a unique row-standard representative[51].

[51] A *representative* of an equivalence class just means an element of this class.

But this is easy: We can pick any representative of our $n$-tabloid, and then sort each row in increasing order (from left to right) so that the representative becomes row-standard. This row-standard representative is furthermore unique, since the row-standardness requirement forces the entries in each row to appear in increasing order and therefore binds each entry to a unique position. Thus, we have shown that our $n$-tabloid has a unique row-standard representative. As explained above, this completes the proof of Proposition 5.3.8. □

### 5.3.4. The symmetric group acts on tableaux and tabloids

Now, let us define an action of the symmetric group $S_n$ on sets of tableaux:

**Definition 5.3.9.** Let $D$ be any diagram. Recall that an $n$-tableau of shape $D$ is defined as a map from $D$ to $\{1,2,3,\ldots\}$ that is injective and whose image is $[n]$. Thus, an $n$-tableau of shape $D$ can be viewed as a bijection from $D$ to $[n]$. Hence, if $T$ is an $n$-tableau of shape $D$, and if $w \in S_n$ is a permutation of $[n]$, then the composition $w \circ T$ is a bijection from $D$ to $[n]$, that is, an $n$-tableau of shape $D$ again.

Now, we define a left $S_n$-action on the set $\{n\text{-tableaux of shape } D\}$ by setting

$$w \rightharpoonup T := w \circ T \qquad \text{for all } w \in S_n \text{ and all } n\text{-tableaux } T$$

(recalling that $T$ is a map from $D$ to $[n]$). This really is a well-defined action of $S_n$, since composition of maps is associative. Explicitly, this action can be described as follows: The tableau $w \rightharpoonup T$ is obtained from $T$ by applying $w$ to each entry (i.e., replacing each entry $i$ by $w(i)$). Thus, this action is called the *action on entries*.

For example, for $D = Y(2,1)$ and $w \in S_3$, we have

$$w \rightharpoonup \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} = \begin{array}{|c|c|}\hline w(1) & w(2) \\ \hline w(3) \\ \cline{1-1}\end{array} \qquad \text{and} \qquad w \rightharpoonup \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} = \begin{array}{|c|c|}\hline w(2) & w(1) \\ \hline w(3) \\ \cline{1-1}\end{array} .$$

(The cells have been enlarged merely to fit the entries into them; they are still meant to be $1 \times 1$-squares.)

The $S_n$-action on $\{n\text{-tableaux of shape } D\}$ defined in Definition 5.3.9 can be furthermore "pulled down" to the set $\{n\text{-tabloids of shape } D\}$, in a rather obvious way:

**Definition 5.3.10.** Let $D$ be any diagram. Then, we define a left $S_n$-action on the set $\{n\text{-tabloids of shape } D\}$ by setting

$$w \rightharpoonup \overline{T} := \overline{w \rightharpoonup T} = \overline{w \circ T} \qquad \text{for all } w \in S_n \text{ and all } n\text{-tableaux } T.$$

This action is well-defined (we will justify this in a moment), and is again called the *action on entries* (but this time on tabloids instead of tableaux). This action makes the set $\{n\text{-tabloids of shape } D\}$ into a left $S_n$-set.

Why is this action well-defined? The haughty answer is "by Definition 4.1.45 **(b)**". Indeed, as we know, the $n$-tableaux of shape $D$ are the equivalence classes of the "row-equivalence" relation on the set $\{n\text{-tableaux of shape } D\}$. The latter relation is easily seen to be $S_n$-invariant[52]. Thus, Definition 4.1.45 **(b)** yields that the quotient set

$$\{n\text{-tableaux of shape } D\} \,/\, (\text{row-equivalence})$$

becomes a left $S_n$-set. But this quotient set is precisely the set $\{n\text{-tabloids of shape } D\}$, and the left $S_n$-action we get is precisely the one defined in Definition 5.3.10. Thus, the $S_n$-action in Definition 5.3.10 is well-defined. Of course, we can obtain the same well-definedness "by foot": Observe that if $w \in S_n$ is any permutation, and if $T$ and $S$ are two row-equivalent $n$-tableaux, then the compositions $w \circ T$ and $w \circ S$ are row-equivalent as well. Thus, the $n$-tabloid $\overline{w \circ T}$ depends only on $w$ and on $\overline{T}$ but not on the representative $T$. This shows that $w \rightharpoonup \overline{T}$ is well-defined. Verifying the associativity and unitality axioms is completely straightforward.

As usual with group actions, we abbreviate $w \rightharpoonup T$ by $wT$ (and likewise $w \rightharpoonup \overline{T}$ by $w\overline{T}$) whenever we feel like doing so. Thus, for example, the equality in Definition 5.3.10 can be rewritten as

$$w\overline{T} := \overline{wT} = \overline{w \circ T}. \tag{145}$$

**Remark 5.3.11.** There is also an action on places: Let $D$ be any diagram. Then, the symmetric group $S_D$ (that is, the group of all permutations of the cells of $D$) acts on the sets

$$\{n\text{-tableaux of shape } D\} \qquad \text{and} \qquad \{\text{tableaux of shape } D\}$$

from the **right**. In both cases, the action is defined by

$$T \leftharpoonup u := T \circ u \qquad \text{for all tableaux } T \text{ and all } u \in S_D.$$

These actions are called the *action on places*. Note that the group acting here is $S_D$, not $S_n$.

For instance, for $D = Y(2,1)$ and $u = t_{(1,1),(1,2)} \in S_D$, we have

$$\begin{array}{|c|c|}\hline i & j \\ \hline k \\ \cline{1-1}\end{array} \leftharpoonup u = \begin{array}{|c|c|}\hline j & i \\ \hline k \\ \cline{1-1}\end{array}.$$

No such action can be defined on any set of $n$-tabloids, since the definition of an $n$-tabloid "smears an entry across the entire row" and thus makes it unclear what entries to move where.

[52] This is just saying that if $w \in S_n$ is any permutation, and if $T$ and $S$ are two row-equivalent $n$-tableaux, then the compositions $w \circ T$ and $w \circ S$ are row-equivalent as well. Convince yourself that this is indeed obvious.

We shall next prove a few basic properties of our $S_n$-actions:

**Proposition 5.3.12.** Let $D$ be any diagram. Let $T$ be an $n$-tableau of shape $D$. Let $w \in S_n$. Then:

**(a)** The tableau $w \rightharpoonup T$ is row-equivalent to $T$ if and only if $w \in \mathcal{R}(T)$.

**(b)** The tableau $w \rightharpoonup T$ is column-equivalent to $T$ if and only if $w \in \mathcal{C}(T)$.

**(c)** We have

$$\mathcal{R}(w \rightharpoonup T) = w\mathcal{R}(T) w^{-1}$$

and

$$\mathcal{C}(w \rightharpoonup T) = w\mathcal{C}(T) w^{-1}.$$

*Proof.* **(a)** We must prove the equivalence

$$(w \rightharpoonup T \text{ is row-equivalent to } T) \iff (w \in \mathcal{R}(T)). \tag{146}$$

We will prove the "$\Longleftarrow$" and the "$\Longrightarrow$" directions of this equivalence separately:

$\Longleftarrow$: Assume that $w \in \mathcal{R}(T)$. Thus, the permutation $w$ is horizontal for $T$ (by the definition of $\mathcal{R}(T)$). In other words, for each $i \in [n]$, the number $w(i)$ lies in the same row of $T$ as $i$ does (by the definition of "horizontal").

Recall that $T$ is an $n$-tableau, thus injective (by the definition of an $n$-tableau). Hence, $T$ has no two equal entries. Similarly, $w \rightharpoonup T$ has no two equal entries (since $w \rightharpoonup T$ is an $n$-tableau).

Let $r \in \mathbb{Z}$. Let $t_1, t_2, \ldots, t_k$ be all the entries in the $r$-th row of $T$.

These entries $t_1, t_2, \ldots, t_k$ are distinct (since $T$ has no two equal entries). Hence, their images $w(t_1), w(t_2), \ldots, w(t_k)$ are distinct as well (since $w$ is a permutation and thus injective).

However, we know that for each $i \in [n]$, the number $w(i)$ lies in the same row of $T$ as $i$ does. Thus, the numbers $w(t_1), w(t_2), \ldots, w(t_k)$ lie in the same rows of $T$ as the numbers $t_1, t_2, \ldots, t_k$ do. Since the latter numbers lie in the $r$-th row of $T$, it follows that the former do as well. In other words,

$$\begin{aligned}\{w(t_1), w(t_2), \ldots, w(t_k)\} &\subseteq \{\text{the entries in the } r\text{-th row of } T\} \\ &= \{t_1, t_2, \ldots, t_k\}\end{aligned}$$

(since $t_1, t_2, \ldots, t_k$ are all the entries in the $r$-th row of $T$).

The set $\{w(t_1), w(t_2), \ldots, w(t_k)\}$ has size $k$ (since $w(t_1), w(t_2), \ldots, w(t_k)$ are distinct), and the set $\{t_1, t_2, \ldots, t_k\}$ has size $k$ as well (since $t_1, t_2, \ldots, t_k$ are distinct). Thus, the two finite sets $\{w(t_1), w(t_2), \ldots, w(t_k)\}$ and $\{t_1, t_2, \ldots, t_k\}$ have equal sizes (namely, $k$). It is well-known that if two finite sets $U$ and $V$ have equal sizes and satisfy $U \subseteq V$, then $U = V$. Applying this to $U = \{w(t_1), w(t_2), \ldots, w(t_k)\}$ and $V = \{t_1, t_2, \ldots, t_k\}$, we thus obtain

$$\{w(t_1), w(t_2), \ldots, w(t_k)\} = \{t_1, t_2, \ldots, t_k\} \tag{147}$$

(since $\{w(t_1), w(t_2), \ldots, w(t_k)\} \subseteq \{t_1, t_2, \ldots, t_k\}$).

But the entries in the $r$-th row of $T$ are $t_1, t_2, \ldots, t_k$ (by the definition of $t_1, t_2, \ldots, t_k$). Hence, the entries in the $r$-th row of $w \rightharpoonup T$ are $w(t_1), w(t_2), \ldots, w(t_k)$ (since the tableau $w \rightharpoonup T$ is obtained from $T$ by replacing each entry $i$ by $w(i)$). In other words,

$$\begin{aligned}\{\text{the entries in the } r\text{-th row of } w \rightharpoonup T\} &= \{w(t_1), w(t_2), \ldots, w(t_k)\} \\ &= \{t_1, t_2, \ldots, t_k\} \qquad \text{(by (147))} \\ &= \{\text{the entries in the } r\text{-th row of } T\}\end{aligned}$$

(by the definition of $t_1, t_2, \ldots, t_k$). In other words, the set of entries in the $r$-th row of $w \rightharpoonup T$ is the set of entries in the $r$-th row of $T$. Moreover, we can replace both words "set" in this sentence by "multiset", since the multisets in question contain no element more than once (because neither $w \rightharpoonup T$ nor $T$ has two equal entries). Thus, the multiset of entries in the $r$-th row of $w \rightharpoonup T$ is the multiset of entries in the $r$-th row of $T$.

Forget that we fixed $r$. We thus have shown that for every $r \in \mathbb{Z}$, the multiset of entries in the $r$-th row of $w \rightharpoonup T$ is the multiset of entries in the $r$-th row of $T$. In other words, the tableaux $w \rightharpoonup T$ and $T$ have the same entries in corresponding rows. In other words, these tableaux $w \rightharpoonup T$ and $T$ are row-equivalent. That is, $w \rightharpoonup T$ is row-equivalent to $T$. This proves the "$\Longleftarrow$" direction of the equivalence (146).

$\Longrightarrow$: Assume that $w \rightharpoonup T$ is row-equivalent to $T$.

Now, let $i \in [n]$. Then, $i$ appears in some cell of $T$ (since $T$ is an $n$-tableau). Let $r$ be the row in which this cell lies. Hence, $i$ appears in the $r$-th row of $T$.

Recall that the tableau $w \rightharpoonup T$ is obtained from $T$ by applying $w$ to each entry. Hence, whatever cell of $T$ contains the entry $i$ will contain the entry $w(i)$ in $w \rightharpoonup T$. In particular, this shows that $w(i)$ appears in the $r$-th row of $w \rightharpoonup T$ (since $i$ appears in the $r$-th row of $T$).

But the tableaux $w \rightharpoonup T$ and $T$ are row-equivalent, i.e., have the same entries in corresponding rows (by the definition of "row-equivalent"). Thus, any entry that appears in the $r$-th row of $w \rightharpoonup T$ must also appear in the $r$-th row of $T$. Thus, $w(i)$ appears in the $r$-th row of $T$ (since $w(i)$ appears in the $r$-th row of $w \rightharpoonup T$). Therefore, the number $w(i)$ lies in the same row of $T$ as $i$ does (namely, in the $r$-th row).

Forget that we fixed $i$. We thus have shown that for each $i \in [n]$, the number $w(i)$ lies in the same row of $T$ as $i$ does. In other words, the permutation $w$ is horizontal for $T$ (by the definition of "horizontal"). In other words, $w \in \mathcal{R}(T)$ (by the definition of $\mathcal{R}(T)$). This proves the "$\Longrightarrow$" direction of the equivalence (146).

Now, we have proved both "$\Longleftarrow$" and "$\Longrightarrow$" directions of the equivalence (146). Thus, this equivalence holds. This proves Proposition 5.3.12 **(a)**.

**(b)** This is analogous to part **(a)**. (We just need to replace rows by columns, and the symbol $\mathcal{R}$ by the symbol $\mathcal{C}$.)

**(c)** We have $w \in S_n$. Thus, $w$ is a bijection from $[n]$ to $[n]$.

If $u \in S_n$ is any permutation, then we have the following chain of equivalences:

$$
\begin{aligned}
&(u \in \mathcal{R}(w \rightharpoonup T)) \\
&\iff (\text{the permutation } u \text{ is horizontal for } w \rightharpoonup T) \\
&\qquad (\text{by the definition of } \mathcal{R}(w \rightharpoonup T)) \\
&\iff (\text{for each } i \in [n], \text{ the number } u(i) \text{ lies in the same row of } w \rightharpoonup T \text{ as } i) \\
&\qquad (\text{by the definition of a horizontal permutation}) \\
&\iff (\text{for each } j \in [n], \text{ the number } u(w(j)) \text{ lies in the same row of } w \rightharpoonup T \text{ as } w(j)) \\
&\qquad \left( \begin{array}{c} \text{here, we have substituted } w(j) \text{ for } i \text{ in the "for all"-sentence,} \\ \text{since } w : [n] \to [n] \text{ is a bijection} \end{array} \right) \\
&\iff \left(\text{for each } j \in [n], \text{ the number } w^{-1}(u(w(j))) \text{ lies in the same row of } w \rightharpoonup T \text{ as } j\right) \\
&\qquad \left( \begin{array}{c} \text{indeed, } w \rightharpoonup T \text{ is obtained from } T \text{ by applying } w \text{ to each entry;} \\ \text{thus, the row of } w \rightharpoonup T \text{ that contains } w(j) \text{ is precisely the} \\ \text{row of } T \text{ that contains } j, \text{ whereas the row of } w \rightharpoonup T \text{ that} \\ \text{contains } u(w(j)) \text{ is precisely the row of } T \text{ that contains } w^{-1}(u(w(j))) \end{array} \right) \\
&\iff \left(\text{for each } i \in [n], \text{ the number } w^{-1}(u(w(i))) \text{ lies in the same row of } w \rightharpoonup T \text{ as } i\right) \\
&\qquad (\text{here, we renamed the index } j \text{ as } i) \\
&\iff \left(\text{for each } i \in [n], \text{ the number } \left(w^{-1}uw\right)(i) \text{ lies in the same row of } w \rightharpoonup T \text{ as } i\right) \\
&\qquad \left(\text{since } w^{-1}(u(w(i))) = \left(w^{-1}uw\right)(i) \text{ for each } i \in [n]\right) \\
&\iff \left(\text{the permutation } w^{-1}uw \text{ is horizontal for } T\right) \\
&\qquad (\text{by the definition of a horizontal permutation}) \\
&\iff \left(w^{-1}uw \in \mathcal{R}(T)\right) \qquad (\text{by the definition of } \mathcal{R}(T)) \\
&\iff \left(u \in w\mathcal{R}(T)w^{-1}\right). \qquad (148)
\end{aligned}
$$

But $\mathcal{R}(w \rightharpoonup T)$ is a subset of $S_n$. Hence,

$$
\begin{aligned}
\mathcal{R}(w \rightharpoonup T) &= \{u \in S_n \mid u \in \mathcal{R}(w \rightharpoonup T)\} \\
&= \left\{u \in S_n \mid u \in w\mathcal{R}(T)w^{-1}\right\} \qquad (\text{by the equivalence (148)}) \\
&= w\mathcal{R}(T)w^{-1} \qquad \left(\text{since } w\mathcal{R}(T)w^{-1} \text{ is a subset of } S_n\right).
\end{aligned}
$$

An analogous argument (using columns instead of rows) shows that $\mathcal{C}(w \rightharpoonup T) = w\mathcal{C}(T)w^{-1}$. Thus, Proposition 5.3.12 **(c)** is proved. □

**Proposition 5.3.13.** Let $D$ be any diagram. Let $T$ be an $n$-tableau of shape $D$. Then:

**(a)** We have $\mathcal{R}(T) = \left\{w \in S_n \mid \overline{w \rightharpoonup T} = \overline{T}\right\}$.

**(b)** Any $n$-tableau of shape $D$ can be written as $w \rightharpoonup T$ for some $w \in S_n$.

**(c)** Any $n$-tabloid of shape $D$ can be written as $w \rightharpoonup \overline{T}$ for some $w \in S_n$.

**(d)** The map

$$f : S_n \to \{n\text{-tableaux of shape } D\}, \\ w \mapsto w \rightharpoonup T$$

is a bijection.

*Proof.* **(a)** Let $w \in S_n$. Then, we have the equivalence

$$(w \in \mathcal{R}(T)) \iff (w \rightharpoonup T \text{ is row-equivalent to } T) \qquad (\text{by Proposition 5.3.12 } \textbf{(a)}).$$

But an $n$-tabloid is defined as an equivalence class of $n$-tableaux with respect to row-equivalence. Thus, the $n$-tableaux $w \rightharpoonup T$ and $T$ are row-equivalent if and only if their equivalence classes $\overline{w \rightharpoonup T}$ and $\overline{T}$ are the same. In other words, the equivalence

$$(w \rightharpoonup T \text{ is row-equivalent to } T) \iff \left(\overline{w \rightharpoonup T} = \overline{T}\right)$$

holds. Altogether, we thus have proved the chain of equivalences

$$\begin{aligned}(w \in \mathcal{R}(T)) &\iff (w \rightharpoonup T \text{ is row-equivalent to } T) \\ &\iff \left(\overline{w \rightharpoonup T} = \overline{T}\right). \end{aligned} \tag{149}$$

Now, forget that we fixed $w$. We thus have proved the equivalence (149) for each $w \in S_n$.

But $\mathcal{R}(T)$ is a subset (and in fact a subgroup) of $S_n$. Thus,

$$\begin{aligned}\mathcal{R}(T) &= \{w \in S_n \mid w \in \mathcal{R}(T)\} \\ &= \left\{w \in S_n \mid \overline{w \rightharpoonup T} = \overline{T}\right\} \qquad (\text{by the equivalence (149)}).\end{aligned}$$

This proves Proposition 5.3.13 **(a)**.

**(b)** Let $U$ be any $n$-tableau of shape $D$. We must show that $U = w \rightharpoonup T$ for some $w \in S_n$.

Intuitively, this is clear: Both tableaux $T$ and $U$ contain the entries $1, 2, \ldots, n$ (each exactly once); thus, we can obtain $U$ from $T$ by permuting these entries.

The rigorous proof is not much harder: We know that $T$ is an $n$-tableau of shape $D$, that is, an injective map from $D$ to $[n]$ whose image is $[n]$. Thus, we can view $T$ as a bijection from $D$ to $[n]$. Likewise, we can view $U$ as a bijection from $D$ to $[n]$. Bijections are always invertible, so that $T^{-1}$ is a bijection from $[n]$ to $D$ (since $T$ is a bijection from $D$ to $[n]$). Hence, $U \circ T^{-1}$ is a bijection from $[n]$ to $[n]$ (since a composition of two bijections is a bijection). In other words, $U \circ T^{-1}$ is a permutation of $[n]$. That is, $U \circ T^{-1} \in S_n$.

Now, the definition of the action on entries yields $\left(U \circ T^{-1}\right) \rightharpoonup T = U \circ \underbrace{T^{-1} \circ T}_{=\mathrm{id}} = U$.

Thus, $U = w \rightharpoonup T$ for some $w \in S_n$ (namely, for $w = U \circ T^{-1}$). This proves Proposition 5.3.13 **(b)**.

**(c)** Let $\overline{U}$ be an $n$-tabloid of shape $D$. We must show that $\overline{U} = w \rightharpoonup \overline{T}$ for some $w \in S_n$.

Proposition 5.3.13 **(b)** shows that $U = w \rightharpoonup T$ for some $w \in S_n$. Using this $w$, we then have

$$\begin{aligned} \overline{U} &= \overline{w \rightharpoonup T} && (\text{since } U = w \rightharpoonup T) \\ &= w \rightharpoonup \overline{T} && (\text{by Definition 5.3.10}) . \end{aligned}$$

Thus, Proposition 5.3.13 **(c)** is proved.

**(d)** Consider the map

$$\begin{aligned} f : S_n &\to \{n\text{-tableaux of shape } D\} , \\ w &\mapsto w \rightharpoonup T. \end{aligned}$$

This map is clearly well-defined. Moreover:

- The map $f$ is injective.

  [*Proof:* Let $u, v \in S_n$ be such that $f(u) = f(v)$. We must prove that $u = v$.

  We have $f(u) = u \rightharpoonup T$ (by the definition of $f$). But $u \rightharpoonup T = u \circ T$ (by Definition 5.3.9). Hence, $f(u) = u \rightharpoonup T = u \circ T$. Similarly, $f(v) = v \circ T$. Thus, $u \circ T = f(u) = f(v) = v \circ T$.

  But $T$ is an $n$-tableau of shape $D$. In other words, $T$ is a bijection from $D$ to $[n]$. Hence, the map $T$ has an inverse $T^{-1} : [n] \to D$. Clearly, $u \circ \underbrace{T \circ T^{-1}}_{=\mathrm{id}} = u$, so that $u = \underbrace{u \circ T}_{=v \circ T} \circ T^{-1} = v \circ \underbrace{T \circ T^{-1}}_{=\mathrm{id}} = v$.

  Forget that we fixed $u, v$. We thus have shown that if $u, v \in S_n$ are such that $f(u) = f(v)$, then $u = v$. In other words, the map $f$ is injective.]

- The map $f$ is surjective.

  [*Proof:* Let $Q \in \{n\text{-tableaux of shape } D\}$. Thus, $Q$ is an $n$-tableau of shape $D$. But Proposition 5.3.13 **(b)** shows that any $n$-tableau of shape $D$ can be written as $w \rightharpoonup T$ for some $w \in S_n$. In particular, $Q$ can be written as $w \rightharpoonup T$ for some $w \in S_n$ (since $Q$ is an $n$-tableau of shape $D$). In other words, $Q = w \rightharpoonup T$ for some $w \in S_n$. Consider this $w$.

  Now, the definition of $f$ yields $f(w) = w \rightharpoonup T = Q$. Hence, $Q = f(w) \in f(S_n)$.

  Forget that we fixed $Q$. We thus have shown that $Q \in f(S_n)$ for every $Q \in \{n\text{-tableaux of shape } D\}$. In other words, $\{n\text{-tableaux of shape } D\} \subseteq f(S_n)$. This shows that $f$ is surjective.]

Now, we have shown that $f$ is both injective and surjective. Hence, $f$ is bijective, i.e., is a bijection. This proves Proposition 5.3.13 **(d)**. □

For $n$-tableaux, we can also describe row-equivalence in terms of horizontal permutations, and likewise for column-equivalence:

**Proposition 5.3.14.** Let $D$ be a diagram. Let $T$ and $S$ be two $n$-tableaux of shape $D$. Then:

**(a)** The $n$-tableaux $T$ and $S$ are row-equivalent if and only if there exists some $w \in \mathcal{R}(T)$ such that $S = w \rightharpoonup T$.

**(b)** The $n$-tableaux $T$ and $S$ are column-equivalent if and only if there exists some $w \in \mathcal{C}(T)$ such that $S = w \rightharpoonup T$.

*Proof.* **(a)** We prove the "$\Longleftarrow$" and "$\Longrightarrow$" directions separately:

$\Longleftarrow$: Assume that there exists some $w \in \mathcal{R}(T)$ such that $S = w \rightharpoonup T$. Consider this $w$. The tableau $w \rightharpoonup T$ is row-equivalent to $T$ (by Proposition 5.3.12 **(a)**, since $w \in \mathcal{R}(T)$). In other words, $S$ is row-equivalent to $T$ (since $S = w \rightharpoonup T$). In other words, $T$ and $S$ are row-equivalent. This proves the "$\Longleftarrow$" direction of Proposition 5.3.14 **(a)**.

$\Longrightarrow$: Assume that the tableaux $T$ and $S$ are row-equivalent. Proposition 5.3.13 **(b)** shows that the tableau $S$ can be written as $w \rightharpoonup T$ for some $w \in S_n$. Consider this $w$.

We assumed that the tableaux $T$ and $S$ are row-equivalent. In other words, the tableaux $T$ and $w \rightharpoonup T$ are row-equivalent (since $S = w \rightharpoonup T$). By Proposition 5.3.12 **(a)**, this entails that $w \in \mathcal{R}(T)$. Thus, we have found a $w \in \mathcal{R}(T)$ such that $S = w \rightharpoonup T$. Hence, such a $w$ exists. This proves the "$\Longrightarrow$" direction of Proposition 5.3.14 **(a)**.

**(b)** This is analogous to part **(a)**. □

**Exercise 5.3.2.** $\boxed{2}$ Let $D$ be a diagram with $|D| = n$. Fix an $n$-tableau $T$ of shape $D$. Then, its row group $\mathcal{R}(T)$ is a subgroup of $S_n$, and thus there is a left $S_n$-set $S_n/\mathcal{R}(T)$ consisting of all left cosets of $\mathcal{R}(T)$ in $S_n$ (see Example 4.1.11 for its definition). On the other hand, the set $\{n\text{-tabloids of shape } D\}$ is a left $S_n$-set as well. Prove that these two left $S_n$-sets are isomorphic. To be more specific: Prove that the map

$$\begin{aligned} S_n/\mathcal{R}(T) &\to \{n\text{-tabloids of shape } D\}, \\ w\mathcal{R}(T) &\mapsto \overline{w \rightharpoonup T} \end{aligned}$$

is a left $S_n$-set isomorphism.

### 5.3.5. Young modules

Now, recall the concept of a permutation module (defined in Example 4.3.9), which automatically turns any left $G$-set $X$ (where $G$ is a group) into a left $\mathbf{k}[G]$-module $\mathbf{k}^{(X)}$ by linearization (i.e., by extending the left $G$-action $G \times X \to X$ to a $\mathbf{k}$-bilinear map $\mathbf{k}[G] \times \mathbf{k}^{(X)} \to \mathbf{k}^{(X)}$ in the only reasonable way). We shall now use this technique to turn the left $S_n$-set $\{n\text{-tabloids of shape } D\}$ (as defined in Definition 5.3.10) into a left $\mathbf{k}[S_n]$-module, which is known as a *Young module*:

**Definition 5.3.15.** Let $D$ be a diagram with $|D| = n$. The *Young module* $\mathcal{M}^D$ (sometimes also denoted by $\mathcal{Y}^D$) is the permutation module corresponding to the left $S_n$-set

$$\{n\text{-tabloids of shape } D\}$$

(which was introduced in Definition 5.3.10). As a $\mathbf{k}$-module, it is the free $\mathbf{k}$-module $\mathbf{k}^{(\{n\text{-tabloids of shape } D\})}$ on the set $\{n\text{-tabloids of shape } D\}$. It has a standard basis vector $e_{\overline{T}}$ for each $n$-tabloid $\overline{T}$ of shape $D$, but we shall just denote this basis vector by $\overline{T}$ by the usual abuse of notation. An arbitrary element of $\mathcal{M}^D$ is thus a formal $\mathbf{k}$-linear combination of $n$-tabloids of shape $D$. The symmetric group $S_n$ acts on $\mathcal{M}^D$ by permuting these $n$-tabloids (i.e., the respective standard basis vectors).

**Example 5.3.16.** Assume that $n \geq 2$. Let $D$ be the Young diagram $Y(n-1,1)$. This Young diagram has two rows of lengths $n-1$ and 1, respectively.

As we know from Example 5.3.7 **(b)**, there are $n$ many $n$-tabloids of shape $D$. We can label these $n$-tabloids as $\overline{\underline{1}}$, $\overline{\underline{2}}$, $\ldots$, $\overline{\underline{n}}$, where $\overline{\underline{i}}$ means the $n$-tabloid whose second row has entry $i$ and whose first row contains all other elements of $[n]$.

Thus, the Young module $\mathcal{M}^D$ has basis $\left(\overline{\underline{1}}, \overline{\underline{2}}, \ldots, \overline{\underline{n}}\right)$ as a $\mathbf{k}$-module. The action of $S_n$ on this module is given by

$$w\,\overline{\underline{i}} = \overline{\underline{w(i)}} \qquad \text{for each } w \in S_n \text{ and each } i \in [n].$$

For instance, for $n = 4$, we have

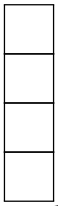
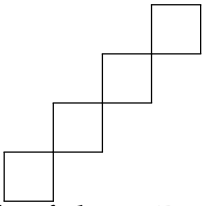
$$\mathrm{cyc}_{1,2,3}\ \overline{\underline{2}} = \mathrm{cyc}_{1,2,3}\,\overline{\underline{\begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline \end{array}\!\!\begin{array}{|c|}\hline 2 \\ \hline \end{array}}} = \overline{\underline{\begin{array}{|c|c|c|}\hline 2 & 1 & 4 \\ \hline \end{array}\!\!\begin{array}{|c|}\hline 3 \\ \hline \end{array}}} = \overline{\underline{\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline \end{array}\!\!\begin{array}{|c|}\hline 3 \\ \hline \end{array}}} = \overline{\underline{3}} = \overline{\underline{\mathrm{cyc}_{1,2,3}(2)}}.$$

The Young module $\mathcal{M}^D$ is just the natural $S_n$-representation $\mathbf{k}^n$ in disguise. Indeed, the $\mathbf{k}$-linear map

$$\begin{aligned} \mathbf{k}^n &\to \mathcal{M}^D, \\ e_i &\mapsto \overline{\underline{i}} \end{aligned}$$

is an isomorphism of $S_n$-representations (since any $w \in S_n$ sends $e_i$ to $e_{w(i)}$ and sends $\overline{\underline{i}}$ to $\overline{\underline{w(i)}}$).

**Example 5.3.17.** Let $D$ be a diagram that has no two cells in the same row and satisfies $|D| = n$. (For example, we can take $D$ to be the Young diagram $Y\left(\underbrace{1,1,\ldots,1}_{n \text{ times}}\right)$, or the skew Young diagram $Y((n,n-1,\ldots,1)/(n-1,n-2,\ldots,1))$. For $n = 4$, these two diagrams are

respectively  and  .)

Then, the $n$-tabloids of shape $D$ are "essentially" the permutations $w \in S_n$. Indeed, no two distinct $n$-tableaux are row-equivalent (since no row has more than one entry). Thus, each $n$-tabloid consists of just one $n$-tableau, and can be identified with this one $n$-tableau that it contains. But there are $n!$ many $n$-tableaux, since we can easily find a bijection from $S_n$ to the set of all $n$-tableaux (by choosing an arbitrary $n$-tableau $T$, and sending each $w \in S_n$ to the $n$-tableau $w \rightharpoonup T$).

The Young module $\mathcal{M}^D$ (for such a diagram $D$) is thus just the left regular representation $\mathbf{k}\left[S_n\right]$ of $S_n$. More precisely, if we fix any $n$-tableau $T$ of shape $D$, then there is a left $\mathbf{k}\left[S_n\right]$-module isomorphism

$$\begin{aligned} \mathbf{k}\left[S_n\right] &\to \mathcal{M}^D, \\ w &\mapsto \overline{w \rightharpoonup T} = w \rightharpoonup \overline{T}. \end{aligned}$$

**Exercise 5.3.3.** [2] Assume that $\mathbf{k}$ is a field. Let $D$ be a diagram that contains at least one row having more than one cell. Let $s \in \mathcal{M}^D$ be the sum of all $n$-tabloids of shape $D$ (or, rather, of the corresponding standard basis vectors). Recall from Exercise 4.3.1 that the 1-dimensional submodule $\operatorname{span}_{\mathbf{k}}\{s\}$ is a trivial subrepresentation of the Young module $\mathcal{M}^D$.

Prove that this submodule is the only 1-dimensional subrepresentation of $\mathcal{M}^D$.

[**Hint:** Any 1-dimensional subrepresentation of $\mathcal{M}^D$ is spanned by a single element, which we can write as $\sum_{\overline{T}} \alpha_{\overline{T}} \overline{T}$. Each $w \in S_n$ must multiply this element by a scalar $\kappa_w$. This gives $\alpha_{\overline{T}} = \kappa_w \alpha_{\overline{wT}}$ for all $T$ and all $w$. This shows that each $\alpha_{\overline{T}}$ must be nonzero (why?), but what else does it show?]

## 5.4. Specht modules

### 5.4.1. Definition and examples

Next, we will define a certain subrepresentation of each Young module, which is called a *Specht module* and will later turn out to be a basic building block of the representation theory of $S_n$.

**Definition 5.4.1.** Let $D$ be a diagram with $|D| = n$.

**(a)** If $T$ is an $n$-tableau of shape $D$, then we define the so-called *polytabloid*

$$\mathbf{e}_T := \sum_{w \in \mathcal{C}(T)} (-1)^w \, \overline{wT} \in \mathcal{M}^D.$$

**(b)** The span of these elements $\mathbf{e}_T$, where $T$ ranges over all $n$-tableaux of shape $D$, is called the *Specht module* $\mathcal{S}^D$. Thus,

$$\mathcal{S}^D = \operatorname{span}_{\mathbf{k}} \left\{ \mathbf{e}_T \mid T \text{ is an } n\text{-tableau of shape } D \right\}.$$

As we will soon see (in Lemma 5.4.6 **(b)**), this Specht module is a left $\mathbf{k}[S_n]$-submodule (i.e., an $S_n$-subrepresentation) of the Young module $\mathcal{M}^D$.

**Warning:** Do not mistake the polytabloid $\mathbf{e}_T$ for the standard basis vector corresponding to the tabloid $\overline{T}$ (which we just call $\overline{T}$). The letter "$\mathbf{e}$" in "$\mathbf{e}_T$" has nothing to do with the notation $e_i$ for a standard basis vector.

Also, keep in mind that tabloids are combinatorial objects (sets of tableaux), while polytabloids are algebraic ones (elements of the $S_n$-representation $\mathcal{M}^D$).

**Example 5.4.2.** Let $n = 3$ and $D = Y(2,1)$. Then, the $n$-tableaux of shape $D$ are

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 3 & 1 \\ \hline 2 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 3 & 2 \\ \hline 1 \\ \cline{1-1}\end{array}.$$

Following Convention 5.2.4, we can just call them $12\backslash\backslash 3$ and $13\backslash\backslash 2$ and so on.

The $n$-tabloids of shape $D$ are

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array},$$

aka $\overline{12\backslash\backslash 3}$ and $\overline{13\backslash\backslash 2}$ and $\overline{23\backslash\backslash 1}$. Each $n$-tabloid contains two $n$-tableaux.

Each of the six $n$-tableaux $T$ of shape $D$ gives rise to a polytabloid. Here are two of them:

$$\begin{aligned}
\mathbf{e}_{12\backslash\backslash 3} &= \sum_{w \in \mathcal{C}(12\backslash\backslash 3)} (-1)^w \, \overline{w\,(12\backslash\backslash 3)} \\
&= \underbrace{(-1)^{\mathrm{id}}}_{=1} \underbrace{\overline{\mathrm{id}\,(12\backslash\backslash 3)}}_{=\overline{12\backslash\backslash 3}} + \underbrace{(-1)^{t_{1,3}}}_{=-1} \underbrace{\overline{t_{1,3}\,(12\backslash\backslash 3)}}_{=\overline{32\backslash\backslash 1}} \\
&\qquad\qquad (\text{since } \mathcal{C}(12\backslash\backslash 3) = \{\mathrm{id}, t_{1,3}\}) \\
&= \overline{12\backslash\backslash 3} - \underbrace{\overline{32\backslash\backslash 1}}_{=\overline{23\backslash\backslash 1}} = \overline{12\backslash\backslash 3} - \overline{23\backslash\backslash 1}
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{e}_{21\backslash\backslash 3} &= \sum_{w \in \mathcal{C}(21\backslash\backslash 3)} (-1)^w \, \overline{w\,(21\backslash\backslash 3)} \\
&= (-1)^{\mathrm{id}} \, \overline{\mathrm{id}\,(21\backslash\backslash 3)} + (-1)^{t_{2,3}} \, \overline{t_{2,3}\,(21\backslash\backslash 3)} \\
&\qquad\qquad (\text{since } \mathcal{C}\,(21\backslash\backslash 3) = \{\mathrm{id}, t_{2,3}\}) \\
&= \overline{21\backslash\backslash 3} - \overline{31\backslash\backslash 2} = \overline{12\backslash\backslash 3} - \overline{13\backslash\backslash 2};
\end{aligned}$$

and there are four more. Rather than compute them one by one, we can just compute them all at once: For any distinct elements $i, j, k \in [3]$, we have

$$\begin{aligned}
\mathbf{e}_{ij\backslash\backslash k} &= \sum_{w \in \mathcal{C}(ij\backslash\backslash k)} (-1)^w \, \overline{w\,(ij\backslash\backslash k)} \\
&= \underbrace{(-1)^{\mathrm{id}}}_{=1} \underbrace{\overline{\mathrm{id}\,(ij\backslash\backslash k)}}_{=\overline{ij\backslash\backslash k}} + \underbrace{(-1)^{t_{i,k}}}_{=-1} \underbrace{\overline{t_{i,k}\,(ij\backslash\backslash k)}}_{=\overline{kj\backslash\backslash i}} \\
&\qquad\qquad (\text{since } \mathcal{C}\,(ij\backslash\backslash k) = \{\mathrm{id}, t_{i,k}\}) \\
&= \overline{ij\backslash\backslash k} - \overline{kj\backslash\backslash i}.
\end{aligned}$$

Using the notation $\overline{\underline{i}}$ from Example 5.3.16, we can rewrite this as

$$\mathbf{e}_{ij\backslash\backslash k} = \overline{\underline{k}} - \overline{\underline{i}}$$

(since $\overline{\underline{k}} = \overline{ij\backslash\backslash k}$ and $\overline{\underline{i}} = \overline{kj\backslash\backslash i}$).

Thus, the Specht module $\mathcal{S}^D$ is spanned by the differences $\overline{\underline{k}} - \overline{\underline{i}}$ for all triples of distinct elements $i, j, k \in [3]$. In other words, it is spanned by the differences $\overline{\underline{k}} - \overline{\underline{i}}$ for all pairs of distinct elements $k \neq i$ in $[3]$. Identifying the Young module $\mathcal{M}^D$ with the natural representation $\mathbf{k}^n$ of $S_n$ for $n = 3$ (as explained in Example 5.3.16), we can see thus view $\mathcal{S}^D$ as the span of the differences $e_k - e_i$ of distinct standard basis vectors. But, as we know, this span is the zero-sum subrepresentation $R\,(\mathbf{k}^n)$.

The same reasoning works not just for $n = 3$ but for any $n \geq 2$. Thus, we can see that the Specht module $\mathcal{S}^{Y(n-1,1)}$ is isomorphic to the zero-sum subrepresentation $R\,(\mathbf{k}^n)$ of the natural $S_n$-representation $\mathbf{k}^n$.

**Example 5.4.3.** Now consider the partition $(n)$ (by which we understand the empty partition if $n = 0$). Let $D = Y\,(n)$. Then, all cells of $D$ lie in the same row. Hence, any two $n$-tableaux of shape $D$ are row-equivalent. Thus, there is only one $n$-tabloid, namely $\overline{\begin{array}{cccc} 1 & 2 & \cdots & n \end{array}}$. The left $S_n$-action on the set of this one $n$-tabloid is trivial, and thus the Young module $\mathcal{M}^{Y(n)}$ is the trivial $S_n$-representation. The Specht module $\mathcal{S}^{Y(n)}$ is also the trivial $S_n$-representation, since $\mathcal{C}\,(T) = \{\mathrm{id}\}$ for any $n$-tableau $T$ of shape $D$.

**Example 5.4.4.** Now consider the partition $(1^n)$ (this is exponential notation for $\underbrace{(1,1,\ldots,1)}_{n \text{ times}}$, as explained in Definition 5.1.8). Let $D = Y(1^n)$. Then, the diagram $D$ consists of a single column, so it has no two cells in the same row. As we saw in Example 5.3.17, this entails that the $n$-tabloids of shape $D$ are all the permutations $w \in S_n$ (encoded as the respective $n$-tabloids $\begin{array}{|c|}\hline w(1) \\ \hline w(2) \\ \hline \vdots \\ \hline w(n) \\ \hline \end{array}$), and the Young module $\mathcal{M}^{Y(1^n)}$ is (isomorphic to) the left regular representation $\mathbf{k}[S_n]$ of $S_n$.

What about the Specht module $\mathcal{S}^{Y(1^n)}$ ? It is spanned by the polytabloids

$$\mathbf{e}_T = \sum_{w \in \mathcal{C}(T)} (-1)^w \overline{wT} = \sum_{w \in S_n} (-1)^w \overline{wT} \qquad (\text{since } \mathcal{C}(T) = S_n \text{ here})$$

for all $n$-tableaux $T$ of shape $Y(1^n)$. These look like a lot of polytabloids ($n!$ to be precise), but I claim that they are all equal up to sign: Indeed, if we identify $\mathcal{M}^{Y(1^n)}$ with the left regular representation $\mathbf{k}[S_n]$, then each of the polytabloids $\mathbf{e}_T$ is either the sign-integral $\nabla^-$ or $-\nabla^-$.

The easiest way to see this is by identifying each permutation $w \in S_n$ with the $n$-tableau $T_w = \begin{array}{|c|}\hline w(1) \\ \hline w(2) \\ \hline \vdots \\ \hline w(n) \\ \hline \end{array}$ and the corresponding $n$-tabloid $\overline{T_w}$. Then, conversely, any $n$-tableau $T$ can be viewed as a permutation in $S_n$ (namely, as the permutation $w \in S_n$ for which $T$ is $T_w$). Hence, for any $n$-tableau $T$, we have

$$\mathbf{e}_T = \sum_{w \in S_n} (-1)^w \underbrace{\overline{wT}}_{\substack{=wT \\ \text{(where we regard } T \\ \text{as a permutation in } S_n)}} = \underbrace{\sum_{w \in S_n} (-1)^w w}_{\substack{=\nabla^- \\ \text{(by the definition of } \nabla^-)}} \quad T = \nabla^- T$$

$$= (-1)^T \nabla^- \qquad (\text{by (18), applied to } T \text{ instead of } w).$$

Therefore, $\mathbf{e}_T$ is $\nabla^-$ or $-\nabla^-$.

Thus, as a left $\mathbf{k}[S_n]$-module, the Specht module $\mathcal{S}^{Y(1^n)}$ is isomorphic to the ($\mathbf{k}$-linear) span of $\nabla^-$. It is furthermore isomorphic to the sign representation of $S_n$, since every $w \in S_n$ satisfies $w\nabla^- = (-1)^w \nabla^-$.

**Example 5.4.5.** Let $D$ be a diagram that has no two cells in the same row and has no two cells in the same column and satisfies $|D| = n$. (For example, $D$ can be the skew Young diagram $Y((n, n-1, \ldots, 1) / (n-1, n-2, \ldots, 1))$, as in Example 5.3.17. But $D$ cannot be a straight Young diagram $Y(\lambda)$, unless $n \leq 1$.)

As we know from Example 5.3.17, the Young module $\mathcal{M}^D$ is (an isomorphic copy of) the left regular representation $\mathbf{k}[S_n]$ of $S_n$, since each $n$-tabloid of shape $D$ contains just one $n$-tableau. Furthermore, since $D$ has no two cells in the same column, each $n$-tableau $T$ of shape $D$ satisfies $\mathcal{C}(T) = \{\mathrm{id}\}$, and thus the polytabloids $\mathbf{e}_T$ are just the corresponding tabloids $\overline{T}$. Hence, the Specht module $\mathcal{S}^D$ is the span of all $n$-tabloids $\overline{T}$, which is of course the whole Young module $\mathcal{M}^D$. Thus, we have

$$\mathcal{S}^D = \mathcal{M}^D \cong \mathbf{k}[S_n]$$

in this case.

### 5.4.2. Properties

It is time to state some general properties of polytabloids and Specht modules:

**Lemma 5.4.6.** Let $D$ be any diagram with $|D| = n$. Then:

**(a)** If $T$ is any $n$-tableau of shape $D$, and if $u \in S_n$, then

$$\mathbf{e}_{uT} = u\mathbf{e}_T.$$

**(b)** The set $\mathcal{S}^D$ is a left $\mathbf{k}[S_n]$-submodule of $\mathcal{M}^D$.

**(c)** If $T$ is any $n$-tableau of shape $D$, and if $u \in \mathcal{C}(T)$, then $\mathbf{e}_{uT} = (-1)^u \mathbf{e}_T$.

**(d)** If $T$ and $S$ are two column-equivalent $n$-tableaux of shape $D$, then $\mathbf{e}_T = \pm\mathbf{e}_S$.

*Proof.* **(a)** Let $T$ be any $n$-tableau of shape $D$. Let $u \in S_n$. Then, Proposition 5.3.12 **(c)** (applied to $w = u$) yields $\mathcal{R}(u \rightharpoonup T) = u\mathcal{R}(T)u^{-1}$ and $\mathcal{C}(u \rightharpoonup T) = u\mathcal{C}(T)u^{-1}$. Since $uT$ is a shorthand for $u \rightharpoonup T$, we can rewrite these two equalities as $\mathcal{R}(uT) = u\mathcal{R}(T)u^{-1}$ and $\mathcal{C}(uT) = u\mathcal{C}(T)u^{-1}$.

From $\mathcal{C}(uT) = u\mathcal{C}(T)u^{-1}$, we see that the map

$$\mathcal{C}(T) \to \mathcal{C}(uT),$$
$$w \mapsto uwu^{-1}$$

is well-defined and surjective. This map is injective as well (since $S_n$ is a group), and thus is a bijection.

But Definition 5.4.1 **(a)** yields

$$\mathbf{e}_T = \sum_{w \in \mathcal{C}(T)} (-1)^w \, \overline{wT} \tag{150}$$

and

$$\begin{aligned}
\mathbf{e}_{uT} &= \sum_{w \in \mathcal{C}(uT)} (-1)^w \, \overline{wuT} = \sum_{w \in \mathcal{C}(T)} \underbrace{(-1)^{uwu^{-1}}}_{\substack{=(-1)^w \\ \text{(by (10))}}} \overline{u \underbrace{wu^{-1}}_{=w} u \, T} \\
&\qquad \left( \begin{array}{c} \text{here, we have substituted } uwu^{-1} \text{ for } w \text{ in the sum,} \\ \text{since the map } \mathcal{C}(T) \to \mathcal{C}(uT), \ w \mapsto uwu^{-1} \\ \text{is a bijection} \end{array} \right) \\
&= \sum_{w \in \mathcal{C}(T)} (-1)^w \underbrace{\overline{uwT}}_{\substack{=u\overline{wT} \\ \text{(by (145))}}} = \sum_{w \in \mathcal{C}(T)} (-1)^w \, u\overline{wT} = u \underbrace{\sum_{w \in \mathcal{C}(T)} (-1)^w \, \overline{wT}}_{\substack{=\mathbf{e}_T \\ \text{(by (150))}}} = u\mathbf{e}_T.
\end{aligned}$$

This proves Lemma 5.4.6 **(a)**.

**(b)** Clearly, $\mathcal{S}^D$ is a $\mathbf{k}$-submodule of $\mathcal{M}^D$ (being defined as a span). Next, we shall show that $\mathcal{S}^D$ is an $S_n$-subset of $\mathcal{M}^D$.

To show this, we must prove that $ux \in \mathcal{S}^D$ for each $u \in S_n$ and each $x \in \mathcal{S}^D$.

So let us prove this. Let $u \in S_n$ and $x \in \mathcal{S}^D$. We must show that $ux \in \mathcal{S}^D$. Since $\mathcal{S}^D$ is spanned by the polytabloids $\mathbf{e}_T$, we can WLOG assume (by linearity) that $x$ is one of these polytabloids $\mathbf{e}_T$. Assume this. Thus, $x = \mathbf{e}_T$, so that $ux = u\mathbf{e}_T = \mathbf{e}_{uT}$ (by Lemma 5.4.6 **(a)**). Thus, $ux$ is a polytabloid, and therefore belongs to $\mathcal{S}^D$ (since $\mathcal{S}^D$ is the span of all polytabloids).

Forget that we fixed $u$ and $x$. We thus have shown that $ux \in \mathcal{S}^D$ for each $u \in S_n$ and each $x \in \mathcal{S}^D$. In other words, $\mathcal{S}^D$ is an $S_n$-subset of $\mathcal{M}^D$. Since $\mathcal{S}^D$ is also a $\mathbf{k}$-submodule of $\mathcal{M}^D$, we thus conclude that $\mathcal{S}^D$ is a subrepresentation of $\mathcal{M}^D$. Hence, by Proposition 4.3.17, we conclude that $\mathcal{S}^D$ is a left $\mathbf{k}[S_n]$-submodule of $\mathcal{M}^D$. This proves Lemma 5.4.6 **(b)**.

**(c)** Let $T$ be any $n$-tableau of shape $D$. Let $u \in \mathcal{C}(T)$.

As in the proof of part **(a)**, we can see that $\mathcal{C}(uT) = u\mathcal{C}(T)\,u^{-1}$. However, $\mathcal{C}(T)$ is a group, and $u$ is one of its elements (since $u \in \mathcal{C}(T)$). Thus, $u\mathcal{C}(T)\,u^{-1}$ is just $\mathcal{C}(T)$. Hence, we can rewrite the equality $\mathcal{C}(uT) = u\mathcal{C}(T)\,u^{-1}$ as $\mathcal{C}(uT) = \mathcal{C}(T)$.

Moreover, $u$ is an element of the group $\mathcal{C}(T)$. Thus, so is its inverse $u^{-1}$. Hence, the map $\mathcal{C}(T) \to \mathcal{C}(T), \ w \mapsto wu^{-1}$ is a bijection (with inverse map given by $w \mapsto wu$).

Now, Definition 5.4.1 **(a)** yields

$$\mathbf{e}_{uT} = \sum_{w \in \mathcal{C}(uT)} (-1)^w \, \overline{wuT} = \sum_{w \in \mathcal{C}(T)} (-1)^w \, \overline{wuT} \qquad (\text{since } \mathcal{C}(uT) = \mathcal{C}(T))$$

$$= \sum_{w \in \mathcal{C}(T)} \underbrace{(-1)^{wu^{-1}}}_{\substack{=(-1)^w (-1)^{u^{-1}} \\ \text{(by the multiplicativity} \\ \text{of the sign)}}} \underbrace{\overline{wu^{-1}u}}_{=w} T$$

$$\left( \begin{array}{c} \text{here, we have substituted } wu^{-1} \text{ for } w \text{ in the sum,} \\ \text{since the map } \mathcal{C}(T) \to \mathcal{C}(T),\ w \mapsto wu^{-1} \\ \text{is a bijection} \end{array} \right)$$

$$= \sum_{w \in \mathcal{C}(T)} (-1)^w (-1)^{u^{-1}} \, \overline{wT} = (-1)^{u^{-1}} \underbrace{\sum_{w \in \mathcal{C}(T)} (-1)^w \, \overline{wT}}_{\substack{=\mathbf{e}_T \\ \text{(by (150))}}}$$

$$= \underbrace{(-1)^{u^{-1}}}_{\substack{=(-1)^u \\ \text{(by the basic properties} \\ \text{of signs of permutations)}}} \mathbf{e}_T = (-1)^u \, \mathbf{e}_T.$$

This proves Lemma 5.4.6 **(c)**.

**(d)** Let $T$ and $S$ be two column-equivalent $n$-tableaux of shape $D$. Then, Proposition 5.3.14 **(b)** shows that there exists some $w \in \mathcal{C}(T)$ such that $S = w \rightharpoonup T$. Consider this $w$. Then, $S = w \rightharpoonup T = wT$. But Lemma 5.4.6 **(c)** (applied to $u = w$) yields $\mathbf{e}_{wT} = (-1)^w \, \mathbf{e}_T$. In other words, $\mathbf{e}_S = (-1)^w \, \mathbf{e}_T$ (since $S = wT$). Hence, $\mathbf{e}_S = \pm \mathbf{e}_T$, so that $\mathbf{e}_T = \pm \mathbf{e}_S$. This proves Lemma 5.4.6 **(d)**. □

### 5.4.3. More examples

Our above examples show that the Specht modules $\mathcal{S}^{Y(\lambda)}$ for various partitions $\lambda$ include

- the trivial representation of $S_n$ (obtained for $\lambda = (n)$, as seen in Example 5.4.3),
- the sign representation of $S_n$ (obtained for $\lambda = (1^n)$, as seen in Example 5.4.4), and
- the zero-sum representation $R(\mathbf{k}^n)$ (obtained for $\lambda = (n-1, 1)$, as seen in Example 5.4.2).

For $n = 3$, all the irreducible representations of $S_n$ (over a field of characteristic 0) are contained in this list. For $n = 4$, there are others. Can we also identify them as Specht modules?

**Example 5.4.7.** Let $n = 4$ and $D = Y(2,2)$. This diagram $D$ has the form $\begin{array}{|c|c|}\hline \ & \ \\ \hline \ & \ \\ \hline \end{array}$.

There are six $n$-tabloids of shape $D$, corresponding bijectively to the 2-element subsets of $[4]$. Indeed, any $n$-tabloid of shape $D$ is uniquely determined by the set of the two entries in its second row, and will be denoted $\overline{\underline{i\ j}}$, where $i$ and $j$ are these two entries. For instance,

$$\overline{\underline{1\ 4}} = \overline{\begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & 4 \\ \hline \end{array}} = \overline{23\backslash\backslash 14}.$$

(This can also be written as $\overline{\underline{4\ 1}}$, since the entries of a row in an $n$-tabloid do not come ordered. But this is not the same $n$-tabloid as $\overline{\underline{2\ 3}}$.)

Thus, the Young module $\mathcal{M}^D$ is a free $\mathbf{k}$-module of rank 6, with basis $\left(\overline{\underline{1\ 2}},\ \overline{\underline{1\ 3}},\ \overline{\underline{1\ 4}},\ \overline{\underline{2\ 3}},\ \overline{\underline{2\ 4}},\ \overline{\underline{3\ 4}}\right)$. What is the Specht module $\mathcal{S}^D$ ? For any $n$-tableau $ij\backslash\backslash k\ell = \begin{array}{|c|c|}\hline i & j \\ \hline k & \ell \\ \hline \end{array}$, the corresponding polytabloid $\mathbf{e}_{ij\backslash\backslash k\ell}$ is

$$\begin{aligned}
\mathbf{e}_{ij\backslash\backslash k\ell} &= \sum_{w \in \mathcal{C}(ij\backslash\backslash k\ell)} (-1)^w \, \overline{w\,(ij\backslash\backslash k\ell)} \\
&= (-1)^{\mathrm{id}} \, \overline{\mathrm{id}\,(ij\backslash\backslash k\ell)} + (-1)^{t_{i,k}} \, \overline{t_{i,k}\,(ij\backslash\backslash k\ell)} \\
&\qquad + (-1)^{t_{j,\ell}} \, \overline{t_{j,\ell}\,(ij\backslash\backslash k\ell)} + (-1)^{t_{i,k}t_{j,\ell}} \, \overline{t_{i,k}t_{j,\ell}\,(ij\backslash\backslash k\ell)} \\
&\qquad\qquad \left(\text{since } \mathcal{C}\,(ij\backslash\backslash k\ell) = \left\{\mathrm{id},\ t_{i,k},\ t_{j,\ell},\ t_{i,k}t_{j,\ell}\right\}\right) \\
&= \overline{ij\backslash\backslash k\ell} - \overline{kj\backslash\backslash i\ell} - \overline{i\ell\backslash\backslash kj} + \overline{k\ell\backslash\backslash ij} \\
&= \overline{\underline{k\ \ell}} - \overline{\underline{i\ \ell}} - \overline{\underline{k\ j}} + \overline{\underline{i\ j}}.
\end{aligned}$$

Thus, the Specht module $\mathcal{S}^D$ is the span of all these elements $\overline{\underline{k\ \ell}} - \overline{\underline{i\ \ell}} - \overline{\underline{k\ j}} + \overline{\underline{i\ j}}$, where $(i,j,k,\ell)$ ranges over all 4-tuples of distinct elements of $[4]$. Many of these elements $\mathbf{e}_{ij\backslash\backslash k\ell}$ are equal (up to sign), but even the distinct ones are still linearly dependent. Their span $\mathcal{S}^D$ turns out to be only 2-dimensional, i.e., a free $\mathbf{k}$-module of rank 2. One basis of it is

$$\left(\mathbf{e}_{12\backslash\backslash 34},\ \mathbf{e}_{13\backslash\backslash 24}\right).$$

Do the two tableaux in the subscripts remind you of something? For the answer, look at Example 5.2.3.

**Exercise 5.4.1.** [2] Assume that $\mathbf{k}$ is a field. Show that the Specht module $\mathcal{S}^D$ in Example 5.4.7 is an irreducible representation of $S_4$ if and only if $\operatorname{char}\mathbf{k} \neq 3$.

[**Hint:** Use Exercise 5.3.3.]

We have not seen this representation before.

We could analyze the Specht module $\mathcal{S}^{Y(2,1,1)}$ similarly, although this would be more laborious (the Young module $\mathcal{M}^{Y(2,1,1)}$ is 12-dimensional). It is yet another irreducible representation of $S_4$ in characteristic 0. Do we have all of them now? We will see – but not very soon (Theorem 5.13.1). We have quite a way to go until we understand Specht modules well enough to answer this.

The partition $(2,2)$ in Example 5.4.7 is a particular case of the partition $(n-2,2)$, which exists for each $n \geq 4$. The Specht module corresponding to the latter partition is the subject of the following exercise:

**Exercise 5.4.2.** Let $n \geq 4$, and let $\lambda$ be the partition $(n-2,2)$. Let $D = Y(\lambda)$. As in Example 5.4.7, let us write $\overline{\underline{i\ j}}$ for the $n$-tabloid of shape $D$ whose second row has entries $i$ and $j$.

**(a)** $\boxed{1}$ Prove that

$$\mathcal{S}^D = \operatorname{span}_{\mathbf{k}} \left\{ \overline{\underline{k\ \ell}} - \overline{\underline{i\ \ell}} - \overline{\underline{k\ j}} + \overline{\underline{i\ j}} \ \mid\ i,j,k,\ell \in [n] \text{ distinct} \right\}.$$

**(b)** $\boxed{1}$ Let $E = Y(n-1,1)$. As in Example 5.3.16, let us write $\overline{\underline{i}}$ for the $n$-tabloid of shape $E$ whose second row has entry $i$. Consider the $\mathbf{k}$-linear map

$$\begin{aligned} \psi_1 : \mathcal{M}^D &\to \mathcal{M}^E, \\ \overline{\underline{i\ j}} &\mapsto \overline{\underline{i}} + \overline{\underline{j}} \end{aligned}$$

and the $\mathbf{k}$-linear map

$$\begin{aligned} \psi_0 : \mathcal{M}^D &\to \mathbf{k}, \\ \overline{\underline{i\ j}} &\mapsto 1. \end{aligned}$$

Prove that both these maps $\psi_1$ and $\psi_0$ are left $\mathbf{k}[S_n]$-linear (where $\mathbf{k}$ is equipped with the trivial $S_n$-action).

**(c)** $\boxed{4}$ Prove that $\mathcal{S}^D = (\operatorname{Ker}\psi_1) \cap (\operatorname{Ker}\psi_0)$.

**(d)** $\boxed{1}$ Prove that $\mathcal{S}^D = \operatorname{Ker}\psi_1$ if 2 is invertible in $\mathbf{k}$.

**(e)** $\boxed{1}$ Assume that both $n-1$ and $\dfrac{n(n-1)}{2}$ equal 0 in $\mathbf{k}$. Prove that $\mathcal{S}^D$ contains the element $\sum\limits_{1 \leq i < j \leq n} \overline{\underline{i\ j}}$ (this is the sum of all $n$-tabloids of shape $D$), and thus contains a trivial subrepresentation spanned by this element.

[**Hint:** For part **(c)**, first show that each element of $\mathcal{M}^D$ can be written as some element of $\mathcal{S}^D$ plus some element of the form $\sum\limits_{i=3}^{n} \alpha_i \overline{\underline{1\ i}} + \sum\limits_{i=3}^{n} \beta_i \overline{\underline{2\ i}} + \gamma \overline{\underline{1\ 2}}$ with $\alpha_i, \beta_i, \gamma \in \mathbf{k}$; then analyze the conditions under which an element of the latter form belongs to $\operatorname{Ker}\psi_1$ and to $\operatorname{Ker}\psi_0$.]

**Exercise 5.4.3.** [2] Let $D$ be any diagram. Prove that $\mathcal{S}^D = \mathcal{M}^D$ holds if and only if the ring $\mathbf{k}$ is trivial or the diagram $D$ has no two cells in the same column.

[**Hint:** In all other cases, find a nontrivial $\mathbf{k}$-linear map $f : \mathcal{M}^D \to \mathbf{k}$ that sends every polytabloid $\mathbf{e}_T$ to 0.]

## 5.5. Symmetrizers, antisymmetrizers and the left ideal avatar

We will now take a look at the Specht modules from a different direction. This relies on a few preliminary results about $\mathbf{k}[S_n]$.

### 5.5.1. Row symmetrizers and column antisymmetrizers

We begin by introducing some more elements of the group algebra $\mathbf{k}[S_n]$:

**Definition 5.5.1.** Let $T$ be an $n$-tableau of any shape $D$. We define its *row symmetrizer*

$$\nabla_{\mathrm{Row}\, T} := \sum_{w \in \mathcal{R}(T)} w \in \mathbf{k}[S_n]$$

and its *column antisymmetrizer*

$$\nabla^-_{\mathrm{Col}\, T} := \sum_{w \in \mathcal{C}(T)} (-1)^w w \in \mathbf{k}[S_n].$$

(We could similarly define the column symmetrizer $\nabla_{\mathrm{Col}\, T}$ and the row antisymmetrizer $\nabla^-_{\mathrm{Row}\, T}$, but we probably won't need them.)

**Example 5.5.2.** For the 5-tableau $12\backslash\backslash34\backslash\backslash5 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 \\ \cline{1-1}\end{array}$ of shape $Y(2,2,1)$, we have

$$\nabla_{\mathrm{Row}(12\backslash\backslash34\backslash\backslash5)} = 1 + t_{1,2} + t_{3,4} + t_{1,2}t_{3,4} \qquad \text{and}$$
$$\nabla^-_{\mathrm{Col}(12\backslash\backslash34\backslash\backslash5)} = 1 - t_{1,3} - t_{1,5} - t_{3,5} + \mathrm{cyc}_{1,3,5} + \mathrm{cyc}_{1,5,3} - t_{2,4} + t_{1,3}t_{2,4} + t_{1,5}t_{2,4} + t_{3,5}t_{2,4} - \mathrm{cyc}_{1,3,5}\, t_{2,4} - \mathrm{cyc}_{1,5,3}\, t_{2,4}.$$

The letter $\nabla$ was chosen because the definitions of $\nabla_{\mathrm{Row}\, T}$ and $\nabla^-_{\mathrm{Col}\, T}$ are rather similar to those of the $X$-integrals $\nabla_X$ and the $X$-sign-integrals $\nabla^-_X$ (see Definition 3.7.1). There is in fact a direct relation between the former and the latter:

**Proposition 5.5.3.** Let $T$ be an $n$-tableau of any shape $D$. Then:

**(a)** For any $i \in \mathbb{Z}$, let $\operatorname{Row}(i,T)$ denote the set of all entries in the $i$-th row of $T$. If $\ell \in \mathbb{N}$ is chosen such that all cells of $D$ lie in rows $1,2,\ldots,\ell$, then

$$\nabla_{\operatorname{Row} T} = \prod_{i=1}^{\ell} \nabla_{\operatorname{Row}(i,T)}.$$

Here we are not specifying the order of the factors in the product, since it does not matter (the factors commute).

**(b)** For any $j \in \mathbb{Z}$, let $\operatorname{Col}(j,T)$ denote the set of all entries in the $j$-th column of $T$. If $k \in \mathbb{N}$ is chosen such that all cells of $D$ lie in columns $1,2,\ldots,k$, then

$$\nabla^{-}_{\operatorname{Col} T} = \prod_{j=1}^{k} \nabla^{-}_{\operatorname{Col}(j,T)}.$$

Here, again, we are not specifying the order of the factors, since it does not matter.

**Example 5.5.4.** Let $n = 8$ and $T = 123\backslash\backslash45\backslash\backslash67\backslash\backslash8 =$ $\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2} 6 & 7 \\ \cline{1-2} 8 \\ \cline{1-1}\end{array}$. Then,

Proposition 5.5.3 **(a)** (for $\ell = 4$) yields

$$\nabla_{\operatorname{Row}(123\backslash\backslash45\backslash\backslash67\backslash\backslash8)} = \nabla_{\{1,2,3\}} \nabla_{\{4,5\}} \nabla_{\{6,7\}} \underbrace{\nabla_{\{8\}}}_{=1} = \nabla_{\{1,2,3\}} \nabla_{\{4,5\}} \nabla_{\{6,7\}}.$$

(We could also take $\ell$ larger than 4, but this would merely introduce extra factors $\nabla_{\operatorname{Row}(i,T)}$ corresponding to empty rows, and all these factors are $\nabla_{\varnothing} = 1$.)

Furthermore, Proposition 5.5.3 **(b)** (for $k = 3$) yields

$$\nabla^{-}_{\operatorname{Col}(123\backslash\backslash45\backslash\backslash67\backslash\backslash8)} = \nabla^{-}_{\{1,4,6,8\}} \nabla^{-}_{\{2,5,7\}} \underbrace{\nabla^{-}_{\{3\}}}_{=1}.$$

(Here, again, we could take $k$ larger than 3, but this would merely introduce factors of the form $\nabla^{-}_{\varnothing} = 1$ corresponding to empty columns.)

*Proof of Proposition 5.5.3.* **(b)** We know that $T$ is an $n$-tableau. Hence, the entries of $T$ are the numbers $1,2,\ldots,n$, and each of these numbers appears exactly once (i.e., the tableau $T$ has no two equal entries).

Let $k \in \mathbb{N}$ be such that all cells of $D$ lie in columns $1,2,\ldots,k$. Then, each entry of $T$ lies in exactly one of the columns $1,2,\ldots,k$ (indeed, not more than

one, since $T$ has no two equal entries). In other words, each entry of $T$ lies in exactly one of the sets $\operatorname{Col}(1,T)$, $\operatorname{Col}(2,T)$, $\ldots$, $\operatorname{Col}(k,T)$.

In other words, each of the numbers $1,2,\ldots,n$ lies in exactly one of the sets $\operatorname{Col}(1,T)$, $\operatorname{Col}(2,T)$, $\ldots$, $\operatorname{Col}(k,T)$ (since the entries of $T$ are the numbers $1,2,\ldots,n$). Thus, $\operatorname{Col}(1,T)$, $\operatorname{Col}(2,T)$, $\ldots$, $\operatorname{Col}(k,T)$ are $k$ disjoint subsets of $[n]$ and satisfy $\operatorname{Col}(1,T) \cup \operatorname{Col}(2,T) \cup \cdots \cup \operatorname{Col}(k,T) = [n]$. Hence, the equality (50) from Proposition 3.7.13 **(b)** (applied to $X_j = \operatorname{Col}(j,T)$) yields

$$\sum_{\substack{w \in S_n; \\ w(\operatorname{Col}(i,T)) \subseteq \operatorname{Col}(i,T) \text{ for all } i \in [k]}} (-1)^w w = \nabla^-_{\operatorname{Col}(1,T)} \nabla^-_{\operatorname{Col}(2,T)} \cdots \nabla^-_{\operatorname{Col}(k,T)}.$$

Moreover, Proposition 3.7.13 **(c)** (applied to $X_j = \operatorname{Col}(j,T)$) yields that the factors $\nabla^-_{\operatorname{Col}(1,T)}, \nabla^-_{\operatorname{Col}(2,T)}, \ldots, \nabla^-_{\operatorname{Col}(k,T)}$ commute. Thus, we can write their product as $\prod_{j=1}^{k} \nabla^-_{\operatorname{Col}(j,T)}$ without specifying the order of multiplication.

However, for any permutation $w \in S_n$, we have the logical equivalence

$$(w \in \mathcal{C}(T)) \iff (w(\operatorname{Col}(i,T)) \subseteq \operatorname{Col}(i,T) \text{ for all } i \in [k])$$

[53]. Hence,

$$\sum_{w \in \mathcal{C}(T)} (-1)^w w = \sum_{\substack{w \in S_n; \\ w(\mathrm{Col}(i,T)) \subseteq \mathrm{Col}(i,T) \text{ for all } i \in [k]}} (-1)^w w.$$

But the definition of $\nabla^-_{\mathrm{Col}\, T}$ yields

$$\nabla^-_{\mathrm{Col}\, T} = \sum_{w \in \mathcal{C}(T)} (-1)^w w = \sum_{\substack{w \in S_n; \\ w(\mathrm{Col}(i,T)) \subseteq \mathrm{Col}(i,T) \text{ for all } i \in [k]}} (-1)^w w$$

$$= \nabla^-_{\mathrm{Col}(1,T)} \nabla^-_{\mathrm{Col}(2,T)} \cdots \nabla^-_{\mathrm{Col}(k,T)} = \prod_{j=1}^{k} \nabla^-_{\mathrm{Col}(j,T)}.$$

This proves Proposition 5.5.3 **(b)**.

[53] *Proof.* Let $w \in S_n$. Then, we have the following chain of logical equivalences:

$(w \in \mathcal{C}(T))$

$\iff$ ($w$ is a vertical permutation for $T$) (by the definition of $\mathcal{C}(T)$)

$\iff$ (for each $i \in [n]$, the number $w(i)$ lies in the same column of $T$ as $i$)
(by the definition of vertical permutations)

$\iff$ (for each $i \in [n]$, for each $c \in \mathbb{Z}$, if $i$ lies in the $c$-th column of $T$, then so does $w(i)$)
(here, we just restated the "lies in the same column" part in words)

$\iff$ (for each $i \in [n]$, for each $c \in [k]$, if $i$ lies in the $c$-th column of $T$, then so does $w(i)$)
(here, we have restricted the index $c$ to the set $[k]$, since the only columns of $T$ that can possibly contain $i$ are the columns $1, 2, \ldots, k$ (because all cells of $D$ lie in columns $1, 2, \ldots, k$))

$\iff$ (for each $c \in [k]$, for each $i \in [n]$, if $i$ lies in the $c$-th column of $T$, then so does $w(i)$)
(here, we swapped the two "for all"-quantifiers)

$\iff$ (for each $c \in [k]$, for each $i \in [n]$, if $i \in \mathrm{Col}(c,T)$, then $w(i) \in \mathrm{Col}(c,T)$)
(since "$i$ lies in the $c$-th column of $T$" is equivalent to "$i \in \mathrm{Col}(c,T)$", and "$w(i)$ lies in the $c$-th column of $T$" is equivalent to "$w(i) \in \mathrm{Col}(c,T)$")

$\iff$ (for each $c \in [k]$, for each $i \in \mathrm{Col}(c,T)$, we have $w(i) \in \mathrm{Col}(c,T)$)
(since the elements $i \in [n]$ that satisfy $i \in \mathrm{Col}(c,T)$ are precisely the elements of $\mathrm{Col}(c,T)$ (because $\mathrm{Col}(c,T) \subseteq [n]$))

$\iff$ (for each $c \in \mathbb{Z}$, we have $w(\mathrm{Col}(c,T)) \subseteq \mathrm{Col}(c,T)$)
(since the statement "for each $i \in [n]$, we have $w(i) \in \mathrm{Col}(c,T)$" is equivalent to "$w(\mathrm{Col}(c,T)) \subseteq \mathrm{Col}(c,T)$")

$\iff$ ($w(\mathrm{Col}(c,T)) \subseteq \mathrm{Col}(c,T)$ for all $c \in [k]$)

$\iff$ ($w(\mathrm{Col}(i,T)) \subseteq \mathrm{Col}(i,T)$ for all $i \in [k]$) (here, we renamed the index $c$ as $i$).

**(a)** Analogous to part **(b)**, but without the signs. □

Some other properties of row symmetrizers and column antisymmetrizers also resemble analogous properties of integrals and sign-integrals:

**Proposition 5.5.5.** Let $T$ be an $n$-tableau of any shape $D$. Then:

**(a)** We have $w\nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T} w = \nabla_{\operatorname{Row} T}$ for each $w \in \mathcal{R}(T)$.

**(b)** We have $w\nabla^{-}_{\operatorname{Col} T} = \nabla^{-}_{\operatorname{Col} T} w = (-1)^{w} \nabla^{-}_{\operatorname{Col} T}$ for each $w \in \mathcal{C}(T)$.

*Proof.* Analogous to Proposition 3.3.3 (but using the group $\mathcal{R}(T)$ or $\mathcal{C}(T)$ instead of $S_n$). □

**Corollary 5.5.6.** Let $T$ be an $n$-tableau of any shape $D$. Then:

**(a)** We have $\left(\nabla_{\operatorname{Row} T}\right)^2 = |\mathcal{R}(T)| \cdot \nabla_{\operatorname{Row} T}$.

**(b)** We have $\left(\nabla^{-}_{\operatorname{Col} T}\right)^2 = |\mathcal{C}(T)| \cdot \nabla^{-}_{\operatorname{Col} T}$.

*Proof.* Each of the two parts is proved similarly to Corollary 3.3.4, but using Proposition 5.5.5 instead of Proposition 3.3.3. (Part **(b)** involves signs as well, but they do not complicate the computation.) □

We note that the factors $|\mathcal{R}(T)|$ and $|\mathcal{C}(T)|$ in Corollary 5.5.6 can be computed rather explicitly:

**Proposition 5.5.7.** Let $T$ be an $n$-tableau of any shape $D$. Then:

**(a)** For any $i \in \mathbb{Z}$, let $\operatorname{Row}(i, T)$ denote the set of all entries in the $i$-th row of $T$. If $\ell \in \mathbb{N}$ is chosen such that all cells of $D$ lie in rows $1, 2, \ldots, \ell$, then

$$|\mathcal{R}(T)| = \prod_{i=1}^{\ell} |\operatorname{Row}(i, T)|!.$$

**(b)** For any $j \in \mathbb{Z}$, let $\operatorname{Col}(j, T)$ denote the set of all entries in the $j$-th column of $T$. If $k \in \mathbb{N}$ is chosen such that all cells of $D$ lie in columns $1, 2, \ldots, k$, then

$$|\mathcal{C}(T)| = \prod_{j=1}^{k} |\operatorname{Col}(j, T)|!.$$

*Proof.* **(a)** This is intuitively obvious, but the slickest proof uses a bit of algebra: Recall the counit $\varepsilon : \mathbf{k}[S_n] \to \mathbf{k}$ (which was defined in Definition 3.12.1). This counit $\varepsilon$ is a $\mathbf{k}$-algebra morphism (by Theorem 3.12.3). Now, let $\ell \in \mathbb{N}$ be chosen such that all cells of $D$ lie in rows $1, 2, \ldots, \ell$. Then, Proposition 5.5.3 **(a)** yields

$$\nabla_{\operatorname{Row} T} = \prod_{i=1}^{\ell} \nabla_{\operatorname{Row}(i,T)}.$$

Applying the map $\varepsilon$ to both sides of this equality, we find

$$\begin{aligned}\varepsilon\left(\nabla_{\operatorname{Row} T}\right) &= \varepsilon\left(\prod_{i=1}^{\ell} \nabla_{\operatorname{Row}(i,T)}\right) \\ &= \prod_{i=1}^{\ell} \underbrace{\varepsilon\left(\nabla_{\operatorname{Row}(i,T)}\right)}_{\substack{=|\operatorname{Row}(i,T)|! \\ \text{(by Proposition 3.12.6,} \\ \text{applied to } X=\operatorname{Row}(i,T))}} \qquad \text{(since } \varepsilon \text{ is a } \mathbf{k}\text{-algebra morphism)} \\ &= \prod_{i=1}^{\ell} |\operatorname{Row}(i,T)|!.\end{aligned}$$

Comparing this with

$$\begin{aligned}\varepsilon\left(\nabla_{\operatorname{Row} T}\right) &= \varepsilon\left(\sum_{w \in \mathcal{R}(T)} w\right) \qquad \left(\text{since } \nabla_{\operatorname{Row} T} = \sum_{w \in \mathcal{R}(T)} w\right) \\ &= \sum_{w \in \mathcal{R}(T)} \underbrace{\varepsilon(w)}_{\substack{=1 \\ \text{(by the definition of } \varepsilon, \\ \text{since } w \in \mathcal{R}(T) \subseteq S_n)}} \qquad \text{(since the map } \varepsilon \text{ is } \mathbf{k}\text{-linear)} \\ &= \sum_{w \in \mathcal{R}(T)} 1 = |\mathcal{R}(T)| \cdot 1 = |\mathcal{R}(T)|,\end{aligned}$$

we obtain $|\mathcal{R}(T)| = \prod_{i=1}^{\ell} |\operatorname{Row}(i,T)|!$. This proves Proposition 5.5.7 **(a)**.

**(b)** This is analogous to part **(a)**. Namely, we have to define an element $\nabla_{\operatorname{Col} T} := \sum_{w \in \mathcal{C}(T)} w$ of $\mathbf{k}[S_n]$, and we have to show that this element satisfies $\nabla_{\operatorname{Col} T} = \prod_{j=1}^{k} \nabla_{\operatorname{Col}(j,T)}$ (this is an analogue of Proposition 5.5.3 **(a)**, and can be proved in the same way). Then, we have to apply $\varepsilon$ as in our above proof of Proposition 5.5.7 **(a)**. □

Next, let us show some ways to factor out simple factors from $\nabla_{\operatorname{Row} T}$ and $\nabla^{-}_{\operatorname{Col} T}$, similarly to Proposition 3.7.4 **(d)**:

**Proposition 5.5.8.** Let $T$ be an $n$-tableau of any shape $D$. Let $i$ and $j$ be two distinct elements of $[n]$ that lie in the same row of $T$. Then:

**(a)** The transposition $t_{i,j}$ belongs to $\mathcal{R}(T)$.

**(b)** We have

$$\nabla_{\operatorname{Row} T} = \nabla^{\text{even}}_{\operatorname{Row} T} \cdot \left(1 + t_{i,j}\right) = \left(1 + t_{i,j}\right) \cdot \nabla^{\text{even}}_{\operatorname{Row} T},$$

where we set

$$\nabla^{\text{even}}_{\operatorname{Row} T} := \sum_{\substack{w \in \mathcal{R}(T); \\ (-1)^w = 1}} w.$$

**(c)** We have $t_{i,j} \nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T} t_{i,j} = \nabla_{\operatorname{Row} T}$.

*Proof.* **(a)** It is easy to see that the transposition $t_{i,j} \in S_n$ is a horizontal permutation for $T$. (*Proof:* We must show that for each $k \in [n]$, the number $t_{i,j}(k)$ lies in the same row of $T$ as $k$ does. This is true for $k = i$ (since $t_{i,j}(i) = j$ lies in the same row of $T$ as $i$ does) and also true for $k = j$ (for similar reasons), but it is even more obvious for all remaining values of $k$ (since $t_{i,j}(k) = k$ for all those values). Thus, this is true for all $k \in [n]$. It follows that the permutation $t_{i,j}$ is horizontal for $T$.)

So we have shown that $t_{i,j}$ is a horizontal permutation for $T$. In other words, $t_{i,j}$ belongs to $\mathcal{R}(T)$ (since $\mathcal{R}(T)$ is the set of all permutations $w \in S_n$ that are horizontal for $T$). This proves Proposition 5.5.8 **(a)**.

**(b)** We know that $\mathcal{R}(T)$ is a group, and $t_{i,j}$ is one of its elements (by Proposition 5.5.8 **(a)**). From here on, the proof of Proposition 5.5.8 **(b)** proceeds exactly like the proof of Proposition 3.7.4 **(d)** (more precisely, the proof of (31)), except that $S_{n,X}$, $\nabla_X$ and $\nabla_X^{\text{even}}$ must be replaced by $\mathcal{R}(T)$, $\nabla_{\text{Row}\,T}$ and $\nabla_{\text{Row}\,T}^{\text{even}}$. Thus, Proposition 5.5.8 **(b)** is proved.

**(c)** Proposition 5.5.8 **(a)** shows that $t_{i,j} \in \mathcal{R}(T)$. Thus, Proposition 5.5.5 **(a)** (applied to $w = t_{i,j}$) yields $t_{i,j}\nabla_{\text{Row}\,T} = \nabla_{\text{Row}\,T} t_{i,j} = \nabla_{\text{Row}\,T}$. This proves Proposition 5.5.8 **(c)**. □

**Proposition 5.5.9.** Let $T$ be an $n$-tableau of any shape $D$. Let $i$ and $j$ be two distinct elements of $[n]$ that lie in the same column of $T$. Then:

**(a)** The transposition $t_{i,j}$ belongs to $\mathcal{C}(T)$.

**(b)** We have

$$\nabla_{\text{Col}\,T}^{-} = \nabla_{\text{Col}\,T}^{\text{even}} \cdot \left(1 - t_{i,j}\right) = \left(1 - t_{i,j}\right) \cdot \nabla_{\text{Col}\,T}^{\text{even}},$$

where we set

$$\nabla_{\text{Col}\,T}^{\text{even}} := \sum_{\substack{w \in \mathcal{C}(T); \\ (-1)^w = 1}} w.$$

**(c)** We have $t_{i,j}\nabla_{\text{Col}\,T} = \nabla_{\text{Col}\,T} t_{i,j} = -\nabla_{\text{Col}\,T}$.

*Proof.* This is analogous to Proposition 5.5.8, except that the signs need to be taken into account (which is easy, since we know that $(-1)^{t_{i,j}} = -1$). □

The antipode $S$ preserves both row symmetrizers and column antisymmetrizers:

**Proposition 5.5.10.** Let $T$ be an $n$-tableau of any shape $D$. Then, the antipode $S$ of $\mathbf{k}[S_n]$ satisfies

$$S\left(\nabla_{\text{Row}\,T}\right) = \nabla_{\text{Row}\,T} \qquad \text{and} \qquad S\left(\nabla_{\text{Col}\,T}^{-}\right) = \nabla_{\text{Col}\,T}^{-}.$$

*Proof.* The proof of $S\left(\nabla^-_{\operatorname{Col} T}\right) = \nabla^-_{\operatorname{Col} T}$ is analogous to the proof of $S\left(\nabla^-\right) = \nabla^-$ in Example 3.11.13 **(c)** (indeed, $\mathcal{C}(T)$ is a group, so that the map $\mathcal{C}(T) \to \mathcal{C}(T)$, $w \mapsto w^{-1}$ is a bijection). The proof of $S\left(\nabla_{\operatorname{Row} T}\right) = \nabla_{\operatorname{Row} T}$ is analogous but easier. □

Next, let us see how row symmetrizers and column antisymmetrizers are affected when a permutation $w \in S_n$ acts on an $n$-tableau $T$:

**Proposition 5.5.11.** Let $T$ be an $n$-tableau of any shape $D$. Let $w \in S_n$. Then,

$$\nabla_{\operatorname{Row}(w \rightharpoonup T)} = w \nabla_{\operatorname{Row} T} w^{-1} \qquad \text{and}$$
$$\nabla^-_{\operatorname{Col}(w \rightharpoonup T)} = w \nabla^-_{\operatorname{Col} T} w^{-1}.$$

*Proof.* Proposition 5.3.12 **(c)** yields $\mathcal{R}(w \rightharpoonup T) = w\mathcal{R}(T) w^{-1}$ and $\mathcal{C}(w \rightharpoonup T) = w\mathcal{C}(T) w^{-1}$. The latter equality shows that the map

$$\mathcal{C}(T) \to \mathcal{C}(w \rightharpoonup T),$$
$$u \mapsto wuw^{-1}$$

is well-defined and surjective. This map is also injective[54], and thus is a bijection.

However, the definition of $\nabla^-_{\operatorname{Col} T}$ says that $\nabla^-_{\operatorname{Col} T} = \sum_{u \in \mathcal{C}(T)} (-1)^u u$. For the same reason, we have

$$\nabla^-_{\operatorname{Col}(w \rightharpoonup T)} = \sum_{u \in \mathcal{C}(w \rightharpoonup T)} (-1)^u u = \sum_{u \in \mathcal{C}(T)} \underbrace{(-1)^{wuw^{-1}}}_{\substack{=(-1)^u \\ \text{(by (10))}}} wuw^{-1}$$
$$\left( \begin{array}{c} \text{here, we have substituted } wuw^{-1} \text{ for } u \text{ in the sum,} \\ \text{since the map } \mathcal{C}(T) \to \mathcal{C}(w \rightharpoonup T),\ u \mapsto wuw^{-1} \\ \text{is a bijection} \end{array} \right)$$
$$= \sum_{u \in \mathcal{C}(T)} (-1)^u wuw^{-1} = w \underbrace{\left( \sum_{u \in \mathcal{C}(T)} (-1)^u u \right)}_{=\nabla^-_{\operatorname{Col} T}} w^{-1} = w \nabla^-_{\operatorname{Col} T} w^{-1}.$$

A similar computation (but simpler due to the absence of signs) shows that $\nabla_{\operatorname{Row}(w \rightharpoonup T)} = w \nabla_{\operatorname{Row} T} w^{-1}$. Proposition 5.5.11 is thus proved. □

We can use column antisymmetrizers to rewrite the definition of polytabloids in the following compact form:

[54] Indeed, if two elements $u, v \in \mathcal{C}(T)$ satisfy $wuw^{-1} = wvw^{-1}$, then $u = v$ (because $S_n$ is a group, and thus we can cancel both $w$ and $w^{-1}$ from the equality $wuw^{-1} = wvw^{-1}$).

**Proposition 5.5.12.** Let $T$ be an $n$-tableau of any shape $D$. Then, in the Young module $\mathcal{M}^D$, we have

$$\mathbf{e}_T = \nabla^-_{\operatorname{Col} T} \overline{T}.$$

*Proof.* The definition of $\nabla^-_{\operatorname{Col} T}$ yields $\nabla^-_{\operatorname{Col} T} = \sum_{w \in \mathcal{C}(T)} (-1)^w w$. Thus,

$$\nabla^-_{\operatorname{Col} T} \overline{T} = \sum_{w \in \mathcal{C}(T)} (-1)^w \underbrace{w\overline{T}}_{\substack{= \overline{wT} \\ \text{(by (145))}}} = \sum_{w \in \mathcal{C}(T)} (-1)^w \overline{wT} = \mathbf{e}_T$$

(by the definition of $\mathbf{e}_T$). This proves Proposition 5.5.12. □

### 5.5.2. The left ideal avatars of Young and Specht modules

Using row symmetrizers and column antisymmetrizers, we can find new avatars (i.e., isomorphic copies) of the Young and Specht modules as subrepresentations of the left regular representation $\mathbf{k}[S_n]$:

**Theorem 5.5.13** (left ideal avatars of Young and Specht modules)**.** Let $D$ be any diagram with $|D| = n$. Let $T$ be an $n$-tableau of shape $D$. Then:

**(a)** The Young module $\mathcal{M}^D$ is isomorphic (as an $S_n$-representation) to the left ideal

$$\mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T} = \{\mathbf{a} \nabla_{\operatorname{Row} T} \mid \mathbf{a} \in \mathbf{k}[S_n]\}$$

of $\mathbf{k}[S_n]$. To be more specific: There is a unique left $\mathbf{k}[S_n]$-linear map

$$\alpha : \mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T} \to \mathcal{M}^D$$

that sends $\nabla_{\operatorname{Row} T}$ to $\overline{T}$, and this map is an isomorphism of $S_n$-representations (i.e., of left $\mathbf{k}[S_n]$-modules).

**(b)** This isomorphism $\alpha$ sends the submodule

$$\mathbf{k}[S_n] \cdot \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} = \left\{\mathbf{a} \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \mid \mathbf{a} \in \mathbf{k}[S_n]\right\}$$

of $\mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T}$ to the Specht module $\mathcal{S}^D$. Thus, the Specht module $\mathcal{S}^D$ is isomorphic to the left ideal $\mathbf{k}[S_n] \cdot \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$ of $\mathbf{k}[S_n]$.

The product $\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$ in Theorem 5.5.13 **(b)** is often called the *Young symmetrizer* of $T$ (although sometimes the word is used instead for a certain scalar multiple of it, which we will see later). We will denote it by $\mathbf{E}_T$ in a later section.

**Example 5.5.14.** Let $n = 5$ and $D = Y(2,2,1)$ and $T = 12\backslash\backslash 34\backslash\backslash 5$. Then,

$$\nabla^-_{\operatorname{Col} T} = \nabla^-_{\{1,3,5\}} \nabla^-_{\{2,4\}} \qquad \text{and}$$
$$\nabla_{\operatorname{Row} T} = \nabla_{\{1,2\}} \nabla_{\{3,4\}} \underbrace{\nabla_{\{5\}}}_{=1} = \nabla_{\{1,2\}} \nabla_{\{3,4\}},$$

so we obtain the left ideal

$$\mathbf{k}[S_n] \cdot \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} = \mathbf{k}[S_5] \cdot \nabla^-_{\{1,3,5\}} \nabla^-_{\{2,4\}} \nabla_{\{1,2\}} \nabla_{\{3,4\}}.$$

Theorem 5.5.13 **(b)** claims that this left ideal of $\mathbf{k}[S_5]$ is isomorphic to the Specht module $\mathcal{S}^D = \mathcal{S}^{Y(2,2,1)}$. Note that taking a different $n$-tableau $T$ of shape $D$ would yield a different but isomorphic left ideal. Thus, Theorem 5.5.13 **(b)** produces not one, but many different isomorphic copies of a Specht module inside $\mathbf{k}[S_n]$. And Theorem 5.5.13 **(a)** does likewise for Young modules.

There are many ways to prove Theorem 5.5.13 (note that part **(b)** follows easily from part **(a)**, but part **(a)** presents some – at least notational – difficulties). In particular, we could generalize it to arbitrary groups, since the Young module $\mathcal{M}^D$ is the permutation module corresponding to a left $S_n$-set that can be identified with a quotient of $S_n$. However, we shall take a different road, which will net us some helpful results along the way that will be used later on.

### 5.5.3. The row-to-row sums

This road leads us through what we call the *row-to-row sums*:

**Definition 5.5.15.** Let $D$ be any diagram. Let $\overline{S}$ and $\overline{T}$ be two $n$-tabloids of shape $D$. Then, we define the *row-to-row sum* $\nabla_{\overline{S},\overline{T}} \in \mathbf{k}[S_n]$ by

$$\nabla_{\overline{S},\overline{T}} := \sum_{\substack{u \in S_n; \\ u\overline{T} = \overline{S}}} u.$$

(Of course, we could just as well rewrite $u\overline{T}$ as $u \rightharpoonup \overline{T}$ or as $\overline{uT}$ here.)

**Example 5.5.16.** Let $n = 4$ and $D = Y(2,2)$ and $\overline{S} = \overline{\begin{array}{cc} 2 & 4 \\ 1 & 3 \end{array}}$ and $\overline{T} = \overline{\begin{array}{cc} 1 & 2 \\ 3 & 4 \end{array}}$.

Then,

$$\nabla_{\overline{S},\overline{T}} = \sum_{\substack{u \in S_n; \\ u\overline{T} = \overline{S}}} u = \sum_{\substack{u \in S_n; \\ u(\{1,2\}) = \{2,4\}; \\ u(\{3,4\}) = \{1,3\}}} u$$
$$= \operatorname{oln}(2413) + \operatorname{oln}(2431) + \operatorname{oln}(4213) + \operatorname{oln}(4231).$$

(See Subsection 2.4.2 for the meaning of oln.)

We begin by proving some easy properties of these elements. First, we observe that row-to-row sums include row symmetrizers as a particular case:

**Proposition 5.5.17.** Let $D$ be any diagram. Let $\overline{T}$ be any $n$-tabloid of shape $D$. Then,

$$\nabla_{\overline{T},\overline{T}} = \nabla_{\operatorname{Row} T}.$$

*Proof.* Definition 5.5.1 yields $\nabla_{\operatorname{Row} T} = \sum_{w \in \mathcal{R}(T)} w$. But Proposition 5.3.13 **(a)** shows that

$$\mathcal{R}(T) = \left\{ w \in S_n \mid \overline{w \rightharpoonup T} = \overline{T} \right\} = \left\{ w \in S_n \mid w\overline{T} = \overline{T} \right\}$$

(since $\overline{w \rightharpoonup T} = w \rightharpoonup \overline{T} = w\overline{T}$ for each $w \in S_n$). Thus, the summation sign $\sum_{w \in \mathcal{R}(T)}$ can be rewritten as $\sum_{\substack{w \in S_n; \\ w\overline{T} = \overline{T}}}$. Hence,

$$\sum_{w \in \mathcal{R}(T)} w = \sum_{\substack{w \in S_n; \\ w\overline{T} = \overline{T}}} w = \sum_{\substack{u \in S_n; \\ u\overline{T} = \overline{T}}} u$$

(here, we have renamed the summation index $w$ as $u$). Therefore,

$$\nabla_{\operatorname{Row} T} = \sum_{w \in \mathcal{R}(T)} w = \sum_{\substack{u \in S_n; \\ u\overline{T} = \overline{T}}} u = \nabla_{\overline{T},\overline{T}}$$

(since $\nabla_{\overline{T},\overline{T}}$ was defined as $\sum_{\substack{u \in S_n; \\ u\overline{T} = \overline{T}}} u$). This proves Proposition 5.5.17. □

The next property is a kind of equivariance:

**Proposition 5.5.18.** Let $D$ be any diagram. Let $\overline{S}$ and $\overline{T}$ be two $n$-tabloids of shape $D$. Let $g \in S_n$ and $h \in S_n$ be two permutations. Then,

$$\nabla_{g\overline{S},h\overline{T}} = g \nabla_{\overline{S},\overline{T}} h^{-1}.$$

*Proof.* Recall that $S_n$ is a group. Thus, the map

$$\begin{aligned} S_n &\to S_n, \\ u &\mapsto guh^{-1} \end{aligned}$$

is a bijection (with inverse given by $u \mapsto g^{-1}uh$). The definition of $\nabla_{\overline{S},\overline{T}}$ yields

$$\nabla_{\overline{S},\overline{T}} = \sum_{\substack{u \in S_n; \\ u\overline{T} = \overline{S}}} u. \tag{151}$$

The same definition (but applied to $g\overline{S}$ and $h\overline{T}$ instead of $\overline{S}$ and $\overline{T}$) yields

$$\nabla_{g\overline{S},h\overline{T}} = \sum_{\substack{u \in S_n; \\ uh\overline{T} = g\overline{S}}} u = \sum_{\substack{u \in S_n; \\ guh^{-1}h\overline{T} = g\overline{S}}} guh^{-1} \tag{152}$$

(here, we have substituted $guh^{-1}$ for $u$ in the sum, since the map $S_n \to S_n$, $u \mapsto guh^{-1}$ is a bijection). However, for any permutation $u \in S_n$, we have the chain of equivalences

$$\left( g \underbrace{uh^{-1}h}_{=u} \overline{T} = g\overline{S} \right) \iff \left( gu\overline{T} = g\overline{S} \right) \iff \left( u\overline{T} = \overline{S} \right)$$

(since the statement $\left(u\overline{T} = \overline{S}\right)$ implies $\left(gu\overline{T} = g\overline{S}\right)$ by left-multiplying it by $g$, whereas conversely, the statement $\left(gu\overline{T} = g\overline{S}\right)$ implies $\left(u\overline{T} = \overline{S}\right)$ by left-multiplying it by $g^{-1}$). Hence, we can replace the condition "$guh^{-1}h\overline{T} = g\overline{S}$" under the summation sign in (152) by "$u\overline{T} = \overline{S}$". Thus, (152) rewrites as

$$\nabla_{g\overline{S},h\overline{T}} = \sum_{\substack{u \in S_n; \\ u\overline{T} = \overline{S}}} guh^{-1} = g \underbrace{\left( \sum_{\substack{u \in S_n; \\ u\overline{T} = \overline{S}}} u \right)}_{\substack{= \nabla_{\overline{S},\overline{T}} \\ \text{(by (151))}}} h^{-1} = g\nabla_{\overline{S},\overline{T}}h^{-1}.$$

This proves Proposition 5.5.18. □

**Corollary 5.5.19.** Let $D$ be any diagram. Let $\overline{T}$ be an $n$-tabloid of shape $D$. Let $w \in S_n$. Then, $\nabla_{w\overline{T},\overline{T}} = w\nabla_{\operatorname{Row} T}$.

*Proof.* Proposition 5.5.18 (applied to $\overline{S} = \overline{T}$, $g = w$ and $h = \mathrm{id}$) yields $\nabla_{w\overline{T},\mathrm{id}\,\overline{T}} = w\nabla_{\overline{T},\overline{T}} \underbrace{\mathrm{id}^{-1}}_{=\mathrm{id}} = w \underbrace{\nabla_{\overline{T},\overline{T}}}_{\substack{= \nabla_{\operatorname{Row} T} \\ \text{(by Proposition 5.5.17)}}} = w\nabla_{\operatorname{Row} T}$. In view of $\mathrm{id}\,\overline{T} = \overline{T}$, we can rewrite this as $\nabla_{w\overline{T},\overline{T}} = w\nabla_{\operatorname{Row} T}$. This proves Corollary 5.5.19. □

**Corollary 5.5.20.** Let $D$ be any diagram. Let $\overline{S}$ and $\overline{T}$ be two $n$-tabloids of shape $D$. Then, $\nabla_{\overline{S},\overline{T}} \in \mathbf{k}\left[S_n\right] \cdot \nabla_{\operatorname{Row} T}$.

*Proof.* Proposition 5.3.13 **(c)** shows that the $n$-tabloid $\overline{S}$ can be written as $w \rightharpoonup \overline{T}$ for some $w \in S_n$. Consider this $w$. Thus, $\overline{S} = w \rightharpoonup \overline{T} = w\overline{T}$, so that

$$\begin{aligned}\nabla_{\overline{S},\overline{T}} = \nabla_{w\overline{T},\overline{T}} &= \underbrace{w}_{\in \mathbf{k}[S_n]} \nabla_{\mathrm{Row}\, T} \qquad \text{(by Corollary 5.5.19)} \\ &\in \mathbf{k}\left[S_n\right] \cdot \nabla_{\mathrm{Row}\, T}.\end{aligned}$$

This proves Corollary 5.5.20. □

### 5.5.4. Proving the left ideal avatars

Now, we shall define a map $\zeta : \mathcal{M}^D \to \mathbf{k}\left[S_n\right] \cdot \nabla_{\mathrm{Row}\, T}$ that will later turn out to be inverse to the map $\alpha$ in Theorem 5.5.13 **(a)**:

**Proposition 5.5.21.** Let $D$ be any diagram with $|D| = n$. Let $T$ be an $n$-tableau of shape $D$. Consider the $\mathbf{k}$-linear map

$$\zeta : \mathcal{M}^D \to \mathbf{k}\left[S_n\right] \cdot \nabla_{\mathrm{Row}\, T}$$

that sends each $n$-tabloid $\overline{S}$ of shape $D$ to $\nabla_{\overline{S},\overline{T}}$. This map $\zeta$ is well-defined and is a left $\mathbf{k}\left[S_n\right]$-module isomorphism.

*Proof.* Let us first show that the map $\zeta$ is well-defined. First of all, we recall that the $n$-tabloids form a basis of the $\mathbf{k}$-module $\mathcal{M}^D$ (by the definition of $\mathcal{M}^D$). Hence, we can uniquely define a $\mathbf{k}$-linear map from $\mathcal{M}^D$ by specifying its values on the $n$-tabloids. Thus, in particular, the definition of $\zeta$ given in Proposition 5.5.21 is legitimate, once we show that the value $\nabla_{\overline{S},\overline{T}}$ really belongs to $\mathbf{k}\left[S_n\right] \cdot \nabla_{\mathrm{Row}\, T}$ for each $n$-tabloid $\overline{S}$ of shape $D$. But the latter claim is just Corollary 5.5.20. Thus, the definition of $\zeta$ is legitimate – i.e., the map $\zeta$ is well-defined.

Next, let us show some further properties of $\zeta$:

*Claim 1:* For each $w \in S_n$, we have $\zeta\left(\overline{wT}\right) = w\nabla_{\mathrm{Row}\, T}$.

*Proof of Claim 1.* Let $w \in S_n$. The definition of $\zeta$ yields

$$\begin{aligned}\zeta\left(\overline{wT}\right) &= \nabla_{\overline{wT},\overline{T}} = \nabla_{w\overline{T},\overline{T}} \qquad \left(\text{since } \overline{wT} = w\overline{T}\right) \\ &= w\nabla_{\mathrm{Row}\, T} \qquad \left(\text{by Corollary 5.5.19}\right).\end{aligned}$$

This proves Claim 1. □

*Claim 2:* The map $\zeta$ is left $\mathbf{k}\left[S_n\right]$-linear.

*Proof of Claim 2.* Let $w \in S_n$, and let $\mathbf{a} \in \mathcal{M}^D$. We shall show that $\zeta\left(w\mathbf{a}\right) = w\zeta\left(\mathbf{a}\right)$.

Indeed, both sides of this equality are $\mathbf{k}$-linear in $\mathbf{a}$. Hence, we can WLOG assume that $\mathbf{a}$ is an $n$-tabloid of shape $D$ (since the $\mathbf{k}$-module $\mathcal{M}^D$ is spanned

by such $n$-tabloids). Assume this. Thus, $\mathbf{a} = \overline{S}$ for some $n$-tabloid $\overline{S}$ of shape $D$. Consider this $\overline{S}$. From $\mathbf{a} = \overline{S}$, we obtain $\zeta(\mathbf{a}) = \zeta\left(\overline{S}\right) = \nabla_{\overline{S},\overline{T}}$ (by the definition of $\zeta$). Thus,

$$w\zeta(\mathbf{a}) = w\nabla_{\overline{S},\overline{T}}. \tag{153}$$

However, from $\mathbf{a} = \overline{S}$, we also obtain $w\mathbf{a} = w\overline{S} = \overline{wS}$ and therefore

$$\begin{aligned}
\zeta(w\mathbf{a}) &= \zeta\left(\overline{wS}\right) = \nabla_{\overline{wS},\overline{T}} && \text{(by the definition of } \zeta\text{)} \\
&= \nabla_{w\overline{S},\mathrm{id}\,\overline{T}} && \left(\text{since } \overline{wS} = w\overline{S} \text{ and } \overline{T} = \mathrm{id}\,\overline{T}\right) \\
&= w\nabla_{\overline{S},\overline{T}}\,\underbrace{\mathrm{id}^{-1}}_{=\mathrm{id}=1} && \left(\begin{array}{c}\text{by Proposition 5.5.18,}\\ \text{applied to } g = w \text{ and } h = \mathrm{id}\end{array}\right) \\
&= w\nabla_{\overline{S},\overline{T}} = w\zeta(\mathbf{a}) && \text{(by (153))}.
\end{aligned}$$

Forget that we fixed $w$ and $\mathbf{a}$. We thus have proved that $\zeta(w\mathbf{a}) = w\zeta(\mathbf{a})$ for all $w \in S_n$ and all $\mathbf{a} \in \mathcal{M}^D$. In other words, the map $\zeta$ is $S_n$-equivariant. Hence, $\zeta$ is a morphism of $S_n$-representations (since $\zeta$ is $\mathbf{k}$-linear). By Proposition 4.3.18, this entails that $\zeta$ is a left $\mathbf{k}[S_n]$-module morphism, i.e., a left $\mathbf{k}[S_n]$-linear map. This proves Claim 2. □

*Claim 3:* The map $\zeta$ is surjective.

*Proof of Claim 3.* Let $\mathbf{a} \in \mathbf{k}[S_n]$. We shall show that $\mathbf{a}\nabla_{\mathrm{Row}\,T} \in \zeta\left(\mathcal{M}^D\right)$.

Indeed, the definition of $\zeta$ yields $\zeta\left(\overline{T}\right) = \nabla_{\overline{T},\overline{T}} = \nabla_{\mathrm{Row}\,T}$ (by Proposition 5.5.17). However, the map $\zeta$ is left $\mathbf{k}[S_n]$-linear (by Claim 2), and thus we have $\zeta\left(\mathbf{a}\overline{T}\right) = \mathbf{a}\,\underbrace{\zeta\left(\overline{T}\right)}_{=\nabla_{\mathrm{Row}\,T}} = \mathbf{a}\nabla_{\mathrm{Row}\,T}$. Thus, $\mathbf{a}\nabla_{\mathrm{Row}\,T} = \zeta\left(\mathbf{a}\overline{T}\right) \in \zeta\left(\mathcal{M}^D\right)$.

Forget that we fixed $\mathbf{a}$. We thus have shown that $\mathbf{a}\nabla_{\mathrm{Row}\,T} \in \zeta\left(\mathcal{M}^D\right)$ for each $\mathbf{a} \in \mathbf{k}[S_n]$. In other words, $\{\mathbf{a}\nabla_{\mathrm{Row}\,T} \mid \mathbf{a} \in \mathbf{k}[S_n]\} \subseteq \zeta\left(\mathcal{M}^D\right)$. Now,

$$\mathbf{k}[S_n] \cdot \nabla_{\mathrm{Row}\,T} = \{\mathbf{a}\nabla_{\mathrm{Row}\,T} \mid \mathbf{a} \in \mathbf{k}[S_n]\} \subseteq \zeta\left(\mathcal{M}^D\right).$$

In other words, the map $\zeta$ is surjective. This proves Claim 3. □

*Claim 4:* Let $a_{\overline{S}} \in \mathbf{k}$ be a scalar for each $n$-tabloid $\overline{S}$ of shape $D$. Then,

$$\zeta\left(\sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid}\\ \text{of shape } D}} a_{\overline{S}}\overline{S}\right) = \sum_{u \in S_n} a_{u\overline{T}}u.$$

*Proof of Claim 4.* The map $\zeta$ is $\mathbf{k}$-linear. Thus,

$$\begin{aligned}
&\zeta\left(\sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} a_{\overline{S}}\overline{S}\right) \\
&= \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} a_{\overline{S}} \underbrace{\zeta\left(\overline{S}\right)}_{=\nabla_{\overline{S},\overline{T}}} \\
&\qquad \text{(by the definition of } \zeta\text{)} \\
&= \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} a_{\overline{S}} \underbrace{\nabla_{\overline{S},\overline{T}}}_{=\sum\limits_{\substack{u\in S_n; \\ u\overline{T}=\overline{S}}} u} = \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} a_{\overline{S}} \sum_{\substack{u\in S_n; \\ u\overline{T}=\overline{S}}} u \\
&\qquad \text{(by the definition of } \nabla_{\overline{S},\overline{T}}\text{)} \\
&= \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} \sum_{\substack{u\in S_n; \\ u\overline{T}=\overline{S}}} \underbrace{a_{\overline{S}}}_{=a_{u\overline{T}}} u = \underbrace{\sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} \sum_{\substack{u\in S_n; \\ u\overline{T}=\overline{S}}}}_{=\sum\limits_{u\in S_n}} a_{u\overline{T}}u \\
&\qquad \text{(since } \overline{S}=u\overline{T}\text{)} \\
&\qquad \text{(since each } u\in S_n \text{ appears in the inner sum for exactly one } \overline{S}\text{)} \\
&= \sum_{u\in S_n} a_{u\overline{T}}u.
\end{aligned}$$

This proves Claim 4. □

*Claim 5:* The map $\zeta$ is injective.

*Proof of Claim 5.* Let $\mathbf{s} \in \operatorname{Ker}\zeta$. We shall show that $\mathbf{s} = 0$.

We have $\mathbf{s} \in \operatorname{Ker}\zeta \subseteq \mathcal{M}^D$. Hence, we can write $\mathbf{s}$ as a $\mathbf{k}$-linear combination of $n$-tabloids of shape $D$ (since $\mathcal{M}^D$ is spanned by such $n$-tabloids). In other words, we can write $\mathbf{s}$ as

$$\mathbf{s} = \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} a_{\overline{S}}\overline{S} \tag{154}$$

for some coefficients $a_{\overline{S}} \in \mathbf{k}$. Consider these $a_{\overline{S}}$.

Applying $\zeta$ to the equality (154), we find

$$\begin{aligned}
\zeta\left(\mathbf{s}\right) &= \zeta\left(\sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} a_{\overline{S}}\overline{S}\right) = \sum_{u\in S_n} a_{u\overline{T}}u \qquad \text{(by Claim 4)} \\
&= \left(a_{u\overline{T}}\right)_{u\in S_n} \qquad \text{(by (12))}.
\end{aligned}$$

Thus, $\left(a_{u\overline{T}}\right)_{u\in S_n} = \zeta\left(\mathbf{s}\right) = 0$ (since $\mathbf{s} \in \operatorname{Ker}\zeta$). In other words,

$$a_{u\overline{T}} = 0 \qquad \text{for each } u \in S_n. \tag{155}$$

Now, let $\overline{S}$ be any $n$-tabloid of shape $D$. Then, Proposition 5.3.13 **(c)** shows that the $n$-tabloid $\overline{S}$ can be written as $w \rightharpoonup \overline{T}$ for some $w \in S_n$. Consider this $w$. Thus, $\overline{S} = w \rightharpoonup \overline{T} = w\overline{T}$, so that $a_{\overline{S}} = a_{w\overline{T}} = 0$ (by (155), applied to $u = w$).

Forget that we fixed $\overline{S}$. We thus have shown that $a_{\overline{S}} = 0$ for each $n$-tabloid $\overline{S}$ of shape $D$. Hence, we can rewrite (154) as $\mathbf{s} = \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} 0\overline{S} = 0$.

Forget that we fixed $\mathbf{s}$. We thus have shown that $\mathbf{s} = 0$ for each $\mathbf{s} \in \operatorname{Ker} \zeta$. In other words, $\operatorname{Ker} \zeta = 0$. Since $\zeta$ is a $\mathbf{k}$-linear map, this entails that $\zeta$ is injective. This proves Claim 5. □

The map $\zeta$ is injective (by Claim 5) and surjective (by Claim 3); thus, it is bijective, i.e., invertible. Since it is furthermore left $\mathbf{k}[S_n]$-linear (by Claim 2), we thus conclude that $\zeta$ is a left $\mathbf{k}[S_n]$-module isomorphism. This proves Proposition 5.5.21. □

We can now easily prove Theorem 5.5.13:

*Proof of Theorem 5.5.13.* **(a)** We must prove the following two claims:

*Claim 1:* There is a unique left $\mathbf{k}[S_n]$-linear map

$$\alpha : \mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T} \to \mathcal{M}^D$$

that sends $\nabla_{\operatorname{Row} T}$ to $\overline{T}$.

*Claim 2:* This map $\alpha$ is an isomorphism of $S_n$-representations.

*Proof of Claim 1.* The uniqueness of $\alpha$ is easy to see: If $\alpha : \mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T} \to \mathcal{M}^D$ is a left $\mathbf{k}[S_n]$-linear map that sends $\nabla_{\operatorname{Row} T}$ to $\overline{T}$, then each $\mathbf{a} \in \mathbf{k}[S_n]$ satisfies

$$\begin{aligned} \alpha\left(\mathbf{a}\nabla_{\operatorname{Row} T}\right) &= \mathbf{a} \underbrace{\alpha\left(\nabla_{\operatorname{Row} T}\right)}_{\substack{=\overline{T} \\ (\text{since } \alpha \text{ sends } \nabla_{\operatorname{Row} T} \text{ to } \overline{T})}} && (\text{since } \alpha \text{ is left } \mathbf{k}[S_n]\text{-linear}) \\ &= \mathbf{a}\overline{T}, \end{aligned}$$

and this determines all values of $\alpha$ (since each element of $\mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T}$ has the form $\mathbf{a}\nabla_{\operatorname{Row} T}$ for some $\mathbf{a} \in \mathbf{k}[S_n]$). Thus, the map $\alpha$ is unique. It remains to prove that such a map exists.

To do so, we construct it explicitly: Consider the left $\mathbf{k}[S_n]$-module isomorphism $\zeta : \mathcal{M}^D \to \mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T}$ constructed in Proposition 5.5.21. The inverse $\zeta^{-1}$ of $\zeta$ is then a left $\mathbf{k}[S_n]$-module isomorphism from $\mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T}$ to $\mathcal{M}^D$. Moreover, the definition of $\zeta$ yields $\zeta\left(\overline{T}\right) = \nabla_{\overline{T},\overline{T}} = \nabla_{\operatorname{Row} T}$ (by Proposition 5.5.17). Thus, $\zeta^{-1}\left(\nabla_{\operatorname{Row} T}\right) = \overline{T}$. In other words, $\zeta^{-1}$ sends $\nabla_{\operatorname{Row} T}$ to $\overline{T}$. Hence, there exists a left $\mathbf{k}[S_n]$-linear map $\alpha : \mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T} \to \mathcal{M}^D$ that sends $\nabla_{\operatorname{Row} T}$ to $\overline{T}$: namely, $\zeta^{-1}$ qualifies as such a map $\alpha$. This proves the existence of $\alpha$, and thus Claim 1 is fully proved. □

*Proof of Claim 2.* Consider the left $\mathbf{k}[S_n]$-module isomorphism $\zeta : \mathcal{M}^D \to \mathbf{k}[S_n] \cdot \nabla_{\mathrm{Row}\, T}$ constructed in Proposition 5.5.21. Then, its inverse $\zeta^{-1}$ is also a left $\mathbf{k}[S_n]$-module isomorphism, i.e., an isomorphism of $S_n$-representations.

However, this inverse $\zeta^{-1}$ is precisely the map $\alpha$ whose existence was claimed in Claim 1 (because we saw in the proof of Claim 1 that $\zeta^{-1}$ qualifies as such a map $\alpha$, but since $\alpha$ is unique, this means that $\zeta^{-1}$ **is** $\alpha$). Thus, the map $\alpha$ is an isomorphism of $S_n$-representations (since $\zeta^{-1}$ is an isomorphism of $S_n$-representations). This proves Claim 2. □

The proof of Theorem 5.5.13 **(a)** is now complete.

**(b)** We must prove that $\alpha$ sends $\mathbf{k}[S_n] \cdot \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}$ to $\mathcal{S}^D$.

The Specht module $\mathcal{S}^D$ is defined as the span of all polytabloids of shape $D$. Thus,

$$\mathcal{S}^D = \mathrm{span}_{\mathbf{k}} \{\mathbf{e}_S \mid S \text{ is an } n\text{-tableau of shape } D\}. \tag{156}$$

However, each $n$-tableau of shape $D$ can be written as $w \rightharpoonup T$ for some $w \in S_n$ (by Proposition 5.3.13 **(b)**). Thus,

$$\{n\text{-tableaux of shape } D\} = \{w \rightharpoonup T \mid w \in S_n\} = \{wT \mid w \in S_n\}$$

(since $wT$ is a shorthand for $w \rightharpoonup T$). Therefore,

$$\{\mathbf{e}_S \mid S \text{ is an } n\text{-tableau of shape } D\} = \{\mathbf{e}_{wT} \mid w \in S_n\}.$$

Thus, we can rewrite (156) as

$$\mathcal{S}^D = \mathrm{span}_{\mathbf{k}} \{\mathbf{e}_{wT} \mid w \in S_n\}. \tag{157}$$

On the other hand, $\mathbf{k}[S_n] = \mathrm{span}_{\mathbf{k}} \{w \mid w \in S_n\}$. Multiplying both sides of this equality with $\nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}$ on the right, we find

$$\begin{aligned} \mathbf{k}[S_n] \cdot \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T} &= \mathrm{span}_{\mathbf{k}} \{w \mid w \in S_n\} \cdot \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T} \\ &= \mathrm{span}_{\mathbf{k}} \left\{w \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T} \mid w \in S_n\right\} \end{aligned}$$

(since $\mathrm{span}_{\mathbf{k}} \{a_1, a_2, \ldots, a_k\} \cdot b = \mathrm{span}_{\mathbf{k}} \{a_1 b, a_2 b, \ldots, a_k b\}$ for any elements $a_1, a_2, \ldots, a_k, b$ of any $\mathbf{k}$-algebra). Applying the map $\alpha$ to this equality, we find

$$\begin{aligned} &\alpha\left(\mathbf{k}[S_n] \cdot \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}\right) \\ &= \alpha\left(\mathrm{span}_{\mathbf{k}} \left\{w \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T} \mid w \in S_n\right\}\right) \\ &= \mathrm{span}_{\mathbf{k}} \left\{\alpha\left(w \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}\right) \mid w \in S_n\right\} \end{aligned} \tag{158}$$

(since the map $\alpha$ is $\mathbf{k}$-linear and thus respects $\mathbf{k}$-linear combinations). But the map $\alpha$ is left $\mathbf{k}[S_n]$-linear. Hence,

$$\alpha\left(\nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}\right) = \nabla^-_{\mathrm{Col}\, T} \underbrace{\alpha\left(\nabla_{\mathrm{Row}\, T}\right)}_{= \overline{T}} = \nabla^-_{\mathrm{Col}\, T} \overline{T} = \mathbf{e}_T$$

(since $\alpha$ sends $\nabla_{\mathrm{Row}\, T}$ to $\overline{T}$)

(by Proposition 5.5.12). Moreover, let us recall again that the map $\alpha$ is left $\mathbf{k}[S_n]$-linear. Hence, each $w \in S_n$ satisfies

$$\alpha\left(w\nabla^{-}_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}\right) = w\underbrace{\alpha\left(\nabla^{-}_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}\right)}_{=\mathbf{e}_T} = w\mathbf{e}_T = \mathbf{e}_{wT}$$

(since Lemma 5.4.6 **(a)** yields $\mathbf{e}_{wT} = w\mathbf{e}_T$). Thus, we can rewrite (158) as

$$\alpha\left(\mathbf{k}[S_n]\cdot\nabla^{-}_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}\right) = \mathrm{span}_{\mathbf{k}}\left\{\mathbf{e}_{wT} \mid w \in S_n\right\}.$$

Comparing this with (157), we obtain

$$\alpha\left(\mathbf{k}[S_n]\cdot\nabla^{-}_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}\right) = \mathcal{S}^D.$$

In other words, $\alpha$ sends $\mathbf{k}[S_n]\cdot\nabla^{-}_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}$ to $\mathcal{S}^D$. Since $\alpha$ is an isomorphism of $S_n$-representations (by part **(a)**), this entails that $\mathcal{S}^D$ is isomorphic to $\mathbf{k}[S_n]\cdot\nabla^{-}_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}$ as $S_n$-representations. Thus, the proof of Theorem 5.5.13 **(b)** is complete. □

### 5.5.5. Explicit formulas for $\alpha^{-1}$

From our above proof of Theorem 5.5.13, we can recover some explicit formulas for the inverse $\alpha^{-1}$ of the isomorphism $\alpha$:

**Proposition 5.5.22.** Let $D$ be any diagram with $|D| = n$. Let $T$ be an $n$-tableau of shape $D$. Consider the isomorphism

$$\alpha : \mathbf{k}[S_n]\cdot\nabla_{\mathrm{Row}\,T} \to \mathcal{M}^D$$

from Theorem 5.5.13 **(a)**. Being an isomorphism, it has an inverse $\alpha^{-1} : \mathcal{M}^D \to \mathbf{k}[S_n]\cdot\nabla_{\mathrm{Row}\,T}$. Then:

**(a)** This $\alpha^{-1}$ is the isomorphism $\zeta$ from Proposition 5.5.21.

**(b)** We have

$$\alpha^{-1}\left(\overline{S}\right) = \nabla_{\overline{S},\overline{T}} \qquad \text{for each } n\text{-tabloid } \overline{S} \text{ of shape } D.$$

**(c)** Let $a_{\overline{S}} \in \mathbf{k}$ be a scalar for each $n$-tabloid $\overline{S}$ of shape $D$. Then,

$$\alpha^{-1}\left(\sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} a_{\overline{S}}\overline{S}\right) = \sum_{u\in S_n} a_{u\overline{T}}u.$$

*Proof.* Proposition 5.5.21 shows that the map $\zeta : \mathcal{M}^D \to \mathbf{k}[S_n] \cdot \nabla_{\mathrm{Row}\, T}$ is a left $\mathbf{k}[S_n]$-module isomorphism. Hence, its inverse $\zeta^{-1} : \mathbf{k}[S_n] \cdot \nabla_{\mathrm{Row}\, T} \to \mathcal{M}^D$ is left $\mathbf{k}[S_n]$-linear as well. Moreover, the definition of $\zeta$ yields $\zeta\left(\overline{T}\right) = \nabla_{\overline{T},\overline{T}} = \nabla_{\mathrm{Row}\, T}$ (by Proposition 5.5.17). Thus, $\zeta^{-1}\left(\nabla_{\mathrm{Row}\, T}\right) = \overline{T}$. In other words, $\zeta^{-1}$ sends $\nabla_{\mathrm{Row}\, T}$ to $\overline{T}$.

But $\alpha$ was defined as the unique left $\mathbf{k}[S_n]$-linear map

$$\alpha : \mathbf{k}[S_n] \cdot \nabla_{\mathrm{Row}\, T} \to \mathcal{M}^D$$

that sends $\nabla_{\mathrm{Row}\, T}$ to $\overline{T}$. Hence, if $\omega : \mathbf{k}[S_n] \cdot \nabla_{\mathrm{Row}\, T} \to \mathcal{M}^D$ is any left $\mathbf{k}[S_n]$-linear map that sends $\nabla_{\mathrm{Row}\, T}$ to $\overline{T}$, then $\omega = \alpha$ (by the uniqueness part of the preceding sentence). Applying this to $\omega = \zeta^{-1}$, we conclude that $\zeta^{-1} = \alpha$ (since $\zeta^{-1} : \mathbf{k}[S_n] \cdot \nabla_{\mathrm{Row}\, T} \to \mathcal{M}^D$ is a left $\mathbf{k}[S_n]$-linear map that sends $\nabla_{\mathrm{Row}\, T}$ to $\overline{T}$). In other words, $\alpha$ is $\zeta^{-1}$. This proves Proposition 5.5.22 **(a)**.

**(b)** Let $\overline{S}$ be an $n$-tabloid of shape $D$. The definition of $\zeta$ shows that $\zeta\left(\overline{S}\right) = \nabla_{\overline{S},\overline{T}}$. Since $\zeta = \alpha^{-1}$ (because $\zeta^{-1} = \alpha$), we can rewrite this as $\alpha^{-1}\left(\overline{S}\right) = \nabla_{\overline{S},\overline{T}}$. This proves Proposition 5.5.22 **(b)**.

**(c)** Claim 4 in the above proof of Proposition 5.5.21 shows that

$$\zeta\left(\sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} a_{\overline{S}}\overline{S}\right) = \sum_{u \in S_n} a_{u\overline{T}}u.$$

Since $\zeta = \alpha^{-1}$ (because $\zeta^{-1} = \alpha$), we can rewrite this as

$$\alpha^{-1}\left(\sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} a_{\overline{S}}\overline{S}\right) = \sum_{u \in S_n} a_{u\overline{T}}u.$$

This proves Proposition 5.5.22 **(c)**. □

### 5.5.6. A restatement of the left ideal avatar of the Specht module

Let us restate part of Theorem 5.5.13 in a form that involves only the Specht module $\mathcal{S}^D$ (not the Young module $\mathcal{M}^D$); this will be convenient to us later.

**Corollary 5.5.23.** Let $D$ be any diagram with $|D| = n$. Let $T$ be an $n$-tableau of shape $D$. Then, there is a left $\mathbf{k}[S_n]$-module isomorphism

$$\overline{\alpha} : \mathbf{k}[S_n] \cdot \nabla^{-}_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T} \to \mathcal{S}^D$$

that sends $\nabla^{-}_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}$ to $\mathbf{e}_T$.

*Proof.* Theorem 5.5.13 **(a)** shows that there is a unique left $\mathbf{k}[S_n]$-linear map

$$\alpha : \mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T} \to \mathcal{M}^D$$

that sends $\nabla_{\operatorname{Row} T}$ to $\overline{T}$, and that this map is an isomorphism of $S_n$-representations (i.e., of left $\mathbf{k}[S_n]$-modules). Consider this isomorphism $\alpha$.

Theorem 5.5.13 **(b)** shows that this isomorphism $\alpha$ sends the submodule $\mathbf{k}[S_n] \cdot \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$ of $\mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T}$ to the Specht module $\mathcal{S}^D$. Thus, by restricting $\alpha$ to the former submodule $\mathbf{k}[S_n] \cdot \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$, we obtain an isomorphism

$$\begin{aligned} \overline{\alpha} : \mathbf{k}[S_n] \cdot \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} &\to \mathcal{S}^D, \\ \mathbf{x} &\mapsto \alpha(\mathbf{x}). \end{aligned}$$

Consider this isomorphism $\overline{\alpha}$.

The map $\alpha : \mathbf{k}[S_n] \cdot \nabla_{\operatorname{Row} T} \to \mathcal{M}^D$ is a morphism of $S_n$-representations, and thus is a morphism of left $\mathbf{k}[S_n]$-modules (since Proposition 4.3.18 shows that any morphism of $S_n$-representations is a morphism of left $\mathbf{k}[S_n]$-modules). Hence, its restriction $\overline{\alpha} : \mathbf{k}[S_n] \cdot \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \to \mathcal{S}^D$ is a morphism of left $\mathbf{k}[S_n]$-modules as well. Since $\overline{\alpha}$ is furthermore invertible (because it is an isomorphism of $S_n$-representations), we thus conclude that $\overline{\alpha}$ is an isomorphism of left $\mathbf{k}[S_n]$-modules.

It remains to show that $\overline{\alpha}$ sends $\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$ to $\mathbf{e}_T$. For this purpose, we observe that $\alpha(\nabla_{\operatorname{Row} T}) = \overline{T}$ (since $\alpha$ sends $\nabla_{\operatorname{Row} T}$ to $\overline{T}$). But $\overline{\alpha}$ was defined as a restriction of $\alpha$; therefore,

$$\begin{aligned} &\overline{\alpha}\left(\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}\right) \\ &= \alpha\left(\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}\right) \\ &= \nabla^-_{\operatorname{Col} T} \underbrace{\alpha\left(\nabla_{\operatorname{Row} T}\right)}_{=\overline{T}} && \text{(since } \alpha \text{ is a morphism of left } \mathbf{k}[S_n]\text{-modules)} \\ &= \nabla^-_{\operatorname{Col} T} \overline{T} = \mathbf{e}_T && \text{(by Proposition 5.5.12)}. \end{aligned}$$

In other words, $\overline{\alpha}$ sends $\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$ to $\mathbf{e}_T$. This completes the proof of Corollary 5.5.23. □

## 5.6. Determinantal avatars of Young and Specht modules

There are several other avatars (i.e., isomorphic versions) of the Specht modules $\mathcal{S}^D$. Two such avatars involve matrix algebra. They are constructed by taking certain matrices (which consist of indeterminates or polynomials) and assigning a certain product of determinants (specifically, of minors – i.e., determinants of submatrices) to each $n$-tableau of shape $D$.

Both avatars cover not only the Specht module $\mathcal{S}^D$ but also the Young module $\mathcal{M}^D$, and indeed are easiest to define starting with $\mathcal{M}^D$. We shall do so now.

First, however, we recall a notion of matrices that is slightly more general than the one commonly studied in linear algebra. Namely, if $P$ and $Q$ are two finite sets, then a $P \times Q$-*matrix* means a family $\left(a_{p,q}\right)_{(p,q)\in P\times Q}$ of elements $a_{p,q}$ of a ring indexed by all pairs $(p,q) \in P \times Q$. When $P = [k]$ and $Q = [\ell]$ for some nonnegative integers $k$ and $\ell$, this notion coincides with the classical notion of a $k \times \ell$-matrix. A $P \times Q$-matrix is said to be *square* if $P = Q$. For a square matrix $\left(a_{p,q}\right)_{(p,q)\in P\times P}$, we can also use the shorthand notation $\left(a_{p,q}\right)_{p,q\in P}$. The determinant $\det A$ of a square matrix $\left(a_{p,q}\right)_{(p,q)\in P\times P}$ is defined to be the sum

$$\sum_{\sigma\in S_P} (-1)^\sigma \prod_{p\in P} a_{p,\sigma(p)};$$

this is the same sum that would be obtained if we relabeled the elements of $P$ as $1, 2, \ldots, k$ (for $k = |P|$) and then took the (usual) determinant of the resulting $k \times k$-matrix.

### 5.6.1. The Specht–Vandermonde avatar

Now we come to the two matrix avatars. We begin with the more classical one, which we call the *Specht–Vandermonde avatar*, since it involves something similar to Vandermonde determinants and was introduced by Specht in [Specht35]:

**Theorem 5.6.1** (Specht–Vandermonde avatar of Young and Specht modules)**.** Let $D$ be a diagram with $|D| = n$. Assume that all $(i,j) \in D$ satisfy $i > 0$.

Consider the polynomial ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$. The symmetric group $S_n$ acts on this ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$ by permuting the variables (that is, $g \rightharpoonup x_i = x_{g(i)}$ for all $g \in S_n$ and $i \in [n]$).

We define a $\mathbf{k}$-linear map

$$\begin{aligned}\beta : \mathcal{M}^D &\to \mathbf{k}[x_1, x_2, \ldots, x_n], \\ \overline{T} &\mapsto \prod_{(i,j)\in D} x_{T(i,j)}^{i-1} = \prod_{i\geq 1} \prod_{m\in \mathrm{Row}(i,T)} x_m^{i-1},\end{aligned}$$

where $\mathrm{Row}(i,T)$ denotes the set of all entries in the $i$-th row of $T$. Then:

**(a)** This map $\beta$ is an injective left $\mathbf{k}[S_n]$-module morphism.

**(b)** This map $\beta$ sends any polytabloid $\mathbf{e}_T \in \mathcal{M}^D$ to the product

$$\prod_{j\in\mathbb{Z}} \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle}, \tag{159}$$

where $D\langle j\rangle$ is the set of all integers $i$ that satisfy $(i,j) \in D$. (Note that the matrix $\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle}$ is thus a square matrix, whose rows and columns are indexed by the elements of $D\langle j\rangle$. Such square matrices have well-defined

determinants, since $D\langle j\rangle$ is a finite set. When $D\langle j\rangle$ is empty, this matrix is a $0\times 0$-matrix and thus has determinant 1, by definition. Hence, all but finitely many factors of the product in (159) equal 1.)

Thus, the map $\beta$ sends the Specht module $\mathcal{S}^D$ to the span of the products shown in (159).

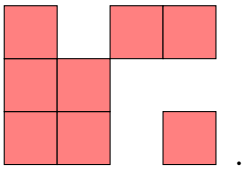
**(c)** Assume that $D$ is the Young diagram $Y(\lambda)$ of a partition $\lambda$. Let $\lambda^t$ be the conjugate of the partition $\lambda$ (as defined in Theorem 5.1.10). Then, the map $\beta$ sends any polytabloid $\mathbf{e}_T$ to

$$\prod_{j\geq 1}\prod_{1\leq i_1<i_2\leq\lambda_j^t}\left(x_{T(i_2,j)}-x_{T(i_1,j)}\right).$$

**Example 5.6.2.** Let $n=8$ and

$$D=\{(1,1),\ (1,3),\ (1,4),\ (2,1),\ (2,2),\ (3,1),\ (3,2),\ (3,4)\}.$$

Visually, $D$ looks as follows:



Define an $n$-tableau $T$ of shape $D$ by

$$T=\begin{array}{|c|c|c|c|}\hline 4 & & 1 & 3\\\hline 2 & 8 & & \multicolumn{1}{c}{}\\\cline{1-2} 7 & 6 & \multicolumn{1}{c}{} & 5\\\cline{1-2}\end{array}\ .$$

Then, the map $\beta$ from Theorem 5.6.1 sends the $n$-tabloid $\overline{T}\in\mathcal{M}^D$ to the monomial

$$\begin{aligned}\prod_{(i,j)\in D}x_{T(i,j)}^{i-1} &= x_{T(1,1)}^0x_{T(1,3)}^0x_{T(1,4)}^0x_{T(2,1)}^1x_{T(2,2)}^1x_{T(3,1)}^2x_{T(3,2)}^2x_{T(3,4)}^2\\ &= x_4^0x_1^0x_3^0x_2^1x_8^1x_7^2x_6^2x_5^2=x_2x_8x_7^2x_6^2x_5^2,\end{aligned}$$

and (by Theorem 5.6.1 **(b)**) sends the polytabloid $\mathbf{e}_T$ to the product

$$\begin{aligned}
&\prod_{j\in\mathbb{Z}} \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle} \\
&= \det\left(x_{T(i,1)}^{p-1}\right)_{i,p\in D\langle 1\rangle} \cdot \det\left(x_{T(i,2)}^{p-1}\right)_{i,p\in D\langle 2\rangle} \\
&\qquad \cdot \det\left(x_{T(i,3)}^{p-1}\right)_{i,p\in D\langle 3\rangle} \cdot \det\left(x_{T(i,4)}^{p-1}\right)_{i,p\in D\langle 4\rangle} \\
&\qquad\qquad \left(\begin{array}{c}\text{because for all } j\notin\{1,2,3,4\}\text{, the set } D\langle j\rangle \text{ is empty,}\\ \text{and thus } \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle} = 1\end{array}\right) \\
&= \det\begin{pmatrix} x_4^0 & x_4^1 & x_4^2 \\ x_2^0 & x_2^1 & x_2^2 \\ x_7^0 & x_7^1 & x_7^2 \end{pmatrix} \cdot \det\begin{pmatrix} x_8^1 & x_8^2 \\ x_6^2 & x_6^2 \end{pmatrix} \cdot \det\begin{pmatrix} x_1^0 \end{pmatrix} \cdot \det\begin{pmatrix} x_3^0 & x_3^2 \\ x_5^0 & x_5^2 \end{pmatrix} \\
&\qquad\qquad \left(\begin{array}{c}\text{since } D\langle 1\rangle = \{1,2,3\} \text{ and } D\langle 2\rangle = \{2,3\}\\ \text{and } D\langle 3\rangle = \{1\} \text{ and } D\langle 4\rangle = \{1,3\}\end{array}\right).
\end{aligned}$$

**Example 5.6.3.** Let $n = 6$ and $D = Y(3,2,1)$, and let $T$ be the tableau $\begin{array}{|c|c|c|}\hline 1&2&6\\\hline 3&4\\\cline{1-2} 5\\\cline{1-1}\end{array}$ of shape $D$. Then, Theorem 5.6.1 **(c)** says that

$$\beta(\mathbf{e}_T) = \underbrace{(x_3-x_1)(x_5-x_1)(x_5-x_3)}_{\text{from column 1}} \cdot \underbrace{(x_4-x_2)}_{\text{from column 2}}.$$

Note that the third column of $T$ contributes nothing (or, to be more precise, an empty product) to this product, since a product of the form $\prod_{1\le i_1<i_2\le\lambda_j^t}$ is empty when $\lambda_j^t \le 1$. More generally, any column of $Y(\lambda)$ that contains only one cell (or no cells at all) contributes an empty product to the product in Theorem 5.6.1 **(c)**.

*Proof of Theorem 5.6.1.* **(a)** First, we need to show that the map $\beta$ is well-defined. This requires proving the following two claims:

*Claim 1:* If $T$ is an $n$-tableau of shape $D$, then the product $\prod_{i\ge1}\prod_{m\in\mathrm{Row}(i,T)} x_m^{i-1}$ depends only on the $n$-tabloid $\overline{T}$, not on the $n$-tableau $T$.

*Claim 2:* If $T$ is an $n$-tableau of shape $D$, then $\prod_{(i,j)\in D} x_{T(i,j)}^{i-1} = \prod_{i\ge1}\prod_{m\in\mathrm{Row}(i,T)} x_m^{i-1}$.

*Proof of Claim 1.* Let $T$ be an $n$-tableau of shape $D$. Then, for each $i \in \mathbb{Z}$, the set $\operatorname{Row}(i,T)$ is the set of all entries in the $i$-th row of $T$, and thus depends only on the $n$-tabloid $\overline{T}$, not on the $n$-tableau $T$ (because row-equivalent tableaux have the same set of entries in their $i$-th row). Therefore, the product $\prod_{i\geq 1} \prod_{m\in\operatorname{Row}(i,T)} x_m^{i-1}$ also depends only on the $n$-tabloid $\overline{T}$, not on the $n$-tableau $T$. This proves Claim 1. □

*Proof of Claim 2.* Let $T$ be an $n$-tableau of shape $D$. Then, $T$ is an injective tableau, i.e., all the entries of $T$ are distinct. Moreover, all $(i,j) \in D$ satisfy $i > 0$ and thus $i \geq 1$ (since $i$ is an integer). Hence, we can rewrite the product sign $\prod_{(i,j)\in D}$ as $\prod_{i\geq 1} \prod_{\substack{j\in\mathbb{Z};\\ (i,j)\in D}}$. Hence,

$$\prod_{(i,j)\in D} x_{T(i,j)}^{i-1} = \prod_{i\geq 1} \prod_{\substack{j\in\mathbb{Z};\\ (i,j)\in D}} x_{T(i,j)}^{i-1}. \tag{160}$$

Now, let $i$ be any integer. We shall show that

$$\prod_{\substack{j\in\mathbb{Z};\\ (i,j)\in D}} x_{T(i,j)}^{i-1} = \prod_{m\in\operatorname{Row}(i,T)} x_m^{i-1}. \tag{161}$$

Indeed, $\operatorname{Row}(i,T)$ is the set of all entries in the $i$-th row of $T$ (by the definition of $\operatorname{Row}(i,T)$), and these entries are all distinct (since all the entries of $T$ are distinct). Thus, the right hand side of (161) is the product of the $x_m^{i-1}$, where $m$ ranges over all entries in the $i$-th row of $T$. But the left hand side is the same product as well (since the entries in the $i$-th row of $T$ are precisely the $T(i,j)$ for all integers $j$ satisfying $(i,j) \in D$). Thus, the two sides of (161) are equal. Hence, (161) is proved.

Forget that we fixed $i$. We thus have proved (161) for each integer $i$. Using (161), we can rewrite (160) as

$$\prod_{(i,j)\in D} x_{T(i,j)}^{i-1} = \prod_{i\geq 1} \prod_{m\in\operatorname{Row}(i,T)} x_m^{i-1}.$$

Thus, Claim 2 is proved. □

Claim 1 and Claim 2 yield that the map $\beta$ is well-defined. It remains to prove the following two claims:

*Claim 3:* The map $\beta$ is injective.

*Claim 4:* The map $\beta$ is left $\mathbf{k}[S_n]$-linear.

*Proof of Claim 3.* Let $\mathfrak{M}$ denote the set of all monomials in the $n$ indeterminates $x_1, x_2, \ldots, x_n$. Then, the $\mathbf{k}$-module $\mathbf{k}[x_1, x_2, \ldots, x_n]$ is the free $\mathbf{k}$-module $\mathbf{k}^{(\mathfrak{M})}$. Meanwhile, $\mathcal{M}^D$ is the free $\mathbf{k}$-module $\mathbf{k}^{(\{n\text{-tabloids of shape } D\})}$ (by its definition). The $\mathbf{k}$-linear map $\beta : \mathcal{M}^D \to \mathbf{k}[x_1, x_2, \ldots, x_n]$ sends each standard basis vector $\overline{T}$ of $\mathcal{M}^D$ to the standard basis vector $\prod_{(i,j)\in D} x_{T(i,j)}^{i-1}$ of $\mathbf{k}[x_1, x_2, \ldots, x_n]$; thus, it is the linearization[55] $f_*$ of the map

$$f : \{n\text{-tabloids of shape } D\} \to \mathfrak{M}, \qquad \overline{T} \mapsto \prod_{(i,j)\in D} x_{T(i,j)}^{i-1}.$$

Consider this latter map $f$. If we can show that $f$ is injective, then it will follow (by Proposition 3.12.8) that its linearization $f_*$ is injective as well, and this will mean that $\beta$ is injective (since $\beta$ is the linearization $f_*$); thus, Claim 3 will be proven. Hence, it suffices to show that $f$ is injective.

So let us prove this. Let $\overline{T}$ and $\overline{S}$ be two $n$-tabloids of shape $D$ such that $f\left(\overline{T}\right) = f\left(\overline{S}\right)$. We must show that $\overline{T} = \overline{S}$.

We know that $T$ is an $n$-tableau, i.e., an injective map from $D$ to $\{1, 2, 3, \ldots\}$ whose image is $[n]$. Thus, we can view $T$ as a bijection from $D$ to $[n]$. Similarly, we can view $S$ as a bijection from $D$ to $[n]$. This is how we shall view $T$ and $S$ henceforth.

For each $m \in [n]$, let us denote the cell $T^{-1}(m) \in D$ (that is, the cell of $T$ that contains $m$) by $(a_m, b_m)$, and let us denote the cell $S^{-1}(m) \in D$ (that is, the cell of $S$ that contains $m$) by $(c_m, d_m)$.

Now, the definition of $f$ easily yields that

$$f\left(\overline{T}\right) = \prod_{m=1}^{n} x_m^{a_m - 1} \tag{162}$$

[55] See Definition 3.12.7 for the definition of "linearization".

[56]. Similarly,

$$f\left(\overline{S}\right) = \prod_{m=1}^{n} x_m^{c_m - 1}.$$

The left hand sides of these two equalities are equal (since $f\left(\overline{T}\right) = f\left(\overline{S}\right)$). Thus, so are their right hand sides. In other words, $\prod_{m=1}^{n} x_m^{a_m - 1} = \prod_{m=1}^{n} x_m^{c_m - 1}$. Therefore,

$$a_m - 1 = c_m - 1 \qquad \text{for each } m \in [n]$$

(since equal monomials must have equal powers on each indeterminate). In other words,

$$a_m = c_m \qquad \text{for each } m \in [n]\,. \tag{164}$$

Now, fix any $i \in \mathbb{Z}$. Note that the entries in the $i$-th row of $T$ are distinct (since $T$ is injective). Likewise, so are the entries in the $i$-th row of $S$.

For any $m \in [n]$, we have the logical equivalences

$$\begin{aligned}
&(m \in \operatorname{Row}(i,T)) \\
&\iff (m \text{ appears in the } i\text{-th row of } T) \\
&\qquad \left( \begin{array}{c} \text{since } \operatorname{Row}(i,T) \text{ is defined to be the} \\ \text{set of all entries in the } i\text{-th row of } T \end{array} \right) \\
&\iff (\text{the cell of } T \text{ that contains } m \text{ lies in the } i\text{-th row}) \\
&\iff (\text{the cell } (a_m, b_m) \text{ lies in the } i\text{-th row}) \\
&\qquad \left(\text{since the cell of } T \text{ that contains } m \text{ is } T^{-1}(m) = (a_m, b_m)\right) \\
&\iff (a_m = i)
\end{aligned} \tag{165}$$

and similarly the equivalence

$$(m \in \operatorname{Row}(i,S)) \iff (c_m = i)\,. \tag{166}$$

[56] *Proof.* Let $(i,j) \in D$. For each $m \in [n]$, we have $(a_m, b_m) = T^{-1}(m)$ (by the definition of $(a_m, b_m)$). Applying this to $m = T(i,j)$, we obtain $\left(a_{T(i,j)}, b_{T(i,j)}\right) = T^{-1}(T(i,j)) = (i,j)$. In other words, $a_{T(i,j)} = i$ and $b_{T(i,j)} = j$.

Forget that we fixed $(i,j)$. We thus have shown that all $(i,j) \in D$ satisfy

$$a_{T(i,j)} = i \qquad \text{and} \qquad b_{T(i,j)} = j. \tag{163}$$

But the definition of $f$ yields

$$f\left(\overline{T}\right) = \prod_{(i,j)\in D} \underbrace{x_{T(i,j)}^{i-1}}_{\substack{= x_{T(i,j)}^{a_{T(i,j)}-1} \\ \text{(since (163) yields } a_{T(i,j)} = i \\ \text{and thus } i = a_{T(i,j)}\text{)}}} = \prod_{(i,j)\in D} x_{T(i,j)}^{a_{T(i,j)}-1} = \prod_{m\in[n]} x_m^{a_m - 1}$$

(here, we have substituted $m$ for $T(i,j)$ in the product, since the map $T : D \to [n]$ is a bijection). Qed.

Thus, for any $m \in [n]$, we have the logical equivalences

$$\begin{aligned}
(m \in \operatorname{Row}(i,T)) &\iff (a_m = i) && \text{(by (165))} \\
&\iff (c_m = i) && \text{(since (164) yields } a_m = c_m\text{)} \\
&\iff (m \in \operatorname{Row}(i,S)) && \text{(by (166))}\,.
\end{aligned}$$

Hence, $\operatorname{Row}(i,T) = \operatorname{Row}(i,S)$ (since $\operatorname{Row}(i,T)$ and $\operatorname{Row}(i,S)$ are subsets of $[n]$). In other words, the set of entries in the $i$-th row of $T$ equals the set of entries in the $i$-th row of $S$ (since these two sets are precisely $\operatorname{Row}(i,T)$ and $\operatorname{Row}(i,S)$). We can furthermore replace the word "set" by "multiset" here (since the entries in the $i$-th row of $T$ are distinct, and so are the entries in the $i$-th row of $S$, and therefore the multisets of these entries do not have any nontrivial multiplicities). Thus, we conclude that the multiset of entries in the $i$-th row of $T$ equals the multiset of entries in the $i$-th row of $S$.

Forget that we fixed $i$. We thus have shown that for each $i \in \mathbb{Z}$, the multiset of entries in the $i$-th row of $T$ equals the multiset of entries in the $i$-th row of $S$. In other words, the $n$-tableaux $T$ and $S$ are row-equivalent. In other words, $\overline{T} = \overline{S}$ (since an $n$-tabloid is an equivalence class with respect to row-equivalence).

Forget that we fixed $\overline{T}$ and $\overline{S}$. We thus have shown that if $\overline{T}$ and $\overline{S}$ are two $n$-tabloids of shape $D$ such that $f\left(\overline{T}\right) = f\left(\overline{S}\right)$, then $\overline{T} = \overline{S}$. In other words, $f$ is injective. As explained above, this completes the proof of Claim 3. □

*Proof of Claim 4.* We shall first show that $\beta$ is $S_n$-equivariant.

For this purpose, we must prove that $\beta(w\mathbf{a}) = w\beta(\mathbf{a})$ for all $w \in S_n$ and all $\mathbf{a} \in \mathcal{M}^D$.

Fix $w \in S_n$ and $\mathbf{a} \in \mathcal{M}^D$. We must prove that $\beta(w\mathbf{a}) = w\beta(\mathbf{a})$. Both sides of this equality are $\mathbf{k}$-linear in $\mathbf{a}$. Thus, by linearity, we can WLOG assume that $\mathbf{a}$ is an $n$-tabloid of shape $D$ (since $\mathcal{M}^D$ is spanned by such $n$-tabloids). Assume this. Thus, $\mathbf{a} = \overline{T}$ for some $n$-tabloid $\overline{T}$ of shape $D$. Consider this $\overline{T}$. From $\mathbf{a} = \overline{T}$, we obtain $\beta(\mathbf{a}) = \beta\left(\overline{T}\right) = \prod\limits_{(i,j)\in D} x_{T(i,j)}^{i-1}$ (by the definition of $\beta$), so that

$$w\beta(\mathbf{a}) = w \prod_{(i,j)\in D} x_{T(i,j)}^{i-1} = \prod_{(i,j)\in D} x_{w(T(i,j))}^{i-1} \tag{167}$$

(since the action of $w$ transforms each indeterminate $x_k$ into $x_{w(k)}$). On the other hand, from $\mathbf{a} = \overline{T}$, we obtain $w\mathbf{a} = w\overline{T} = \overline{wT}$, so that

$$\begin{aligned}
\beta(w\mathbf{a}) = \beta\left(\overline{wT}\right) &= \prod_{(i,j)\in D} \underbrace{x_{(wT)(i,j)}^{i-1}}_{\substack{= x_{w(T(i,j))}^{i-1} \\ \text{(since } wT = w\circ T \text{ and} \\ \text{thus } (wT)(i,j) = w(T(i,j))\text{)}}} && \text{(by the definition of } \beta\text{)} \\
&= \prod_{(i,j)\in D} x_{w(T(i,j))}^{i-1}.
\end{aligned}$$

Comparing this with (167), we find $\beta(w\mathbf{a}) = w\beta(\mathbf{a})$.

Forget that we fixed $w$ and $\mathbf{a}$. We thus have shown that $\beta(w\mathbf{a}) = w\beta(\mathbf{a})$ for all $w \in S_n$ and $\mathbf{a} \in \mathcal{M}^D$. In other words, the map $\beta$ is $S_n$-equivariant. Since $\beta$ is also $\mathbf{k}$-linear, we thus conclude that $\beta$ is a morphism of $S_n$-representations. Therefore (by Proposition 4.3.18), the map $\beta$ is a morphism of left $\mathbf{k}[S_n]$-modules, i.e., a left $\mathbf{k}[S_n]$-linear map. This proves Claim 4. □

Claim 3 and Claim 4 show that $\beta$ is an injective left $\mathbf{k}[S_n]$-module morphism. Theorem 5.6.1 **(a)** is thus proved.

**(b)** Let $T$ be an $n$-tableau of shape $D$. Then, the definition of $\mathbf{e}_T$ yields

$$\mathbf{e}_T = \sum_{w \in \mathcal{C}(T)} (-1)^w \overline{wT}.$$

Hence,

$$\begin{aligned}
\beta(\mathbf{e}_T) &= \beta\left(\sum_{w \in \mathcal{C}(T)} (-1)^w \overline{wT}\right) \\
&= \sum_{w \in \mathcal{C}(T)} (-1)^w \beta\left(\overline{wT}\right) \qquad \text{(since the map } \beta \text{ is } \mathbf{k}\text{-linear)} \\
&= \sum_{w \in \mathcal{C}(T)} (-1)^w \prod_{(i,j) \in D} x_{(wT)(i,j)}^{i-1} \qquad (168)
\end{aligned}$$

(since the definition of $\beta$ yields $\beta\left(\overline{wT}\right) = \prod_{(i,j) \in D} x_{(wT)(i,j)}^{i-1}$ for every $w \in \mathcal{C}(T)$).

Now, let us WLOG assume that all cells $(i,j) \in D$ satisfy $j > 0$. (We can always achieve this by a horizontal translation of $D$ (since $D$ is finite); this does not affect the validity of the theorem we are proving.) Since $D$ is finite, there exists some $k \in \mathbb{N}$ such that all cells of $D$ lie in columns $1, 2, \ldots, k$. Pick such a $k$. Thus, each cell $(i,j) \in D$ satisfies $j \in [k]$. Therefore, we can rewrite the product sign $\prod_{(i,j) \in D}$ as $\prod_{j=1}^{k} \prod_{\substack{i \in \mathbb{Z}; \\ (i,j) \in D}}$. Furthermore, for any given $j \in [k]$, we can rewrite the product sign $\prod_{\substack{i \in \mathbb{Z}; \\ (i,j) \in D}}$ as $\prod_{i \in D\langle j \rangle}$ (since $D\langle j \rangle$ is the set of all integers $i$ such that $(i,j) \in D$). Altogether, we thus find the following equality of product signs:

$$\prod_{(i,j) \in D} = \prod_{j=1}^{k} \underbrace{\prod_{\substack{i \in \mathbb{Z}; \\ (i,j) \in D}}}_{= \prod_{i \in D\langle j \rangle}} = \prod_{j=1}^{k} \prod_{i \in D\langle j \rangle}. \qquad (169)$$

Furthermore, for each integer $j \notin [k]$, we have $D\langle j\rangle = \varnothing$ (since all cells of $D$ lie in columns $1, 2, \ldots, k$, and thus no cell of $D$ lies in column $j$), and therefore

$$\det \underbrace{\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle}}_{\substack{\text{a } 0\times 0\text{-matrix} \\ (\text{since } D\langle j\rangle = \varnothing)}} = 1.$$

In other words, in the infinite product $\prod_{j\in\mathbb{Z}} \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle}$, all factors with $j \notin [k]$ equal 1. Hence, this infinite product is well-defined, and equals the same product but restricted only to the factors with $j \in [k]$. In other words,

$$\prod_{j\in\mathbb{Z}} \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle} = \prod_{j=1}^{k} \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle}. \tag{170}$$

For any $j \in \mathbb{Z}$, let $\operatorname{Col}(j,T)$ denote the set of all entries in the $j$-th column of $T$. Let us set $X_j := \operatorname{Col}(j,T)$ for each $j \in [k]$. As in the proof of Proposition 5.5.3 **(b)**, we can see that $\operatorname{Col}(1,T)$, $\operatorname{Col}(2,T)$, $\ldots$, $\operatorname{Col}(k,T)$ are $k$ disjoint subsets of $[n]$ and satisfy $\operatorname{Col}(1,T) \cup \operatorname{Col}(2,T) \cup \cdots \cup \operatorname{Col}(k,T) = [n]$. In other words, $X_1, X_2, \ldots, X_k$ are $k$ disjoint subsets of $[n]$ and satisfy $X_1 \cup X_2 \cup \cdots \cup X_k = [n]$ (since $X_j = \operatorname{Col}(j,T)$ for each $j \in [k]$).

Define a subset $K$ of $S_n$ by

$$K := \{w \in S_n \mid w(X_i) \subseteq X_i \text{ for all } i \in [k]\}.$$

Then,

$$\begin{aligned} K &= \{w \in S_n \mid w(X_i) \subseteq X_i \text{ for all } i \in [k]\} \\ &= \{w \in S_n \mid w(\operatorname{Col}(i,T)) \subseteq \operatorname{Col}(i,T) \text{ for all } i \in [k]\} \end{aligned} \tag{171}$$

(since $X_i = \operatorname{Col}(i,T)$ for each $i \in [k]$).

However, for any permutation $w \in S_n$, we have the logical equivalence

$$(w \in \mathcal{C}(T)) \iff (w(\operatorname{Col}(i,T)) \subseteq \operatorname{Col}(i,T) \text{ for all } i \in [k]) \tag{172}$$

(this was already proved back in the proof of Proposition 5.5.3 **(b)**). From $\mathcal{C}(T) \subseteq S_n$, we obtain

$$\begin{aligned} \mathcal{C}(T) &= \{w \in S_n \mid w \in \mathcal{C}(T)\} \\ &= \{w \in S_n \mid w(\operatorname{Col}(i,T)) \subseteq \operatorname{Col}(i,T) \text{ for all } i \in [k]\} \\ &\qquad \text{(by the equivalence (172))} \\ &= K \qquad \text{(by (171))}. \end{aligned}$$

Thus, we can rewrite (168) as

$$\beta(\mathbf{e}_T) = \sum_{w\in K} (-1)^w \prod_{(i,j)\in D} x_{(wT)(i,j)}^{i-1}. \tag{173}$$

For each permutation $w \in K$ and each $j \in [k]$, we define a map $w^{\langle j \rangle} : X_j \to X_j$ by setting

$$w^{\langle j \rangle}(x) = w(x) \qquad \text{for each } x \in X_j.$$

As we know from Lemma 3.12.20 **(a)**, this map $w^{\langle j \rangle}$ is well-defined and belongs to the symmetric group $S_{X_j}$ for each $w \in K$ and each $j \in [k]$. Furthermore, Lemma 3.12.20 **(b)** tells us that the map

$$\begin{aligned} & K \to S_{X_1} \times S_{X_2} \times \cdots \times S_{X_k}, \\ & w \mapsto \left( w^{\langle 1 \rangle}, w^{\langle 2 \rangle}, \ldots, w^{\langle k \rangle} \right) \end{aligned} \tag{174}$$

is a bijection. It is easy to see that each $w \in K$ and each $(i,j) \in D$ satisfy

$$(wT)(i,j) = w^{\langle j \rangle}(T(i,j)) \tag{175}$$

[57].

Finally, each $w \in K$ satisfies

$$\begin{aligned} (-1)^w &= (-1)^{w^{\langle 1 \rangle}} (-1)^{w^{\langle 2 \rangle}} \cdots (-1)^{w^{\langle k \rangle}} \qquad \text{(by Lemma 3.12.20 \textbf{(c)})} \\ &= \prod_{j=1}^{k} (-1)^{w^{\langle j \rangle}}. \end{aligned} \tag{176}$$

[57] *Proof:* Let $w \in K$ and $(i,j) \in D$. Then, $T(i,j)$ is an entry in the $j$-th column of $T$, and thus belongs to $\operatorname{Col}(j,T)$ (by the definition of $\operatorname{Col}(j,T)$). In other words, $T(i,j) \in \operatorname{Col}(j,T) = X_j$ (since $X_j$ is defined to be $\operatorname{Col}(j,T)$). Hence, the definition of $w^{\langle j \rangle}$ yields $w^{\langle j \rangle}(T(i,j)) = w(T(i,j))$. But the definition of the action of $S_n$ on $n$-tableaux shows that $wT = w \circ T$, so that $(wT)(i,j) = (w \circ T)(i,j) = w(T(i,j))$. Comparing this with $w^{\langle j \rangle}(T(i,j)) = w(T(i,j))$, we obtain $(wT)(i,j) = w^{\langle j \rangle}(T(i,j))$. This proves (175).

Hence, (173) becomes

$$\begin{aligned}
\beta\left(\mathbf{e}_T\right) &= \sum_{w\in K} \underbrace{(-1)^w}_{\substack{=\prod_{j=1}^{k}(-1)^{w^{\langle j\rangle}} \\ \text{(by (176))}}} \underbrace{\prod_{(i,j)\in D}}_{\substack{=\prod_{j=1}^{k}\prod_{i\in D\langle j\rangle} \\ \text{(by (169))}}} \underbrace{x^{i-1}_{(wT)(i,j)}}_{\substack{=x^{i-1}_{w^{\langle j\rangle}(T(i,j))} \\ \text{(since (175) shows} \\ \text{that } (wT)(i,j)=w^{\langle j\rangle}(T(i,j)))}} \\
&= \sum_{w\in K} \underbrace{\left(\prod_{j=1}^{k}(-1)^{w^{\langle j\rangle}}\right)\prod_{j=1}^{k}\prod_{i\in D\langle j\rangle} x^{i-1}_{w^{\langle j\rangle}(T(i,j))}}_{=\prod_{j=1}^{k}\left((-1)^{w^{\langle j\rangle}}\prod_{i\in D\langle j\rangle} x^{i-1}_{w^{\langle j\rangle}(T(i,j))}\right)} \\
&= \sum_{w\in K}\prod_{j=1}^{k}\left((-1)^{w^{\langle j\rangle}}\prod_{i\in D\langle j\rangle} x^{i-1}_{w^{\langle j\rangle}(T(i,j))}\right) \\
&= \sum_{(w_1,w_2,\ldots,w_k)\in S_{X_1}\times S_{X_2}\times\cdots\times S_{X_k}}\prod_{j=1}^{k}\left((-1)^{w_j}\prod_{i\in D\langle j\rangle} x^{i-1}_{w_j(T(i,j))}\right) \\
&\qquad\qquad \left(\begin{array}{c}\text{here, we have substituted } (w_1,w_2,\ldots,w_k) \\ \text{for } \left(w^{\langle 1\rangle},w^{\langle 2\rangle},\ldots,w^{\langle k\rangle}\right) \text{ in the sum, since the} \\ \text{map (174) is a bijection}\end{array}\right) \\
&= \prod_{j=1}^{k}\sum_{w\in S_{X_j}}(-1)^w\prod_{i\in D\langle j\rangle} x^{i-1}_{w(T(i,j))} \qquad (177)
\end{aligned}$$

(by the product rule, applied backwards[58]).

Now, we shall prove the following:

*Claim 5.* Let $j\in[k]$. Then,

$$\det\left(x^{p-1}_{T(i,j)}\right)_{i,p\in D\langle j\rangle} = \sum_{w\in S_{X_j}}(-1)^w\prod_{i\in D\langle j\rangle} x^{i-1}_{w(T(i,j))}. \qquad (178)$$

*Proof of Claim 5.* We recall a basic property of determinants: The determinant of a square matrix does not change if we relabel its rows and its columns (as

[58] i.e., since the product rule shows that

$$\begin{aligned}
&\prod_{j=1}^{k}\sum_{w\in S_{X_j}}(-1)^w\prod_{i\in D\langle j\rangle} x^{i-1}_{w(T(i,j))} \\
&= \sum_{(w_1,w_2,\ldots,w_k)\in S_{X_1}\times S_{X_2}\times\cdots\times S_{X_k}}\prod_{j=1}^{k}\left((-1)^{w_j}\prod_{i\in D\langle j\rangle} x^{i-1}_{w_j(T(i,j))}\right)
\end{aligned}$$

long as both rows and columns are relabelled simultaneously using the same bijection). In other words, if $\alpha : P \to Q$ is a bijection between two finite sets $P$ and $Q$, and if $\left(a_{p,q}\right)_{p,q\in Q}$ is a $Q \times Q$-matrix, then

$$\det\left(a_{p,q}\right)_{p,q\in Q} = \det\left(a_{\alpha(p),\alpha(q)}\right)_{p,q\in P}. \tag{179}$$

Next, we recall that $T$ is an $n$-tableau, hence injective. In other words, all entries $T(i,j)$ of $T$ are distinct.

Now, recall that $D\langle j\rangle$ is the set of all integers $i$ such that $(i,j) \in D$. Thus, for each $i \in D\langle j\rangle$, the entry $T(i,j)$ of $T$ is a well-defined entry of the $j$-th column of $T$. In other words, for each $i \in D\langle j\rangle$, we have $T(i,j) \in \operatorname{Col}(j,T)$ (since $\operatorname{Col}(j,T)$ is the set of all entries of the $j$-th column of $T$). In other words, for each $i \in D\langle j\rangle$, we have $T(i,j) \in X_j$ (since $X_j = \operatorname{Col}(j,T)$). Hence, the map

$$\begin{aligned} \alpha : D\langle j\rangle &\to X_j, \\ i &\mapsto T(i,j) \end{aligned}$$

is well-defined. This map $\alpha$ is furthermore injective (since all entries $T(i,j)$ of $T$ are distinct) and surjective (since each $x \in X_j$ is an entry of the $j$-th column of $T$ (because $x \in X_j = \operatorname{Col}(j,T)$) and thus has the form $T(i,j)$ for some $i \in D\langle j\rangle$). Thus, it is bijective. In other words, $\alpha : D\langle j\rangle \to X_j$ is a bijection.

Hence, (179) (applied to $P = D\langle j\rangle$ and $Q = X_j$ and $a_{p,q} = x_p^{\alpha^{-1}(q)-1}$) yields

$$\begin{aligned} &\det\left(x_p^{\alpha^{-1}(q)-1}\right)_{p,q\in X_j} \\ &= \det\left(x_{\alpha(p)}^{\alpha^{-1}(\alpha(q))-1}\right)_{p,q\in D\langle j\rangle} \\ &= \det\left(x_{\alpha(p)}^{q-1}\right)_{p,q\in D\langle j\rangle} \qquad \left(\text{since } \alpha^{-1}(\alpha(q)) = q \text{ for each } q \in D\langle j\rangle\right) \\ &= \det\left(x_{\alpha(i)}^{p-1}\right)_{i,p\in D\langle j\rangle} \qquad \left(\begin{array}{c}\text{here, we have renamed the indices } p \text{ and } q \\ \text{as } i \text{ and } p\end{array}\right) \\ &= \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle} \qquad \left(\begin{array}{c}\text{since } \alpha(i) = T(i,j) \text{ for each } i \in D\langle j\rangle \\ \text{(by the definition of } \alpha)\end{array}\right). \end{aligned}$$

Thus,

$$\det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle} = \det\left(x_p^{\alpha^{-1}(q)-1}\right)_{p,q\in X_j} = \det\left(x_q^{\alpha^{-1}(p)-1}\right)_{p,q\in X_j}$$

(since the determinant of a matrix equals the determinant of its transpose).

Therefore,

$$\begin{aligned}
\det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle} &= \det\left(x_q^{\alpha^{-1}(p)-1}\right)_{p,q\in X_j} \\
&= \sum_{\sigma\in S_{X_j}} (-1)^\sigma \prod_{p\in X_j} x_{\sigma(p)}^{\alpha^{-1}(p)-1} \qquad \left(\begin{array}{c}\text{by the definition of}\\ \text{a determinant}\end{array}\right) \\
&= \sum_{\sigma\in S_{X_j}} (-1)^\sigma \prod_{i\in D\langle j\rangle} \underbrace{x_{\sigma(\alpha(i))}^{\alpha^{-1}(\alpha(i))-1}}_{\substack{=x_{\sigma(T(i,j))}^{i-1} \\ (\text{since } \alpha^{-1}(\alpha(i))=i \\ \text{and } \alpha(i)=T(i,j) \\ (\text{by the definition of } \alpha))}} \\
&\qquad\qquad \left(\begin{array}{c}\text{here, we have substituted } \alpha(i) \text{ for } p \text{ in}\\ \text{the product, since the map } \alpha : D\langle j\rangle \to X_j\\ \text{is a bijection}\end{array}\right) \\
&= \sum_{\sigma\in S_{X_j}} (-1)^\sigma \prod_{i\in D\langle j\rangle} x_{\sigma(T(i,j))}^{i-1} \\
&= \sum_{w\in S_{X_j}} (-1)^w \prod_{i\in D\langle j\rangle} x_{w(T(i,j))}^{i-1}
\end{aligned}$$

(here, we have renamed the summation index $\sigma$ as $w$). This proves Claim 5. □

Now, (177) becomes

$$\begin{aligned}
\beta\left(\mathbf{e}_T\right) &= \prod_{j=1}^{k} \underbrace{\sum_{w\in S_{X_j}} (-1)^w \prod_{i\in D\langle j\rangle} x_{w(T(i,j))}^{i-1}}_{\substack{=\det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle} \\ (\text{by (178)})}} \\
&= \prod_{j=1}^{k} \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle} \\
&= \prod_{j\in\mathbb{Z}} \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle} \qquad (\text{by (170)}).
\end{aligned}$$

In other words, $\beta$ sends the polytabloid $\mathbf{e}_T$ to the product $\prod_{j\in\mathbb{Z}} \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle}$. This proves the first claim of Theorem 5.6.1 **(b)**.

It remains to prove the second claim – which says that the map $\beta$ sends the Specht module $\mathcal{S}^D$ to the span of the products shown in (159). But this is easy: The map $\beta$ is $\mathbf{k}$-linear, and sends each polytabloid $\mathbf{e}_T$ to the product shown in (159) (as we have just shown). Thus, it sends the span of the polytabloids $\mathbf{e}_T$ to the span of the products shown in (159). In other words, it sends the Specht

module $\mathcal{S}^D$ to the span of the products shown in (159) (since the Specht module $\mathcal{S}^D$ is precisely the span of the polytabloids $\mathbf{e}_T$). The proof of Theorem 5.6.1 **(b)** is thus complete.

**(c)** Let $T$ be an $n$-tableau of shape $D$. Define the sets $D\langle j\rangle$ as in Theorem 5.6.1 **(b)**. Note that each integer $j \leq 0$ satisfies $D\langle j\rangle = \varnothing$ (since all cells of $D$ lie in columns $1, 2, 3, \ldots$) and therefore

$$\det \underbrace{\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle}}_{\substack{\text{a } 0\times 0\text{-matrix} \\ (\text{since } D\langle j\rangle = \varnothing)}} = 1. \tag{180}$$

But Theorem 5.6.1 **(b)** yields

$$\beta\left(\mathbf{e}_T\right) = \prod_{j\in\mathbb{Z}} \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle}. \tag{181}$$

In this product, all the factors with $j \leq 0$ equal 1 (by (180)). Hence, we can drop all these factors. Thus, (181) simplifies to

$$\beta\left(\mathbf{e}_T\right) = \prod_{j\geq 1} \det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle}. \tag{182}$$

Now, let $j$ be any positive integer. Then, the cells in the $j$-th column of $Y(\lambda)$ are $(1,j)$, $(2,j)$, $\ldots$, $\left(\lambda_j^t, j\right)$ (by Theorem 5.1.10 **(e)**). In other words, the integers $i$ that satisfy $(i,j) \in Y(\lambda)$ are $1, 2, \ldots, \lambda_j^t$. Since $D = Y(\lambda)$, we can rewrite this as follows: The integers $i$ that satisfy $(i,j) \in D$ are $1, 2, \ldots, \lambda_j^t$. In other words, $D\langle j\rangle = \left\{1, 2, \ldots, \lambda_j^t\right\}$ (since $D\langle j\rangle$ is the set of all integers $i$ that satisfy $(i,j) \in D$).

Set $q := \lambda_j^t$. Thus,

$$D\langle j\rangle = \left\{1, 2, \ldots, \lambda_j^t\right\} = \left[\lambda_j^t\right] = [q] \qquad \left(\text{since } \lambda_j^t = q\right).$$

Hence, the determinant $\det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle}$ is an instance of the well-known *Vandermonde determinant*. Let us recall what this means: The *Vandermonde determinant formula* ([21s, Theorem 6.4.31 **(c)**]) says that if $k \in \mathbb{N}$ is arbitrary, and if $a_1, a_2, \ldots, a_k$ are any $k$ elements of any commutative ring, then

$$\det\begin{pmatrix} a_1^0 & a_1^1 & \cdots & a_1^{k-1} \\ a_2^0 & a_2^1 & \cdots & a_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_k^0 & a_k^1 & \cdots & a_k^{k-1} \end{pmatrix} = \prod_{1\leq i_1<i_2\leq k} \left(a_{i_2} - a_{i_1}\right). \tag{183}$$

Thus,

$$\det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in[q]} = \det\begin{pmatrix} x_{T(1,j)}^{0} & x_{T(1,j)}^{1} & \cdots & x_{T(1,j)}^{q-1} \\ x_{T(2,j)}^{0} & x_{T(2,j)}^{1} & \cdots & x_{T(2,j)}^{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{T(q,j)}^{0} & x_{T(q,j)}^{1} & \cdots & x_{T(q,j)}^{q-1} \end{pmatrix}$$
$$= \prod_{1\le i_1<i_2\le q}\left(x_{T(i_2,j)} - x_{T(i_1,j)}\right)$$

(by (183), applied to $k = q$ and $a_i = x_{T(i,j)}$). In view of $[q] = D\langle j\rangle$ and $q = \lambda_j^t$, we can rewrite this as

$$\det\left(x_{T(i,j)}^{p-1}\right)_{i,p\in D\langle j\rangle} = \prod_{1\le i_1<i_2\le \lambda_j^t}\left(x_{T(i_2,j)} - x_{T(i_1,j)}\right). \tag{184}$$

Forget that we fixed $j$. We thus have proved (184) for each positive integer $j$. Therefore, (182) rewrites as

$$\beta\left(\mathbf{e}_T\right) = \prod_{j\ge 1}\ \prod_{1\le i_1<i_2\le \lambda_j^t}\left(x_{T(i_2,j)} - x_{T(i_1,j)}\right).$$

This proves Theorem 5.6.1 **(c)**. □

The Specht–Vandermonde avatar (Theorem 5.6.1 **(c)**) takes on a particularly simple form when $D$ is the Young diagram $Y(\lambda)$ of a hook partition $\lambda = (i, 1^j)$. In this case, Theorem 5.6.1 **(c)** leads to the following:

**Corollary 5.6.4.** Let $\lambda$ be a hook partition of $n$. (See Definition 5.1.9 for the notion of a hook partition.) Consider the injective $\mathbf{k}[S_n]$-module morphism $\beta : \mathcal{M}^D \to \mathbf{k}[x_1, x_2, \ldots, x_n]$ from Theorem 5.6.1. Then:

**(a)** The map $\beta$ sends any polytabloid $\mathbf{e}_T$ to

$$\prod_{1\le u<v\le \ell(\lambda)}\left(x_{T(v,1)} - x_{T(u,1)}\right).$$

**(b)** The image of the Specht module $\mathcal{S}^D$ under the morphism $\beta$ is the submodule

$$\operatorname{span}_{\mathbf{k}}\left\{\prod_{1\le u<v\le \ell(\lambda)}\left(x_{\sigma(v)} - x_{\sigma(u)}\right) \mid \sigma\in S_n\right\}$$

of $\mathbf{k}[x_1, x_2, \ldots, x_n]$.

*Proof.* Let $\lambda^t$ be the conjugate of the partition $\lambda$ (as defined in Theorem 5.1.10).

We know that $\lambda$ is a hook partition. In other words, $\lambda = (p, 1^q)$ for some positive integer $p$ and some nonnegative integer $q$. Thus, any column of $Y(\lambda)$ except for the first one has at most one cell. In other words, for each integer $j \geq 2$, we have

$$(\#\text{ of cells in the } j\text{-th column of } Y(\lambda)) \leq 1. \tag{185}$$

Hence, for each integer $j \geq 2$, we have

$$\begin{aligned} \lambda_j^t &= (\#\text{ of cells in the } j\text{-th column of } Y(\lambda)) \qquad (\text{by Theorem 5.1.10 } \textbf{(b)}) \\ &\leq 1 \qquad (\text{by (185)}). \end{aligned} \tag{186}$$

Furthermore, Theorem 5.1.10 **(c)** yields $\lambda_1^t = \ell(\lambda)$.

**(a)** The map $\beta$ sends any polytabloid $\mathbf{e}_T$ to

$$\prod_{j\geq 1} \prod_{1\leq i_1<i_2\leq \lambda_j^t} \left(x_{T(i_2,j)} - x_{T(i_1,j)}\right) \qquad (\text{by Theorem 5.6.1 } \textbf{(c)})$$

$$= \left(\prod_{1\leq i_1<i_2\leq \lambda_1^t} \left(x_{T(i_2,1)} - x_{T(i_1,1)}\right)\right) \cdot \prod_{j\geq 2} \underbrace{\prod_{1\leq i_1<i_2\leq \lambda_j^t} \left(x_{T(i_2,j)} - x_{T(i_1,j)}\right)}_{=(\text{empty product})}$$

(since (186) shows that $\lambda_j^t \leq 1$, and thus there are no two integers $i_1$ and $i_2$ satisfying $1 \leq i_1 < i_2 \leq \lambda_j^t$)

$$\left(\begin{array}{c}\text{here, we have split off the factor for } j=1 \text{ from} \\ \text{the outer product}\end{array}\right)$$

$$= \left(\prod_{1\leq i_1<i_2\leq \lambda_1^t} \left(x_{T(i_2,1)} - x_{T(i_1,1)}\right)\right) \cdot \prod_{j\geq 2} \underbrace{(\text{empty product})}_{=1}$$

$$= \left(\prod_{1\leq i_1<i_2\leq \lambda_1^t} \left(x_{T(i_2,1)} - x_{T(i_1,1)}\right)\right) \cdot \underbrace{\prod_{j\geq 2} 1}_{=1}$$

$$= \prod_{1\leq i_1<i_2\leq \lambda_1^t} \left(x_{T(i_2,1)} - x_{T(i_1,1)}\right)$$

$$= \prod_{1\leq u<v\leq \lambda_1^t} \left(x_{T(v,1)} - x_{T(u,1)}\right) \qquad \left(\begin{array}{c}\text{here, we have renamed the} \\ \text{indices } i_1 \text{ and } i_2 \text{ as } u \text{ and } v\end{array}\right)$$

$$= \prod_{1\leq u<v\leq \ell(\lambda)} \left(x_{T(v,1)} - x_{T(u,1)}\right) \qquad (\text{since } \lambda_1^t = \ell(\lambda)).$$

This proves Corollary 5.6.4 **(a)**.

**(b)** We have $D = Y(\lambda)$. Thus, the cells in the 1-st column of $D = Y(\lambda)$ are $(1,1)$, $(2,1)$, $\ldots$, $\left(\lambda_1^t, 1\right)$ (by Theorem 5.1.10 **(e)**, applied to $j = 1$).

Let $T$ be the $n$-tableau of shape $D$ that is obtained by filling the numbers $1, 2, \ldots, \lambda_1^t$ into these cells $(1,1)$, $(2,1)$, $\ldots$, $\left(\lambda_1^t, 1\right)$ and filling the remaining numbers $\lambda_1^t + 1$, $\lambda_1^t + 2$, $\ldots, n$ into the remaining cells of $D$ (say, from left to right, although the exact order is immaterial). Then,

$$T(j,1) = j \qquad \text{for each } j \in \left[\lambda_1^t\right]$$

(since we filled the numbers $1, 2, \ldots, \lambda_1^t$ into the cells $(1,1)$, $(2,1)$, $\ldots$, $\left(\lambda_1^t, 1\right)$). In other words,

$$T(j,1) = j \qquad \text{for each } j \in [\ell(\lambda)] \tag{187}$$

(since $\lambda_1^t = \ell(\lambda)$). Hence, for each $u, v \in [\ell(\lambda)]$, we have

$$\underbrace{x_{T(v,1)}}_{\substack{=x_v \\ \text{(since (187)} \\ \text{yields } T(v,1)=v)}} - \underbrace{x_{T(u,1)}}_{\substack{=x_u \\ \text{(since (187)} \\ \text{yields } T(u,1)=u)}} = x_v - x_u. \tag{188}$$

Now, the map $\beta$ sends the polytabloid $\mathbf{e}_T$ to

$$\prod_{1 \le u < v \le \ell(\lambda)} \underbrace{\left(x_{T(v,1)} - x_{T(u,1)}\right)}_{\substack{=x_v - x_u \\ \text{(by (188))}}} \qquad \text{(by Corollary 5.6.4 \textbf{(a)})}$$
$$= \prod_{1 \le u < v \le \ell(\lambda)} (x_v - x_u).$$

In other words,

$$\beta(\mathbf{e}_T) = \prod_{1 \le u < v \le \ell(\lambda)} (x_v - x_u). \tag{189}$$

But $\beta$ is a left $\mathbf{k}[S_n]$-module morphism. Hence, for each $w \in S_n$, we have

$$\beta(w\mathbf{e}_T) = w \cdot \beta(\mathbf{e}_T) = w \cdot \prod_{1 \le u < v \le \ell(\lambda)} (x_v - x_u) \qquad \text{(by (189))}$$
$$= \prod_{1 \le u < v \le \ell(\lambda)} \left(x_{w(v)} - x_{w(u)}\right) \tag{190}$$

(since $S_n$ acts on the polynomial ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$ by permuting the indeterminates). However, for each $w \in S_n$, we have $\mathbf{e}_{wT} = w\mathbf{e}_T$ (by Lemma 5.4.6 **(a)**) and thus

$$\beta(\mathbf{e}_{wT}) = \beta(w\mathbf{e}_T) = \prod_{1 \le u < v \le \ell(\lambda)} \left(x_{w(v)} - x_{w(u)}\right) \tag{191}$$

(by (190)).

Now, recall the equality (157), which we proved above. It says that

$$\mathcal{S}^D = \operatorname{span}_{\mathbf{k}} \left\{ \mathbf{e}_{wT} \mid w \in S_n \right\}.$$

Applying the map $\beta$ to it, we obtain

$$\begin{aligned}
\beta\left(\mathcal{S}^D\right) &= \beta\left(\operatorname{span}_{\mathbf{k}} \left\{ \mathbf{e}_{wT} \mid w \in S_n \right\}\right) \\
&= \operatorname{span}_{\mathbf{k}} \left\{ \beta\left(\mathbf{e}_{wT}\right) \mid w \in S_n \right\} \qquad \left( \begin{array}{c} \text{since the map } \beta \text{ is } \mathbf{k}\text{-linear,} \\ \text{and thus respects spans} \end{array} \right) \\
&= \operatorname{span}_{\mathbf{k}} \left\{ \prod_{1 \leq u < v \leq \ell(\lambda)} \left( x_{w(v)} - x_{w(u)} \right) \mid w \in S_n \right\} \qquad \text{(by (191))} \\
&= \operatorname{span}_{\mathbf{k}} \left\{ \prod_{1 \leq u < v \leq \ell(\lambda)} \left( x_{\sigma(v)} - x_{\sigma(u)} \right) \mid \sigma \in S_n \right\}
\end{aligned}$$

(here, we have renamed the index $w$ as $\sigma$). Thus, we have proved Corollary 5.6.4 **(b)**. □

Corollary 5.6.4 **(b)** has been used by Peel to study the Specht modules $\mathcal{S}^{Y(\lambda)}$ for hook partitions $\lambda$ in [Peel71].

### 5.6.2. The letterplace avatar

The Specht–Vandermonde avatar represents $n$-tabloids and polytabloids as polynomials in $n$ indeterminates $x_1, x_2, \ldots, x_n$. There is another avatar which is even simpler, in that it only uses multilinear polynomials (i.e., polynomials of degree $\leq 1$ in each variable). However, this simplicity is obtained at the cost of requiring more indeterminates: Namely, it needs $nk$ indeterminates $x_{1,1}, x_{1,2}, \ldots, x_{n,k}$ (the subscripts range over all pairs $(i,j)$ with $i \in [n]$ and $j \in [k]$), where $k$ is chosen such that all cells of $D$ lie in rows $1, 2, \ldots, k$. Other than this, it is very similar to the Specht–Vandermonde avatar, mainly differing in that each $x_{T(i,j)}^{i-1}$ is replaced by $x_{T(i,j),i}$ (so to speak, the exponent $i-1$ is "lowered" to become a second subscript):

**Theorem 5.6.5** (letterplace avatar of Young and Specht modules)**.** Let $D$ be a diagram with $|D| = n$. Let $k \in \mathbb{N}$ be such that all cells of $D$ lie in rows $1, 2, \ldots, k$.

Consider the polynomial ring

$$\begin{aligned}
&\mathbf{k}\left[ x_{i,j} \mid i \in [n] \text{ and } j \in [k] \right] \\
&= \mathbf{k}\left[ x_{1,1},\ x_{1,2},\ \ldots,\ x_{1,k},\ x_{2,1},\ x_{2,2},\ \ldots,\ x_{2,k},\ \ldots,\ x_{n,k} \right]
\end{aligned}$$

in $nk$ indeterminates that are named $x_{i,j}$ and indexed by the pairs $(i,j) \in [n] \times [k]$. This ring is known as the *letterplace algebra* (and the indeterminates

$x_{i,j}$ are called the *letterplace variables*, with the index $i$ being called the *letter* and the index $j$ being called the *place*). Let the symmetric group $S_n$ act on this ring by permuting the variables according to the rule

$$g \rightharpoonup x_{i,j} = x_{g(i),j} \qquad \text{for all } g \in S_n \text{ and } i \in [n] \text{ and } j \in [k].$$

(That is, a permutation $g \in S_n$ acts on a polynomial $p$ by replacing each indeterminate $x_{i,j}$ with $x_{g(i),j}$. This is called the *action on letters*.)

We define a $\mathbf{k}$-linear map

$$\begin{aligned} \gamma : \mathcal{M}^D &\to \mathbf{k}\left[x_{i,j} \mid i \in [n] \text{ and } j \in [k]\right], \\ \overline{T} &\mapsto \prod_{(i,j)\in D} x_{T(i,j),i} = \prod_{i\geq 1} \prod_{m\in \operatorname{Row}(i,T)} x_{m,i}, \end{aligned}$$

where $\operatorname{Row}(i,T)$ denotes the set of all entries in the $i$-th row of $T$. Then:

**(a)** This map $\gamma$ is an injective left $\mathbf{k}[S_n]$-module morphism.

**(b)** This map $\gamma$ sends any polytabloid $\mathbf{e}_T \in \mathcal{M}^D$ to the determinant product

$$\prod_{j\in\mathbb{Z}} \det\left(x_{T(i,j),p}\right)_{i,p\in D\langle j\rangle}, \tag{192}$$

where $D\langle j\rangle$ is the set of all integers $i$ that satisfy $(i,j) \in D$. (Note that the matrix $\left(x_{T(i,j),p}\right)_{i,p\in D\langle j\rangle}$ is thus a square matrix, whose rows and columns are indexed by the elements of $D\langle j\rangle$. Such square matrices have well-defined determinants, since $D\langle j\rangle$ is a finite set. When $D\langle j\rangle$ is empty, this matrix is a $0\times 0$-matrix and thus has determinant 1, by definition. Hence, all but finitely many factors of the product in (159) equal 1.)

Thus, the map $\gamma$ sends the Specht module $\mathcal{S}^D$ to the span of the products shown in (192).

**Example 5.6.6.** Let $n$, $D$ and $T$ be as in Example 5.6.2. Let $k = 3$. Then, the map $\gamma$ from Theorem 5.6.5 sends the $n$-tabloid $\overline{T} \in \mathcal{M}^D$ to the monomial

$$\begin{aligned} \prod_{(i,j)\in D} x_{T(i,j),i} &= x_{T(1,1),1}x_{T(1,3),1}x_{T(1,4),1}x_{T(2,1),2}x_{T(2,2),2}x_{T(3,1),2}x_{T(3,2),3}x_{T(3,4),3} \\ &= x_{4,1}x_{1,1}x_{3,1}x_{2,2}x_{8,2}x_{7,3}x_{6,3}x_{5,3}, \end{aligned}$$

and (by Theorem 5.6.1 **(b)**) sends the polytabloid $\mathbf{e}_T$ to the product

$$\begin{aligned}
&\prod_{j\in\mathbb{Z}} \det\left(x_{T(i,j),p}\right)_{i,p\in D\langle j\rangle} \\
&= \det\left(x_{T(i,1),p}\right)_{i,p\in D\langle 1\rangle} \cdot \det\left(x_{T(i,2),p}\right)_{i,p\in D\langle 2\rangle} \\
&\qquad \cdot \det\left(x_{T(i,3),p}\right)_{i,p\in D\langle 3\rangle} \cdot \det\left(x_{T(i,4),p}\right)_{i,p\in D\langle 4\rangle} \\
&\qquad\qquad \left( \begin{array}{c} \text{because for all } j \notin \{1,2,3,4\}\text{, the set } D\langle j\rangle \text{ is empty,} \\ \text{and thus } \det\left(x_{T(i,j),p}\right)_{i,p\in D\langle j\rangle} = 1 \end{array} \right) \\
&= \det\begin{pmatrix} x_{4,1} & x_{4,2} & x_{4,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \\ x_{7,1} & x_{7,2} & x_{7,3} \end{pmatrix} \cdot \det\begin{pmatrix} x_{8,2} & x_{8,3} \\ x_{6,2} & x_{6,3} \end{pmatrix} \cdot \det\begin{pmatrix} x_{1,1} \end{pmatrix} \cdot \det\begin{pmatrix} x_{3,1} & x_{3,3} \\ x_{5,1} & x_{5,3} \end{pmatrix}.
\end{aligned}$$

*Proof of Theorem 5.6.5.* Analogous to parts **(a)** and **(b)** of Theorem 5.6.1. □

The isomorphic avatar of Specht modules defined in Theorem 5.6.5 (or, rather, an insubstantial generalization thereof) has been used quite successfully to prove their properties, e.g., by Clausen in [Clause91] and by Krob in [Krob95, §2.2].

We have now seen four avatars of Young and Specht modules (and the tableau avatar depends on the choice of an $n$-tableau $T$, so it gives many different isomorphic copies).

## 5.7. The standard polytabloids are linearly independent

### 5.7.1. Introduction and main result

Young modules are easy to understand, at least as $\mathbf{k}$-modules: For any diagram $D$ of size $n$, the Young module $\mathcal{M}^D$ has a basis

$$\left(\overline{T}\right)_{\overline{T} \text{ is an } n\text{-tabloid of shape } D}$$

(as a $\mathbf{k}$-module), so it is a free $\mathbf{k}$-module of rank

$$(\#\text{ of } n\text{-tabloids of shape } D) = \frac{n!}{r_1!r_2!\cdots r_k!},$$

where $r_1, r_2, \ldots, r_k$ are the sizes of the nonempty rows of $D$ (why?). The symmetric group $S_n$ acts on $\mathcal{M}^D$ by permuting the $n$-tabloids.

The Specht module $\mathcal{S}^D$ is much more mysterious. It is a submodule of $\mathcal{M}^D$ spanned by the $n!$ polytabloids $\mathbf{e}_T$, but these polytabloids are heavily linearly

dependent, so it is not clear how many of them are redundant, i.e., what the dimension (rank) of $\mathcal{S}^D$ really is. In fact, it is not even clear whether $\mathcal{S}^D$ has a basis for all commutative rings $\mathbf{k}$.

**Question 5.7.1.** For a general diagram $D$ and a general commutative ring $\mathbf{k}$, is the $\mathbf{k}$-module $\mathcal{S}^D$ free? Does it have a basis independent of $\mathbf{k}$ ? Is its rank independent of $\mathbf{k}$ ? Is it a direct addend of $\mathcal{M}^D$ as a $\mathbf{k}$-module?

These questions have been open since the 1970s, and work has been done for various classes of diagrams $D$.

We shall give a complete (and positive) answer to this question when $D$ is a skew Young diagram: We will show that the so-called *standard polytabloids* form a basis of $\mathcal{S}^D$ in this case. We can define the standard polytabloids for any $D$ (we just cannot expect them to be a basis in this generality):

**Definition 5.7.2.** Let $D$ be any diagram with $|D| = n$. The *standard polytabloids* of shape $D$ will mean the polytabloids $\mathbf{e}_T \in \mathcal{S}^D$, where $T$ is a standard tableau of shape $D$. (Recall from Proposition 5.2.6 that any standard tableau $T$ of shape $D$ is an $n$-tableau, which is why $\mathbf{e}_T \in \mathcal{S}^D$ is well-defined.)

**Example 5.7.3.** The standard polytabloids of shape $Y(2,2)$ are $\mathbf{e}_{12\backslash\backslash 34}$ and $\mathbf{e}_{13\backslash\backslash 24}$, since the standard tableaux of this shape are $12\backslash\backslash 34$ and $13\backslash\backslash 24$.

We shall now make a first major step towards understanding the Specht module $\mathcal{S}^D$ by proving the following theorem:

**Theorem 5.7.4.** Let $D$ be any diagram. Then, the standard polytabloids of shape $D$ are $\mathbf{k}$-linearly independent.

### 5.7.2. The Young last letter order

The proof of Theorem 5.7.4 is not hard. It relies on a "triangularity-like argument" (or "leading term argument"). For this, we need a total order on the set of all $n$-tabloids of shape $D$:

**Definition 5.7.5.** **(a)** If $\overline{T}$ is any $n$-tabloid (of any shape), and if $i \in [n]$, then we let $r_{\overline{T}}(i)$ be the number of the row of $T$ that contains $i$. (Clearly, this depends only on $\overline{T}$, not on $T$, since we could just as well have written "the row of $\overline{T}$ that contains $i$" instead of "the row of $T$ that contains $i$".)

**(b)** Let $D$ be a diagram. Define a binary relation $<$ on the set $\{n\text{-tabloids of shape } D\}$ as follows:

Let $\overline{T}$ and $\overline{S}$ be two $n$-tabloids of shape $D$. Then, we declare that $\overline{T} < \overline{S}$ if and only if

- there exists at least one $i \in [n]$ such that $r_{\overline{T}}(i) \neq r_{\overline{S}}(i)$, and
- the **largest** such $i$ satisfies $r_{\overline{T}}(i) < r_{\overline{S}}(i)$.

In other words, we say that $\overline{T} < \overline{S}$ if the **largest** number that appears in different rows in $\overline{T}$ and $\overline{S}$ appears further north in $\overline{T}$ than in $\overline{S}$.

**Example 5.7.6.** Let $n = 7$ and $D = Y(3,2,2)$. Consider the two $n$-tableaux

$$T = \begin{array}{|c|c|c|}\hline 2 & 3 & 7 \\ \hline 4 & 6 \\ \cline{1-2} 1 & 5 \\ \cline{1-2}\end{array} \qquad \text{and} \qquad S = \begin{array}{|c|c|c|}\hline 3 & 5 & 7 \\ \hline 1 & 4 \\ \cline{1-2} 2 & 6 \\ \cline{1-2}\end{array}.$$

Then, $r_{\overline{T}}(2) = 1$ and $r_{\overline{S}}(2) = 3$ and so on. We have $r_{\overline{T}}(7) = r_{\overline{S}}(7) = 1$ but $r_{\overline{T}}(6) = 2 < 3 = r_{\overline{S}}(6)$. Thus, the **largest** number that appears in different rows in $\overline{T}$ and $\overline{S}$ is 6 and therefore appears further north in $\overline{T}$ than in $\overline{S}$. In other words, $\overline{T} < \overline{S}$.

It is easy to see that the relation $<$ that we have just defined is a total order:

**Proposition 5.7.7.** Let $D$ be a diagram. Then, the relation $<$ introduced in Definition 5.7.5 **(b)** is the smaller relation of a total order on the set $\{n\text{-tabloids of shape } D\}$.

This total order $<$ is called the *Young last letter order*.

The proof of Proposition 5.7.7 is fairly straightforward, and is built upon an even more obvious lemma:

**Lemma 5.7.8.** Let $P$ and $Q$ be two $n$-tableaux. Assume that $r_{\overline{P}}(i) = r_{\overline{Q}}(i)$ for each $i \in [n]$. Then, $P$ and $Q$ are row-equivalent.

*Proof of Lemma 5.7.8.* Let $k \in \mathbb{Z}$. Recall that $P$ is an $n$-tableau, and hence injective. In other words, all entries of $P$ are distinct. In particular, all entries in the $k$-th row of $P$ are distinct. In other words, the multiset of entries in the $k$-th row of $P$ is a set (i.e., contains no nontrivial multiplicities). Similarly, the multiset of entries in the $k$-th row of $Q$ is a set.

Let $u$ be an entry of the $k$-th row of $P$. Thus, the row of $P$ that contains $u$ is the $k$-th row. In other words, $r_{\overline{P}}(u) = k$ (by the definition of $r_{\overline{P}}(u)$). However, we know that $r_{\overline{P}}(i) = r_{\overline{Q}}(i)$ for each $i \in [n]$. Applying this to $i = u$, we obtain $r_{\overline{P}}(u) = r_{\overline{Q}}(u)$. Hence, $r_{\overline{Q}}(u) = r_{\overline{P}}(u) = k$. In other words, the row of $Q$ that contains $u$ is the $k$-th row (by the definition of $r_{\overline{Q}}(u)$). Hence, $u$ is an entry of the $k$-th row of $Q$.

Now, forget that we fixed $u$. We thus have shown that if $u$ is an entry of the $k$-th row of $P$, then $u$ is an entry of the $k$-th row of $Q$. In other words, each entry of the $k$-th row of $P$ is an entry of the $k$-th row of $Q$. In other words, the set of entries in the $k$-th row of $P$ is a subset of the set of entries in the $k$-th row of $Q$. But the same argument (with

the roles of $P$ and $Q$ interchanged) shows that the set of entries in the $k$-th row of $Q$ is a subset of the set of entries in the $k$-th row of $P$. Combining these two containment relations, we conclude that the set of entries in the $k$-th row of $P$ is the set of entries in the $k$-th row of $Q$. We can furthermore replace the word "set" by "multiset" in this sentence (since the multiset of entries in the $k$-th row of $P$ is a set, and the multiset of entries in the $k$-th row of $Q$ is a set). Thus, we have shown that the multiset of entries in the $k$-th row of $P$ is the multiset of entries in the $k$-th row of $Q$.

Now, forget that we fixed $k$. We have shown that for every $k \in \mathbb{Z}$, the multiset of entries in the $k$-th row of $P$ is the multiset of entries in the $k$-th row of $Q$. In other words, the tableaux $P$ and $Q$ have the same entries in corresponding rows. In other words, $P$ and $Q$ are row-equivalent. This proves Lemma 5.7.8. □

*Proof of Proposition 5.7.7.* In a nutshell: This total order is actually a disguised lexicographic order: We can encode any $n$-tabloid $\overline{T}$ of shape $D$ by the $n$-tuple $(r_{\overline{T}}(n), r_{\overline{T}}(n-1), \ldots, r_{\overline{T}}(1)) \in \mathbb{Z}^n$ (it is easy to see that this encoding is injective, i.e., we can reconstruct $\overline{T}$ from $(r_{\overline{T}}(n), r_{\overline{T}}(n-1), \ldots, r_{\overline{T}}(1))$), and then our relation $<$ is simply given by

$$\begin{aligned}&\left(\overline{T} < \overline{S}\right)\\ &\iff \left(\text{the encoding of } \overline{T} \text{ is lexicographically smaller than the encoding of } \overline{S}\right).\end{aligned}$$

Thus, $<$ is a total order, since the lexicographic order on $\mathbb{Z}^n$ is a total order.

But let me also give a more direct and detailed proof:

First, we shall show that the relation $<$ is the smaller relation of a partial order on $\{n\text{-tabloids of shape } D\}$. In order to prove this, we need to verify three claims:

> *Claim 1:* The relation $<$ is irreflexive (i.e., there exists no $n$-tabloid $\overline{T}$ such that $\overline{T} < \overline{T}$).
>
> *Claim 2:* The relation $<$ is transitive (i.e., if three $n$-tabloids $\overline{P}$, $\overline{Q}$ and $\overline{R}$ satisfy $\overline{P} < \overline{Q}$ and $\overline{Q} < \overline{R}$, then $\overline{P} < \overline{R}$).
>
> *Claim 3:* The relation $<$ is asymmetric (i.e., no two $n$-tabloids $\overline{P}$ and $\overline{Q}$ satisfy $\overline{P} < \overline{Q}$ and $\overline{Q} < \overline{P}$ at the same time).

*Proof of Claim 1.* Let $\overline{T}$ be an $n$-tabloid such that $\overline{T} < \overline{T}$. We shall derive a contradiction.

Indeed, we have $\overline{T} < \overline{T}$. By the definition of the relation $<$, this means that

- there exists at least one $i \in [n]$ such that $r_{\overline{T}}(i) \neq r_{\overline{T}}(i)$, and
- the **largest** such $i$ satisfies $r_{\overline{T}}(i) < r_{\overline{T}}(i)$.

But the first of these two bullet points is clearly absurd (since $r_{\overline{T}}(i) \neq r_{\overline{T}}(i)$ can never happen). Thus, we have found a contradiction.

Forget that we fixed $\overline{T}$. We thus have obtained a contradiction for each $n$-tabloid $\overline{T}$ such that $\overline{T} < \overline{T}$. Hence, there exists no such $\overline{T}$. In other words, the relation $<$ is irreflexive. This proves Claim 1. □

*Proof of Claim 2.* Let $\overline{P}$, $\overline{Q}$ and $\overline{R}$ be three $n$-tabloids that satisfy $\overline{P} < \overline{Q}$ and $\overline{Q} < \overline{R}$. We must prove that $\overline{P} < \overline{R}$.

Indeed, we have $\overline{P} < \overline{Q}$. By the definition of the relation $<$, this means that

- there exists at least one $i \in [n]$ such that $r_{\overline{P}}(i) \neq r_{\overline{Q}}(i)$, and
- the **largest** such $i$ satisfies $r_{\overline{P}}(i) < r_{\overline{Q}}(i)$.

Let us denote this largest $i$ by $u$. Thus, we have $r_{\overline{P}}(u) < r_{\overline{Q}}(u)$, whereas

$$\text{each } i > u \text{ satisfies } r_{\overline{P}}(i) = r_{\overline{Q}}(i) \tag{193}$$

(since $u$ is the **largest** $i \in [n]$ that satisfies $r_{\overline{P}}(i) \neq r_{\overline{Q}}(i)$).

Furthermore, we have $\overline{Q} < \overline{R}$. By the definition of the relation $<$, this means that

- there exists at least one $i \in [n]$ such that $r_{\overline{Q}}(i) \neq r_{\overline{R}}(i)$, and
- the **largest** such $i$ satisfies $r_{\overline{Q}}(i) < r_{\overline{R}}(i)$.

Let us denote this largest $i$ by $v$. Thus, we have $r_{\overline{Q}}(v) < r_{\overline{R}}(v)$, whereas

$$\text{each } i > v \text{ satisfies } r_{\overline{Q}}(i) = r_{\overline{R}}(i) \tag{194}$$

(since $v$ is the **largest** $i \in [n]$ that satisfies $r_{\overline{Q}}(i) \neq r_{\overline{R}}(i)$).

Now, let $w := \max\{u, v\}$. Then, $w \geq u$ and $w \geq v$. Hence,

$$\text{each } i > w \text{ satisfies } r_{\overline{P}}(i) = r_{\overline{R}}(i) \tag{195}$$

[59]. Furthermore,

$$r_{\overline{P}}(w) < r_{\overline{R}}(w) \tag{196}$$

[60] and therefore $r_{\overline{P}}(w) \neq r_{\overline{R}}(w)$. Hence, $w$ is an $i \in [n]$ that satisfies $r_{\overline{P}}(i) \neq r_{\overline{R}}(i)$. Moreover, $w$ is the **largest** such $i$ (since (195) shows that all larger $i$'s satisfy $r_{\overline{P}}(i) = r_{\overline{R}}(i)$). Hence, we have now shown that

[59] *Proof.* Let $i \in [n]$ be such that $i > w$. Then, $i > w \geq u$, so that $r_{\overline{P}}(i) = r_{\overline{Q}}(i)$ (by (193)). Also, $i > w \geq v$, so that $r_{\overline{Q}}(i) = r_{\overline{R}}(i)$ (by (194)). Hence, $r_{\overline{P}}(i) = r_{\overline{Q}}(i) = r_{\overline{R}}(i)$. This proves (195).

[60] *Proof of (196):* We are in one of the following three cases:

*Case 1:* We have $u < v$.

*Case 2:* We have $u = v$.

*Case 3:* We have $u > v$.

Let us first consider Case 1. In this case, we have $u < v$. Hence, $w = \max\{u, v\} = v$ (since $u < v$). But we know that $r_{\overline{Q}}(v) < r_{\overline{R}}(v)$. Since $w = v$, we can rewrite this as $r_{\overline{Q}}(w) < r_{\overline{R}}(w)$. Furthermore, $w = v > u$ (since $u < v$) and thus $r_{\overline{P}}(w) = r_{\overline{Q}}(w)$ (by (193), applied to $i = u$). Thus, $r_{\overline{P}}(w) = r_{\overline{Q}}(w) < r_{\overline{R}}(w)$. Hence, (196) is proved in Case 1.

Let us next consider Case 2. In this case, we have $u = v$. Hence, $w = \max\{u, v\} = u$ (since $u = v$). But we know that $r_{\overline{Q}}(v) < r_{\overline{R}}(v)$. Since $w = u = v$, we can rewrite this as $r_{\overline{Q}}(w) < r_{\overline{R}}(w)$. Furthermore, we know that $r_{\overline{P}}(u) < r_{\overline{Q}}(u)$. Since $w = u$, we can rewrite this as $r_{\overline{P}}(w) < r_{\overline{Q}}(w)$. Thus, $r_{\overline{P}}(w) < r_{\overline{Q}}(w) < r_{\overline{R}}(w)$. Hence, (196) is proved in Case 2.

Let us finally consider Case 3. In this case, we have $u > v$. Hence, $w = \max\{u, v\} = u$ (since $u > v$). But we know that $r_{\overline{P}}(u) < r_{\overline{Q}}(u)$. Since $w = u$, we can rewrite this as $r_{\overline{P}}(w) < r_{\overline{Q}}(w)$. Furthermore, $w = u > v$ and thus $r_{\overline{Q}}(w) = r_{\overline{R}}(w)$ (by (194), applied to $i = w$). Thus, $r_{\overline{P}}(w) < r_{\overline{Q}}(w) = r_{\overline{R}}(w)$. Hence, (196) is proved in Case 3.

Thus, (196) is proved in all three Cases 1, 2 and 3.

- there exists at least one $i \in [n]$ such that $r_{\overline{P}}(i) \neq r_{\overline{R}}(i)$ (namely, $i = w$), and
- the **largest** such $i$ satisfies $r_{\overline{P}}(i) < r_{\overline{R}}(i)$ (because the largest such $i$ is $w$, and we know that $w$ satisfies $r_{\overline{P}}(w) < r_{\overline{R}}(w)$).

In other words, $\overline{P} < \overline{R}$ (by the definition of the relation $<$). This proves Claim 2. □

*Proof of Claim 3.* Any binary relation that is irreflexive and transitive is automatically asymmetric (since $\overline{P} < \overline{Q}$ and $\overline{Q} < \overline{P}$ would entail $\overline{P} < \overline{P}$ by transitivity, but this would contradict irreflexivity). Since the relation $<$ is irreflexive (by Claim 1) and transitive (by Claim 2), we thus conclude that it is also asymmetric. This proves Claim 3. □

Having proved Claims 1, 2 and 3, we thus conclude that the relation $<$ is the smaller relation of a partial order. It remains to show that this partial order is a total order. For this purpose, we need to prove the following last claim:

> *Claim 4:* Let $\overline{P}$ and $\overline{Q}$ be two distinct $n$-tabloids of shape $D$. Then, we have $\overline{P} < \overline{Q}$ or $\overline{Q} < \overline{P}$.

*Proof of Claim 4.* We shall first show that there exists some $i \in [n]$ such that $r_{\overline{P}}(i) \neq r_{\overline{Q}}(i)$.

Indeed, assume the contrary. Thus, $r_{\overline{P}}(i) = r_{\overline{Q}}(i)$ for each $i \in [n]$. Hence, Lemma 5.7.8 yields that $P$ and $Q$ are row-equivalent. In other words, $\overline{P} = \overline{Q}$ (since an $n$-tabloid is an equivalence class of $n$-tableaux with respect to row-equivalence). But this contradicts the fact that $\overline{P}$ and $\overline{Q}$ are distinct.

This contradiction shows that our assumption was false. Hence, we have shown that there exists some $i \in [n]$ such that $r_{\overline{P}}(i) \neq r_{\overline{Q}}(i)$.

Consider the **largest** such $i$. Then, this $i$ satisfies $r_{\overline{P}}(i) \neq r_{\overline{Q}}(i)$. Hence, either $r_{\overline{P}}(i) < r_{\overline{Q}}(i)$ or $r_{\overline{P}}(i) > r_{\overline{Q}}(i)$. In the former case, we have $\overline{P} < \overline{Q}$ (by the definition of the relation $<$). In the latter case, we have $\overline{Q} < \overline{P}$ (by the definition of the relation $<$ again[61]). Thus, we have found that $\overline{P} < \overline{Q}$ or $\overline{Q} < \overline{P}$. This proves Claim 4. □

Claim 4 shows that the partial order whose smaller relation is $<$ is actually a total order. The proof of Proposition 5.7.7 is thus complete. □

Now we claim the following:

**Lemma 5.7.9.** Let $D$ be a diagram. Let $T$ be a column-standard $n$-tableau of shape $D$. Let $<$ be the relation introduced in Definition 5.7.5. Then:

**(a)** Every $w \in \mathcal{C}(T)$ satisfying $w \neq \mathrm{id}$ satisfies $\overline{wT} < \overline{T}$.

**(b)** We have

$$\mathbf{e}_T = \overline{T} + \left(\text{a linear combination of } n\text{-tabloids } \overline{S} \text{ with } \overline{S} < \overline{T}\right).$$

[61] *Proof.* Assume that $r_{\overline{P}}(i) > r_{\overline{Q}}(i)$. Thus, $r_{\overline{Q}}(i) < r_{\overline{P}}(i)$. But recall that our $i$ is the **largest** $i \in [n]$ such that $r_{\overline{P}}(i) \neq r_{\overline{Q}}(i)$. In other words, our $i$ is the **largest** $i \in [n]$ such that $r_{\overline{Q}}(i) \neq r_{\overline{P}}(i)$. Hence, there exists an $i \in [n]$ such that $r_{\overline{Q}}(i) \neq r_{\overline{P}}(i)$, and the **largest** such $i$ satisfies $r_{\overline{Q}}(i) < r_{\overline{P}}(i)$. In other words, we have $\overline{Q} < \overline{P}$ (by the definition of the relation $<$), qed.

*Proof.* **(a)** Let $w \in \mathcal{C}(T)$ be such that $w \neq \mathrm{id}$. Then, there exists some $k \in [n]$ such that $w(k) \neq k$ (since $w \neq \mathrm{id}$). Pick the **largest** such $k$. Then,

$$w(i) = i \qquad \text{for all } i > k. \tag{197}$$

Here is the rest of the proof in a nutshell: The permutation $w$ is vertical for $T$, and thus must permute the set of entries of the column of $T$ that contains the entry $k$. All entries south of this entry $k$ stay unchanged under this permutation (by (197)), but the entry $k$ does not (since $w(k) \neq k$). Hence, this entry $k$ must move further north. In other words, the entry $k$ lies further north in $wT$ than it does in $T$. But all entries larger than $i$ lie in the same row (and, in fact, in the same cell) in $wT$ as they do in $T$ (again by (197)). Thus, the **largest** number that appears in different rows in $\overline{wT}$ and $\overline{T}$ is the number $k$, and this number appears further north in $\overline{wT}$ than in $\overline{T}$. By the definition of the relation $<$, this means that $\overline{wT} < \overline{T}$, and thus Lemma 5.7.9 **(a)** is proved.

Here is a rigorous way to formulate this proof:

Let $(u, v)$ be the cell of $T$ that contains the entry $k$. Thus, $T(u, v) = k$. Hence, the number $k$ lies in the $u$-th row and in the $v$-th column of $T$. In particular, $r_{\overline{T}}(k) = u$.

Let $(u', v')$ be the cell of the $n$-tableau $wT$ that contains the entry $k$. Thus, $(wT)(u', v') = k$. Hence, the number $k$ lies in the $u'$-th row and in the $v'$-th column of $wT$. In particular, $r_{\overline{wT}}(k) = u'$.

The definition of the action of $S_n$ on the $n$-tableaux yields $wT = w \circ T$, so that $(wT)(u', v') = w(T(u', v'))$. Hence, $w(T(u', v')) = (wT)(u', v') = k$. Therefore, $T(u', v') = w^{-1}(k)$. This shows that the number $w^{-1}(k)$ lies in the $v'$-th column of $T$.

Note that $w^{-1}(k) \neq k$ (since $w^{-1}(k) = k$ would yield $k = w(k) \neq k$, which is absurd). Hence, $T(u', v') = w^{-1}(k) \neq k = T(u, v)$, and thus $(u', v') \neq (u, v)$.

But the permutation $w$ is vertical for $T$ (since $w \in \mathcal{C}(T)$). Hence, for each $i \in [n]$, the number $w(i)$ lies in the same column of $T$ as $i$ does (by the definition of a vertical permutation). Applying this to $i = w^{-1}(k)$, we conclude that the number $w\left(w^{-1}(k)\right)$ lies in the same column of $T$ as $w^{-1}(k)$ does. In other words, the number $k$ lies in the same column of $T$ as $w^{-1}(k)$ does (since $w\left(w^{-1}(k)\right) = k$). In other words, we have $v' = v$ (since the number $w^{-1}(k)$ lies in the $v$-th column of $T$, whereas the number $k$ lies in the $v$-th column of $T$). Thus, $v = v'$. Hence, $(u', v) = (u', v') \in D$. and $T(u', v) = T(u', v') = w^{-1}(k)$.

If we had $u' = u$, then we would have $(u', v') = (u, v)$ (since $v = v'$), which would contradict $(u', v') \neq (u, v)$. Hence, we cannot have $u' = u$. Thus, $u' \neq u$.

We now claim that $u' < u$. Indeed, assume the contrary. Thus, $u' \geq u$, so that $u' > u$ (since $u' \neq u$). Hence, the cell $(u', v)$ lies strictly south of the cell $(u, v)$ (and we know that both cells belong to $D$). Since the entries of $T$ increase top-to-bottom down each column (because $T$ is column-standard), we thus conclude that $T(u', v) > T(u, v) = k$. Hence, (197) (applied to $i = T(u', v)$) yields $w(T(u', v)) = T(u', v)$. In view of $T(u', v) = w^{-1}(k)$, we can rewrite this as $w\left(w^{-1}(k)\right) = w^{-1}(k)$. Hence, $w^{-1}(k) = w\left(w^{-1}(k)\right) = k$. But this contradicts $w^{-1}(k) \neq k$.

This contradiction shows that our assumption was false. Hence, $u' < u$. In other words,

$$r_{\overline{wT}}(k) < r_{\overline{T}}(k)$$

(since $r_{\overline{T}}(k) = u$ and $r_{\overline{wT}}(k) = u'$). Thus, $r_{\overline{wT}}(k) \neq r_{\overline{T}}(k)$.

Moreover, using (197), we can easily see that each $i \in [n]$ satisfying $i > k$ satisfies $r_{\overline{wT}}(i) = r_{\overline{T}}(i)$ [62].

Thus, we conclude that $k$ is a number $i \in [n]$ such that $r_{\overline{wT}}(i) \neq r_{\overline{T}}(i)$ (since $r_{\overline{wT}}(k) \neq r_{\overline{T}}(k)$), but there exists no number $i \in [n]$ higher than $k$ that satisfies $r_{\overline{wT}}(i) \neq r_{\overline{T}}(i)$ (since each $i \in [n]$ satisfying $i > k$ satisfies $r_{\overline{wT}}(i) = r_{\overline{T}}(i)$). Hence, $k$ is the **largest** $i \in [n]$ such that $r_{\overline{wT}}(i) \neq r_{\overline{T}}(i)$. Thus, we have shown that

- there exists at least one $i \in [n]$ such that $r_{\overline{wT}}(i) \neq r_{\overline{T}}(i)$ (namely, $k$), and
- the **largest** such $i$ satisfies $r_{\overline{wT}}(i) < r_{\overline{T}}(i)$ (since this largest $i$ is $k$, and we know that $r_{\overline{wT}}(k) < r_{\overline{T}}(k)$).

In other words, $\overline{wT} < \overline{T}$ (by the definition of the relation $<$). Lemma 5.7.9 **(a)** is thus proved.

**(b)** The definition of $\mathbf{e}_T$ yields

$$\begin{aligned}
\mathbf{e}_T &= \sum_{w \in \mathcal{C}(T)} (-1)^w \, \overline{wT} \\
&= \underbrace{(-1)^{\mathrm{id}}}_{=1} \underbrace{\overline{\mathrm{id}\, T}}_{=\overline{T}} + \sum_{\substack{w \in \mathcal{C}(T); \\ w \neq \mathrm{id}}} (-1)^w \, \overline{wT} \\
&\qquad\qquad \left( \begin{array}{c} \text{here, we have split off the addend for } w = \mathrm{id} \\ \text{from the sum, since } \mathrm{id} \in \mathcal{C}(T) \end{array} \right) \\
&= \overline{T} + \underbrace{\sum_{\substack{w \in \mathcal{C}(T); \\ w \neq \mathrm{id}}} (-1)^w \, \overline{wT}}_{\substack{= \left( \text{a linear combination of } n\text{-tabloids } \overline{S} \text{ with } \overline{S} < \overline{T} \right) \\ \text{(since Lemma 5.7.9 (a) yields that } \overline{wT} < \overline{T} \\ \text{for each } w \in \mathcal{C}(T) \text{ satisfying } w \neq \mathrm{id}\text{)}}} \\
&= \overline{T} + \left( \text{a linear combination of } n\text{-tabloids } \overline{S} \text{ with } \overline{S} < \overline{T} \right).
\end{aligned}$$

The proof of Lemma 5.7.9 **(b)** is thus complete. □

### 5.7.3. Proof of the linear independence

We can now prove Theorem 5.7.4:

[62] *Proof.* Let $i \in [n]$ be such that $i > k$. We must prove that $r_{\overline{wT}}(i) = r_{\overline{T}}(i)$.

Indeed, let $(x, y)$ be the cell of $T$ that contains the entry $i$. Thus, $T(x, y) = i$ and $r_{\overline{T}}(i) = x$. But $wT = w \circ T$, so that $(wT)(x, y) = w\left(\underbrace{T(x, y)}_{=i}\right) = w(i) = i$ (by (197)). Thus, $(x, y)$ is the cell of $wT$ that contains $i$. Hence, $r_{\overline{wT}}(i) = x$. Comparing this with $r_{\overline{T}}(i) = x$, we obtain $r_{\overline{wT}}(i) = r_{\overline{T}}(i)$, qed.

*Proof of Theorem 5.7.4.* We must prove that the standard polytabloids $\mathbf{e}_T$ are $\mathbf{k}$-linearly independent.

Assume the contrary. Thus, they are $\mathbf{k}$-linearly dependent. That is, there exists a nontrivial relation

$$\sum_{T \in U} \alpha_T \mathbf{e}_T = 0 \tag{198}$$

for some nonempty set $U$ of standard $n$-tableaux of shape $D$ and some nonzero coefficients $\alpha_T \in \mathbf{k}$. Consider this $U$ and these coefficients $\alpha_T$.

Proposition 5.3.8 says that there is a bijection

$$\begin{aligned} &\text{from } \{\text{row-standard } n\text{-tableaux of shape } D\} \\ &\text{to } \{n\text{-tabloids of shape } D\} \end{aligned}$$

that sends each row-standard $n$-tableau $T$ to its tabloid $\overline{T}$. In particular, this bijection is injective. In other words, distinct row-standard $n$-tableaux $T$ of shape $D$ produce distinct $n$-tabloids $\overline{T}$. Hence, distinct standard $n$-tableaux $T$ produce distinct $n$-tabloids $\overline{T}$ (since any standard $n$-tableau is row-standard). Therefore, among the standard $n$-tableaux $T \in U$, there is a unique one for which the $n$-tabloid $\overline{T}$ is maximum with respect to the Young last letter order (since this order is a total order). Let $T_{\max}$ be this $n$-tableau. Thus,

$$\overline{T} < \overline{T_{\max}} \qquad \text{for each } T \in U \text{ satisfying } T \neq T_{\max}. \tag{199}$$

Now, recall that the family $\left(\overline{T}\right)_{\overline{T} \text{ is an } n\text{-tabloid of shape } D}$ is a basis of the $\mathbf{k}$-module $\mathcal{M}^D$. Thus, each element $\mathbf{a} \in \mathcal{M}^D$ can be uniquely expressed as a $\mathbf{k}$-linear combination $\sum_{\overline{S} \text{ is an } n\text{-tabloid of shape } D} a_{\overline{S}} \overline{S}$ of the $n$-tabloids of shape $D$. We shall denote the coefficients $a_{\overline{S}}$ appearing in this $\mathbf{k}$-linear combination by $\left[\overline{S}\right] \mathbf{a}$. In other words, we let $\left[\overline{S}\right] \mathbf{a}$ denote the $\overline{S}$-coordinate of an element $\mathbf{a} \in \mathcal{M}^D$ with respect to the basis $\left(\overline{T}\right)_{\overline{T} \text{ is an } n\text{-tabloid of shape } D}$.

Now, we claim the following:

*Claim 1:* Let $T \in U$ satisfy $T \neq T_{\max}$. Then, $\left[\overline{T_{\max}}\right] (\mathbf{e}_T) = 0$.

*Claim 2:* We have $\left[\overline{T_{\max}}\right] (\mathbf{e}_{T_{\max}}) = 1$.

Both of these claims follow easily from Lemma 5.7.9 **(b)** and (199). Namely:

*Proof of Claim 1.* From (199), we obtain $\overline{T} < \overline{T_{\max}}$, so that $\overline{T} \neq \overline{T_{\max}}$ and therefore $\left[\overline{T_{\max}}\right] \overline{T} = 0$. Moreover, $T$ is a standard $n$-tableau (since $T \in U$), thus a column-standard $n$-tableau. Hence, Lemma 5.7.9 **(b)** yields

$$\mathbf{e}_T = \overline{T} + \left(\text{a linear combination of } n\text{-tabloids } \overline{S} \text{ with } \overline{S} < \overline{T}\right).$$

In other words,

$$\mathbf{e}_T = \overline{T} + \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{with } \overline{S} < \overline{T}}} \gamma_{\overline{S}} \overline{S} \tag{200}$$

for some coefficients $\gamma_{\overline{S}} \in \mathbf{k}$. Using these coefficients $\gamma_{\overline{S}}$, we now obtain

$$\begin{aligned}\left[\overline{T_{\max}}\right](\mathbf{e}_T) &= \left[\overline{T_{\max}}\right]\left(\overline{T} + \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{with } \overline{S}<\overline{T}}} \gamma_{\overline{S}}\overline{S}\right) \qquad \text{(by (200))} \\ &= \underbrace{\left[\overline{T_{\max}}\right]\overline{T}}_{=0} + \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{with } \overline{S}<\overline{T}}} \gamma_{\overline{S}} \underbrace{\left[\overline{T_{\max}}\right]\overline{S}}_{\substack{=0 \\ (\text{since } \overline{S}<\overline{T}<\overline{T_{\max}} \\ \text{and thus } \overline{S}\neq\overline{T_{\max}})}} = 0 + \underbrace{\sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{with } \overline{S}<\overline{T}}} \gamma_{\overline{S}} 0}_{=0} = 0.\end{aligned}$$

This proves Claim 1. □

*Proof of Claim 2.* We know that $T_{\max}$ is a standard $n$-tableau (since $T_{\max} \in U$), thus a column-standard $n$-tableau. Hence, Lemma 5.7.9 **(b)** (applied to $T = T_{\max}$) yields

$$\mathbf{e}_{T_{\max}} = \overline{T_{\max}} + \left(\text{a linear combination of } n\text{-tabloids } \overline{S} \text{ with } \overline{S} < \overline{T_{\max}}\right).$$

In other words,

$$\mathbf{e}_{T_{\max}} = \overline{T_{\max}} + \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{with } \overline{S}<\overline{T_{\max}}}} \gamma_{\overline{S}}\overline{S} \tag{201}$$

for some coefficients $\gamma_{\overline{S}} \in \mathbf{k}$. Using these coefficients $\gamma_{\overline{S}}$, we now obtain

$$\begin{aligned}\left[\overline{T_{\max}}\right](\mathbf{e}_{T_{\max}}) &= \left[\overline{T_{\max}}\right]\left(\overline{T_{\max}} + \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{with } \overline{S}<\overline{T_{\max}}}} \gamma_{\overline{S}}\overline{S}\right) \qquad \text{(by (201))} \\ &= \underbrace{\left[\overline{T_{\max}}\right]\overline{T_{\max}}}_{=1} + \sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{with } \overline{S}<\overline{T_{\max}}}} \gamma_{\overline{S}} \underbrace{\left[\overline{T_{\max}}\right]\overline{S}}_{\substack{=0 \\ (\text{since } \overline{S}<\overline{T_{\max}} \\ \text{and thus } \overline{S}\neq\overline{T_{\max}})}} = 1 + \underbrace{\sum_{\substack{\overline{S} \text{ is an } n\text{-tabloid} \\ \text{with } \overline{S}<\overline{T_{\max}}}} \gamma_{\overline{S}} 0}_{=0} = 1.\end{aligned}$$

This proves Claim 2. □

Now,

$$\begin{aligned}\left[\overline{T_{\max}}\right]\left(\sum_{T\in U} \alpha_T \mathbf{e}_T\right) &= \sum_{T\in U} \alpha_T \left[\overline{T_{\max}}\right](\mathbf{e}_T) \\ &= \alpha_{T_{\max}} \underbrace{\left[\overline{T_{\max}}\right](\mathbf{e}_{T_{\max}})}_{\substack{=1 \\ \text{(by Claim 2)}}} + \sum_{\substack{T\in U; \\ T\neq T_{\max}}} \alpha_T \underbrace{\left[\overline{T_{\max}}\right](\mathbf{e}_T)}_{\substack{=0 \\ \text{(by Claim 1)}}} \\ &\qquad \left(\begin{array}{c}\text{here, we have split off the} \\ \text{addend for } T = T_{\max} \text{ from the sum}\end{array}\right) \\ &= \alpha_{T_{\max}} + \underbrace{\sum_{\substack{T\in U; \\ T\neq T_{\max}}} \alpha_T 0}_{=0} = \alpha_{T_{\max}} \neq 0\end{aligned}$$

(since all our coefficients $\alpha_T$ are nonzero). But this contradicts

$$\left[\overline{T_{\max}}\right] \underbrace{\left(\sum_{T \in U} \alpha_T \mathbf{e}_T\right)}_{\substack{=0 \\ \text{(by (198))}}} = \left[\overline{T_{\max}}\right] 0 = 0.$$

This contradiction shows that our assumption was false. Hence, the standard polytabloids $\mathbf{e}_T$ are $\mathbf{k}$-linearly independent. This proves Theorem 5.7.4. □

**Corollary 5.7.10.** Let $D$ be any diagram with $|D| = n$. Then, $\mathcal{S}^D \neq 0$ (unless $\mathbf{k}$ is trivial).

*Proof.* First, we shall show that there exists at least one standard $n$-tableau $T$ of shape $D$. Indeed, we can construct such a $T$ explicitly:

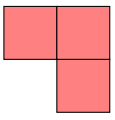
The *depth* of a cell $(i,j) \in D$ shall mean the integer $i+j$. It is clear that the depth of a cell increases (strictly) as we move east or south.

Now, number the $n$ cells of $D$ by $1, 2, \ldots, n$ in the order of increasing depth[63]. Define an $n$-tableau $T$ by filling each cell $c \in D$ with its number (i.e., we set $T(c)$ to be the number that we have just assigned to the cell $c$). This $n$-tableau $T$ is row-standard (because if a cell $(i,j)$ lies strictly west of a cell $(i,j')$, then $j < j'$, so that the cell $(i,j)$ has smaller depth than $(i,j')$, and thus the cell $(i,j)$ has a lower number than $(i,j')$; but this means that $T(i,j) < T(i,j')$) and column-standard (for similar reasons). Thus, it is standard. Hence, we have found a standard $n$-tableau $T$ of shape $D$.

Now consider the corresponding standard polytabloid $\mathbf{e}_T$. By Theorem 5.7.4, this $\mathbf{e}_T$ is part of a $\mathbf{k}$-linearly independent family of vectors. Hence, unless $\mathbf{k}$ is trivial, it follows that $\mathbf{e}_T \neq 0$, so that $\mathcal{S}^D \neq 0$ (since $\mathbf{e}_T$ belongs to $\mathcal{S}^D$). This proves Corollary 5.7.10. □

**Remark 5.7.11.** In general, the standard polytabloids $\mathbf{e}_T$ do **not** form a basis of $\mathcal{S}^D$. We will soon see (Theorem 5.9.1) that they do whenever $D$ is a skew Young diagram $Y(\lambda/\mu)$. But in general, there are not enough of them. For example, if $n = 3$ and

$$D = \{(1,1),\ (1,2),\ (2,2)\} = \quad ,$$



then there is only one standard tableau of shape $D$, but the Specht module $\mathcal{S}^D$ is a free $\mathbf{k}$-module of rank 2, so that a single polytabloid $\mathbf{e}_T$ is not enough to span it.

[63] There might be ties (i.e., distinct cells with equal depths), but these ties can be broken arbitrarily. For instance, you can choose to number cells of equal depth in the order of increasing x-coordinate (i.e., from northernmost to southernmost).

## 5.8. The Garnir relations

As we showed above, the standard polytabloids $\mathbf{e}_T$ in a Specht module $\mathcal{S}^D$ are linearly independent, but don't always span $\mathcal{S}^D$. Our goal is to show that they span $\mathcal{S}^D$ whenever $D$ is a skew Young diagram $Y(\lambda/\mu)$. To do this, we need a useful construction called the *Garnir relations*. This construction underlies much of the modern study of $S_n$-representations; see [Welsh92, §3.4], [Wildon18, §6.2], [James78, §7], [JamKer81, §7.2], [Sagan01, §2.6], [Howe22, §11.4], [LakBro18, §4.5] for various alternative treatments (some only handling straight Young diagrams), and [Liu16, §3.2], [Welsh92], [Murphy81], [DipJam85, §3] for generalizations, variants and applications.

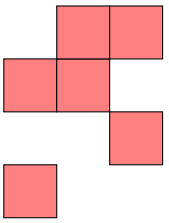
Before we define the Garnir relations, we need to introduce some simple but useful features of arbitrary diagrams.

### 5.8.1. Notations and some harmless lemmas

**Definition 5.8.1.** Let $D$ be a diagram. Let $j \in \mathbb{Z}$. Then, $D\langle j\rangle$ shall mean the set of all integers $i$ that satisfy $(i,j) \in D$.

For example, if

$$D = \{(1,2),\ (1,3),\ (2,1),\ (2,2),\ (3,3),\ (4,1)\} = $$



,

then $D\langle 1\rangle = \{2,4\}$ and $D\langle 2\rangle = \{1,2\}$ and $D\langle 3\rangle = \{1,3\}$ and $D\langle j\rangle = \varnothing$ for all $j \notin [3]$. If $D$ is a skew Young diagram $Y(\lambda/\mu)$, then each set $D\langle j\rangle$ is an integer interval (equalling $\left\{\mu^t_j+1,\ \mu^t_j+2,\ \ldots,\ \lambda^t_j\right\}$ when $j>0$ and being empty otherwise), but of course this does not generally hold for bad-shape diagrams.

**Definition 5.8.2.** Let $T$ be an $n$-tableau (of any shape). Let $j \in \mathbb{Z}$. Then, $\operatorname{Col}(j,T)$ shall denote the set of all entries in the $j$-th column of $T$.

For example, if $D = \begin{array}{|c|c|c|}\hline 2 & 1 & 4 \\ \hline 6 & 5 \\ \cline{1-2}\end{array}$, then $\operatorname{Col}(1,T) = \{2,6\}$ and $\operatorname{Col}(2,T) = \{1,5\}$ and $\operatorname{Col}(3,T) = \{4\}$ and $\operatorname{Col}(j,T) = \varnothing$ for all $j \notin [3]$.

The following should come as no surprise:

**Proposition 5.8.3.** Let $D$ be a diagram. Let $T$ be an $n$-tableau of shape $D$. Let $j \in \mathbb{Z}$. Then, $|\operatorname{Col}(j,T)| = |D\langle j\rangle|$.

*Proof.* We know that $T$ is an $n$-tableau, thus injective. Hence, all entries of $T$ are distinct. In particular, all entries in the $j$-th column of $T$ are distinct.

But $D\langle j\rangle$ is defined as the set of all integers $i$ that satisfy $(i,j) \in D$. Thus, the cells in the $j$-th column of $D$ are precisely the cells $(i,j)$ with $i \in D\langle j\rangle$. Hence, for any $i \in D\langle j\rangle$, the cell $(i,j)$ belongs to the $j$-th column of $D$, and therefore the corresponding entry $T(i,j)$ of $T$ is an entry of the $j$-th column of $T$ and thus belongs to $\mathrm{Col}(j,T)$ (by the definition of $\mathrm{Col}(j,T)$). Therefore, the map

$$\begin{aligned} D\langle j\rangle &\to \mathrm{Col}(j,T), \\ i &\mapsto T(i,j) \end{aligned}$$

is well-defined. This map is furthermore injective (since all entries in the $j$-th column of $T$ are distinct) and surjective (since each element of $\mathrm{Col}(j,T)$ is an entry in the $j$-th column of $T$, and thus has the form $T(i,j)$ for some $i \in D\langle j\rangle$). Thus, this map is bijective, i.e., a bijection. Hence, we have found a bijection from $D\langle j\rangle$ to $\mathrm{Col}(j,T)$. We thus conclude (by the bijection principle) that $|\mathrm{Col}(j,T)| = |D\langle j\rangle|$. This proves Proposition 5.8.3. □

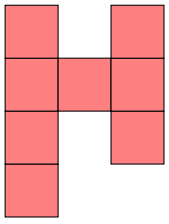
Somewhat (but not much) less obvious is the following equality in a Young module:

**Lemma 5.8.4.** Let $D$ be a diagram. Let $T$ be an $n$-tableau of shape $D$. Let $j$ and $k$ be two integers. Let $Z$ be a subset of $\mathrm{Col}(j,T) \cup \mathrm{Col}(k,T)$ such that $|Z| > |D\langle j\rangle \cup D\langle k\rangle|$. Then, in the Young module $\mathcal{M}^D$, we have

$$\nabla_Z^- \overline{T} = 0.$$

**Example 5.8.5.** Let $n = 8$ and

$$D = \{(1,1),\ (1,3),\ (2,1),\ (2,2),\ (2,3),\ (3,1),\ (3,3),\ (4,1)\}$$

$=$  .

Let $T$ be the $n$-tableau $\begin{array}{|c|c|c|}\hline 3 & & 6 \\ \hline 1 & 5 & 4 \\ \hline 7 & & 2 \\ \hline 8 \\ \cline{1-1}\end{array}$ of shape $D$. Let $Z = \{1,2,6,7,8\}$. Then, $Z$ is a subset of $\mathrm{Col}(1,T) \cup \mathrm{Col}(3,T)$ and satisfies $|Z| > |D\langle 1\rangle \cup D\langle 3\rangle|$ (since $|Z| = 5$ and $|D\langle 1\rangle \cup D\langle 3\rangle| = 4$). Hence, Lemma 5.8.4 yields $\nabla_Z^- \overline{T} = 0$. If

we color the cells of $T$ that contain the entries in $Z$ green, then $T$ looks as follows:

$$T = \begin{array}{|c|c|c|} \hline 3 & & 6 \\ \hline 1 & 5 & 4 \\ \hline 7 & & 2 \\ \hline 8 \\ \cline{1-1} \end{array} .$$

*Proof of Lemma 5.8.4.* We note that $|D| = n$ (since $T$ is an $n$-tableau of shape $D$).

For each $m \in [n]$, we let $r_T(m)$ be the number of the row of $T$ that contains the entry $m$.

We have $Z \subseteq \operatorname{Col}(j, T) \cup \operatorname{Col}(k, T)$ (by the definition of $Z$). Thus, each $m \in Z$ satisfies $r_T(m) \in D\langle j \rangle \cup D\langle k \rangle$ (this is easy to see[64]). Hence, the map

$$\begin{aligned} &Z \to D\langle j \rangle \cup D\langle k \rangle , \\ &m \mapsto r_T(m) \end{aligned}$$

is well-defined. By the pigeonhole principle, this map cannot be injective (since its domain $Z$ is larger than its target $D\langle j \rangle \cup D\langle k \rangle$ [65]). Thus, there exist two distinct elements $u, v \in Z$ satisfying $r_T(u) = r_T(v)$. Consider these $u$ and $v$. Thus, the entries $u$ and $v$ lie in the same row of the tableau $T$ [66]. Hence, the permutation $t_{u,v}$ is horizontal for $T$ [67]. In other words, $t_{u,v} \in \mathcal{R}(T)$. Hence,

[64] *Proof.* Let $m \in Z$. We must show that $r_T(m) \in D\langle j \rangle \cup D\langle k \rangle$.

We have $m \in Z \subseteq \operatorname{Col}(j, T) \cup \operatorname{Col}(k, T)$. Thus, we have $m \in \operatorname{Col}(j, T)$ or $m \in \operatorname{Col}(k, T)$. We WLOG assume that $m \in \operatorname{Col}(j, T)$ (since the other case is analogous). Thus, $m$ is an entry of the $j$-th column of $T$ (by the definition of $\operatorname{Col}(j, T)$). In other words, $m = T(i, j)$ for some $i \in \mathbb{Z}$ satisfying $(i, j) \in D$. Consider this $i$. From $(i, j) \in D$, we obtain $i \in D\langle j \rangle$ (by the definition of $D\langle j \rangle$). But $m = T(i, j)$ shows that the entry $m$ appears in the cell $(i, j)$ of $T$. Hence, the entry $m$ appears in the $i$-th row of $T$. In other words, $r_T(m) = i$ (by the definition of $r_T(m)$). Thus, $r_T(m) = i \in D\langle j \rangle \subseteq D\langle j \rangle \cup D\langle k \rangle$, qed.

[65] In more detail: One form of the pigeonhole principle says that if $X$ and $Y$ are two finite sets satisfying $|X| > |Y|$, then there exists no injective map from $X$ to $Y$. Applying this to $X = Z$ and $Y = D\langle j \rangle \cup D\langle k \rangle$, we see that there exists no injective map from $Z$ to $D\langle j \rangle \cup D\langle k \rangle$ (since $|Z| > |D\langle j \rangle \cup D\langle k \rangle|$). Hence, in particular, the map

$$\begin{aligned} &Z \to D\langle j \rangle \cup D\langle k \rangle , \\ &m \mapsto r_T(m) \end{aligned}$$

cannot be injective.

[66] *Proof.* Recall that $r_T(u)$ is defined as the number of the row of $T$ that contains the entry $u$, whereas $r_T(v)$ is defined as the number of the row of $T$ that contains the entry $v$. Hence, the equality $r_T(u) = r_T(v)$ (which we know to be true) is saying that the entries $u$ and $v$ lie in the same row of the tableau $T$.

[67] *Proof.* We need to show that for each $i \in [n]$, the number $t_{u,v}(i)$ lies in the same row of $T$ as $i$ does. But this is true for $i = u$ (since $t_{u,v}(u) = v$ lies in the same row of $T$ as $u$ does) and true for $i = v$ (for similar reasons), and also true for all remaining values of $i$ (since $t_{u,v}(i) = i$ whenever $i$ is neither $u$ nor $v$). Thus, we have shown that the permutation $t_{u,v}$ is horizontal for $T$.

Proposition 5.3.12 **(a)** (applied to $w = t_{u,v}$) shows that the tableau $t_{u,v} \rightharpoonup T$ is row-equivalent to $T$. In other words, $\overline{t_{u,v} \rightharpoonup T} = \overline{T}$ (since the $n$-tabloids are the equivalence classes of the $n$-tableaux with respect to row-equivalence). Thus, in $\mathcal{M}^D$, we have

$$(1 - t_{u,v})\,\overline{T} = \overline{T} - \underbrace{t_{u,v}\overline{T}}_{=\overline{t_{u,v}T} = \overline{t_{u,v} \rightharpoonup T} = \overline{T}} = \overline{T} - \overline{T} = 0.$$

Let us set

$$\nabla_Z^{\text{even}} := \sum_{\substack{w \in S_{n,Z}; \\ (-1)^w = 1}} w$$

(where $S_{n,Z}$ is defined according to Proposition 3.7.4).

However, $u$ and $v$ are two distinct elements of $Z$. Thus, the equality (32) from Proposition 3.7.4 **(d)** (applied to $Z$, $u$ and $v$ instead of $X$, $i$ and $j$) yields that

$$\nabla_Z^- = \nabla_Z^{\text{even}} \cdot (1 - t_{u,v}) = (1 - t_{u,v}) \cdot \nabla_Z^{\text{even}}.$$

Hence,

$$\underbrace{\nabla_Z^-}_{=\nabla_Z^{\text{even}} \cdot (1 - t_{u,v})}\ \overline{T} = \nabla_Z^{\text{even}} \cdot \underbrace{(1 - t_{u,v})\,\overline{T}}_{=0} = 0.$$

This proves Lemma 5.8.4. □

Lemma 5.8.4 easily yields the following:

**Lemma 5.8.6.** Let $D$ be a diagram. Let $S$ and $T$ be two column-equivalent $n$-tableaux of shape $D$. Let $j$ and $k$ be two integers. Let $Z$ be a subset of $\operatorname{Col}(j,T) \cup \operatorname{Col}(k,T)$ such that $|Z| > |D\langle j\rangle \cup D\langle k\rangle|$. Then, in the Young module $\mathcal{M}^D$, we have

$$\nabla_Z^- \overline{S} = 0.$$

*Proof.* The tableau $T$ is column-equivalent to $S$. Thus, it contains the same entries in its $j$-th column as $S$ does. In other words, $\operatorname{Col}(j,T) = \operatorname{Col}(j,S)$. Similarly, $\operatorname{Col}(k,T) = \operatorname{Col}(k,S)$.

But we assumed that $Z$ is a subset of $\operatorname{Col}(j,T) \cup \operatorname{Col}(k,T)$. In other words, $Z$ is a subset of $\operatorname{Col}(j,S) \cup \operatorname{Col}(k,S)$ (since $\operatorname{Col}(j,T) = \operatorname{Col}(j,S)$ and $\operatorname{Col}(k,T) = \operatorname{Col}(k,S)$). Hence, Lemma 5.8.4 (applied to $S$ instead of $T$) yields $\nabla_Z^- \overline{S} = 0$. This proves Lemma 5.8.6. □

In order to apply Lemma 5.8.6, we need to find a subset $Z$ of $\operatorname{Col}(j,T) \cup \operatorname{Col}(k,T)$ that satisfies the condition $|Z| > |D\langle j\rangle \cup D\langle k\rangle|$. It is easy to see that such a subset $Z$ can only exist if the sets $D\langle j\rangle$ and $D\langle k\rangle$ have a nonempty intersection – i.e., if $D$ has a row that has both a cell in the $j$-th column and a cell in the $k$-th column.

The following lemma will be particularly useful for us as we will apply Lemma 5.8.6 to skew Young tableaux:

**Lemma 5.8.7.** Let $D$ be the skew Young diagram $Y(\lambda/\mu)$ of a skew partition $\lambda/\mu$. Let $T$ be an $n$-tableau of shape $D$. Let $j$ and $k$ be two integers with $j < k$. Let $i \in \mathbb{Z}$ be such that $(i,j) \in D$ and $(i,k) \in D$. Define two sets

$$X = \left\{ T\left(i',j\right) \mid i' \in \mathbb{Z} \text{ with } i' \geq i \text{ and } \left(i',j\right) \in D \right\} \qquad \text{and}$$
$$Y = \left\{ T\left(i',k\right) \mid i' \in \mathbb{Z} \text{ with } i' \leq i \text{ and } \left(i',k\right) \in D \right\}.$$

(That is, $X$ is the set of all entries in the $j$-th column of $T$ at the cell $(i,j)$ and further south, whereas $Y$ is the set of all entries in the $k$-th column of $T$ at the cell $(i,k)$ and further north.) Then:

**(a)** We have $X \subseteq \operatorname{Col}(j,T)$ and $Y \subseteq \operatorname{Col}(k,T)$ and $|X \cup Y| > |D\langle j\rangle \cup D\langle k\rangle|$.

**(b)** If the tableau $T$ is column-standard and satisfies $T(i,j) > T(i,k)$, then any element of $X$ is larger than any element of $Y$.

**Example 5.8.8.** Let $n = 10$ and

$$D = Y((3,3,3,3,1)/(2,1)) = \ydiagram{3,3,3,3,1}*[*(white)]{3,3,3,3,1}*[*(white)]{2,1} \,,$$

and let $T$ be the $n$-tableau $\begin{array}{|c|c|c|}\hline \cdot & \cdot & 4 \\ \hline \cdot & 1 & 6 \\ \hline 2 & 5 & 8 \\ \hline 3 & 7 & 9 \\ \hline 10 \\ \cline{1-1}\end{array}$. Let $j = 2$ and $k = 3$ and $i = 3$. Then, the conditions of Lemma 5.8.7 are satisfied, and the sets $X$ and $Y$ in Lemma 5.8.7 are $X = \{5,7\}$ and $Y = \{4,6,8\}$. If we color the cells of $T$ that contain elements in $X$ green, and color the cells of $T$ that contain elements of $Y$ red, then $T$ looks as follows:

$$T = \begin{array}{|c|c|c|}\hline \cdot & \cdot & 4 \\ \hline \cdot & 1 & 6 \\ \hline 2 & 5 & 8 \\ \hline 3 & 7 & 9 \\ \hline 10 \\ \cline{1-1}\end{array} \,.$$

However, Lemma 5.8.7 **(b)** cannot be applied here, since the condition $T(i,j) > T(i,k)$ is not satisfied.

*Proof of Lemma 5.8.7.* **(a)** The definition of $X$ makes it clear that each element of $X$ is an entry $T(i',j)$ in the $j$-th column of $T$, thus an element of $\operatorname{Col}(j,T)$ (by the definition of $\operatorname{Col}(j,T)$). In other words, $X \subseteq \operatorname{Col}(j,T)$. Similarly, $Y \subseteq \operatorname{Col}(k,T)$.

Set $Z := X \cup Y$. Then, $Z$ is a subset of $\operatorname{Col}(j,T) \cup \operatorname{Col}(k,T)$ (since we have $Z = \underbrace{X}_{\subseteq \operatorname{Col}(j,T)} \cup \underbrace{Y}_{\subseteq \operatorname{Col}(k,T)} \subseteq \operatorname{Col}(j,T) \cup \operatorname{Col}(k,T)$) and satisfies $X \subseteq X \cup Y = Z$ and $Y \subseteq X \cup Y = Z$.

The definition of $X$ yields $T(i,j) \in X$ (since $i \in \mathbb{Z}$ and $i \geq i$ and $(i,j) \in D$). Similarly, $T(i,k) \in Y$. Thus, $T(i,j)$ and $T(i,k)$ are two elements of $Z$ (since $T(i,j) \in X \subseteq Z$ and $T(i,k) \in Y \subseteq Z$).

We know that $T$ is an $n$-tableau. Thus, $T$ is injective, i.e., all entries of $T$ are distinct. Thus, in particular, $T(i,j) \neq T(i,k)$ (since $(i,j) \neq (i,k)$ (because $j < k$)). Consequently, $T(i,j)$ and $T(i,k)$ are two distinct elements of $Z$.

For each $m \in [n]$, we let $r_T(m)$ be the number of the row of $T$ that contains the entry $m$. In particular, $r_T(T(i,j)) = i$ (since the entry $T(i,j)$ obviously lies in the $i$-th row of $T$) and similarly $r_T(T(i,k)) = i$. Hence, $r_T(T(i,j)) = i = r_T(T(i,k))$.

Now we claim the following:

> *Claim 1:* For each $i' \in D\langle j \rangle \cup D\langle k \rangle$, there exists some $m \in Z$ such that $r_T(m) = i'$.

*Proof of Claim 1.* Let $i' \in D\langle j \rangle \cup D\langle k \rangle$. We must prove that there exists some $m \in Z$ such that $r_T(m) = i'$.

We are in one of the following two cases:

*Case 1:* We have $i' \geq i$.

*Case 2:* We have $i' \leq i$.

Let us consider Case 1. In this case, we have $i' \geq i$. Thus, $i \leq i'$. Furthermore, we have $i' \in D\langle j \rangle \cup D\langle k \rangle$. This easily yields $i' \in D\langle j \rangle$ because $D$ is a skew Young diagram[68]. In other words, $(i',j) \in D$ (by the definition of $D\langle j \rangle$). Since $i' \geq i$, this entails that $T(i',j) \in X$ (by the definition of $X$) and therefore $T(i',j) \in X \subseteq Z$. Of course, the entry $T(i',j)$ of $T$ appears in the $i'$-th row. In other words, $r_T(T(i',j)) = i'$. Hence, there exists some $m \in Z$ such that $r_T(m) = i'$ (namely, $m = T(i',j)$). This proves Claim 1 in Case 1.

Now, consider Case 2. In this case, we have $i' \leq i$. Furthermore, we have $i' \in D\langle j \rangle \cup D\langle k \rangle$. This easily yields $i' \in D\langle k \rangle$ because $D$ is a skew Young diagram[69]. In other words, $(i',k) \in D$ (by the definition of $D\langle k \rangle$). Since $i' \leq i$, this entails that

[68] *Proof.* We must prove that $i' \in D\langle j \rangle$.

But we have $i' \in D\langle j \rangle \cup D\langle k \rangle$. Therefore, $i' \in D\langle j \rangle$ or $i' \in D\langle k \rangle$. In the former case, we are immediately done. Thus, we can WLOG assume that we are in the latter case. In other words, $i' \in D\langle k \rangle$. In other words, $(i',k) \in D$ (by the definition of $D\langle k \rangle$).

Now, recall the relation $\leq$ on cells introduced in Definition 5.1.2 **(c)**. The three cells $(i,j)$, $(i',j)$ and $(i',k)$ all belong to $\{1,2,3,\ldots\}^2$ and satisfy $(i,j) \leq (i',j)$ (since $i \leq i'$) and $(i',j) \leq (i',k)$ (since $j \leq k$) and $(i,j) \in D = Y(\lambda/\mu)$ and $(i',k) \in D = Y(\lambda/\mu)$. Hence, Proposition 5.1.6 (applied to $c = (i,j)$ and $d = (i',j)$ and $e = (i',k)$) yields $(i',j) \in Y(\lambda/\mu) = D$. Thus, $i' \in D\langle j \rangle$ (by the definition of $D\langle j \rangle$). This completes our proof.

[69] *Proof.* We must prove that $i' \in D\langle k \rangle$.

But we have $i' \in D\langle j \rangle \cup D\langle k \rangle$. Therefore, $i' \in D\langle j \rangle$ or $i' \in D\langle k \rangle$. In the latter case, we are immediately done. Thus, we can WLOG assume that we are in the former case. In other

$T(i',k) \in Y$ (by the definition of $Y$) and therefore $T(i',k) \in Y \subseteq Z$. Of course, the entry $T(i',k)$ of $T$ appears in the $i'$-th row. In other words, $r_T(T(i',k)) = i'$. Hence, there exists some $m \in Z$ such that $r_T(m) = i'$ (namely, $m = T(i',k)$). This proves Claim 1 in Case 2.

We have now proved Claim 1 in both Cases 1 and 2. Thus, Claim 1 always holds. □

Now, recall that we have already shown that $X \subseteq \operatorname{Col}(j,T)$ and $Y \subseteq \operatorname{Col}(k,T)$. It remains to prove that $|X \cup Y| > |D\langle j\rangle \cup D\langle k\rangle|$.

Assume the contrary. Thus, $|X \cup Y| \leq |D\langle j\rangle \cup D\langle k\rangle|$. In other words, $|Z| \leq |D\langle j\rangle \cup D\langle k\rangle|$ (since $Z = X \cup Y$). Therefore, one form of the pigeonhole principle shows that any surjective map from $Z$ to $D\langle j\rangle \cup D\langle k\rangle$ is bijective[70].

However, $Z$ is a subset of $\operatorname{Col}(j,T) \cup \operatorname{Col}(k,T)$. Thus, the map

$$\begin{aligned} Z &\to D\langle j\rangle \cup D\langle k\rangle, \\ m &\mapsto r_T(m) \end{aligned}$$

is well-defined (this has already been shown in the proof of Lemma 5.8.4). Claim 1 shows that this map is surjective. Hence, it is bijective (since any surjective map from $Z$ to $D\langle j\rangle \cup D\langle k\rangle$ is bijective), thus injective. Therefore, it must send the two distinct elements $T(i,j)$ and $T(i,k)$ of $Z$ to two distinct values. In other words, $r_T(T(i,j)) \neq r_T(T(i,k))$. But this contradicts $r_T(T(i,j)) = r_T(T(i,k))$. This contradiction shows that our assumption was false. Hence, $|X \cup Y| > |D\langle j\rangle \cup D\langle k\rangle|$ is proved. This completes the proof of Lemma 5.8.7 **(a)**.

**(b)** Assume that the tableau $T$ is column-standard and satisfies $T(i,j) > T(i,k)$. We must prove that any element of $X$ is larger than any element of $Y$. In other words, we must prove that $x > y$ for each $x \in X$ and each $y \in Y$.

So let $x \in X$ and $y \in Y$. We must prove that $x > y$.

We have $x \in X$. In other words, $x = T(i',j)$ for some $i' \in \mathbb{Z}$ satisfying $i' \geq i$ and $(i',j) \in D$ (by the definition of $X$). Consider this $i'$. The cell $(i',j)$ lies weakly south of the cell $(i,j)$ in the $j$-th column (since $i' \geq i$). Since the entries of $T$ strictly increase top-to-bottom in each column (because $T$ is column-standard), we thus conclude that $T(i',j) \geq T(i,j)$. Thus, $x = T(i',j) \geq T(i,j)$.

A similar argument can be used to show that $y \leq T(i,k)$ (since the cell of $T$ that contains $y$ lies weakly north of the cell $(i,k)$ in the $k$-th column). Thus, $T(i,k) \geq y$. Altogether, we now have $x \geq T(i,j) > T(i,k) \geq y$. Thus, $x > y$ is proved. As we said, this completes the proof of Lemma 5.8.7 **(b)**. □

---

words, $i' \in D\langle j\rangle$. In other words, $(i',j) \in D$ (by the definition of $D\langle j\rangle$).

Now, recall the relation $\leq$ on cells introduced in Definition 5.1.2 **(c)**. The three cells $(i',j)$, $(i',k)$ and $(i,k)$ all belong to $\{1,2,3,\ldots\}^2$ and satisfy $(i',j) \leq (i',k)$ (since $j \leq k$) and $(i',k) \leq (i,k)$ (since $i' \leq i$) and $(i',j) \in D = Y(\lambda/\mu)$ and $(i,k) \in D = Y(\lambda/\mu)$. Hence, Proposition 5.1.6 (applied to $c = (i',j)$ and $d = (i',k)$ and $e = (i,k)$) yields $(i',k) \in Y(\lambda/\mu) = D$. Thus, $i' \in D\langle k\rangle$ (by the definition of $D\langle k\rangle$). This completes our proof.

[70] *Proof.* One version of the pigeonhole principle says that if $P$ and $Q$ are two finite sets such that $|P| \leq |Q|$, then any surjective map from $P$ to $Q$ is bijective. Applying this to $P = Z$ and $Q = D\langle j\rangle \cup D\langle k\rangle$, we see that any surjective map from $Z$ to $D\langle j\rangle \cup D\langle k\rangle$ is bijective (since $|Z| \leq |D\langle j\rangle \cup D\langle k\rangle|$).

### 5.8.2. Transversals (systems of coset representatives) in groups

We will use two basic notions from group theory: that of a *left transversal*, and that of a *right transversal*. We briefly mentioned the latter in Remark 3.7.7, but let us now give a better definition:

**Definition 5.8.9.** Let $G$ be a group, and let $H$ be a subgroup of $G$. Then:

**(a)** A subset $L$ of $G$ is said to be a *left transversal of $H$ in $G$* (aka a *system of representatives for the left cosets of $H$ in $G$*) if it has the property that each left coset $uH$ of $H$ in $G$ contains exactly one element of $L$ (that is, each $u \in G$ satisfies $|uH \cap L| = 1$).

**(b)** A subset $R$ of $G$ is said to be a *right transversal of $H$ in $G$* (aka a *system of representatives for the right cosets of $H$ in $G$*) if it has the property that each right coset $Hu$ of $H$ in $G$ contains exactly one element of $R$ (that is, each $u \in G$ satisfies $|Hu \cap R| = 1$).

**Example 5.8.10.** Let $G$ be the symmetric group $S_3$, and let $H$ be its subgroup $\{1, s_1\}$. Then:

- The left cosets of $H$ in $G$ are $\{1, s_1\}$ and $\left\{s_2, \mathrm{cyc}_{1,3,2}\right\}$ and $\left\{t_{1,3}, \mathrm{cyc}_{1,2,3}\right\}$.
- The right cosets of $H$ in $G$ are $\{1, s_1\}$ and $\left\{s_2, \mathrm{cyc}_{1,2,3}\right\}$ and $\left\{t_{1,3}, \mathrm{cyc}_{1,3,2}\right\}$.

Thus:

- The set $\left\{1, s_2, \mathrm{cyc}_{1,2,3}\right\}$ is a left transversal of $H$ in $G$, but not a right transversal of $H$ in $G$ (since it contains two elements of the right coset $\left\{s_2, \mathrm{cyc}_{1,2,3}\right\}$ but no elements of the right coset $\left\{t_{1,3}, \mathrm{cyc}_{1,3,2}\right\}$).
- The set $\left\{1, s_2, \mathrm{cyc}_{1,3,2}\right\}$ is a right transversal of $H$ in $G$, but not a left transversal.
- The set $\{s_1, s_2, t_{1,3}\}$ is both a left transversal of $H$ in $G$ and a right transversal of $H$ in $G$.

Clearly, the notions of a left transversal and of a right transversal are mutually analogous, and in fact their roles get swapped if we reverse the order of multiplication in $G$. Thus, any result we can prove about one notion can be translated into a corresponding result about the other.

For instance, here is a nearly obvious (but important) fact:

**Proposition 5.8.11.** Let $G$ be a group, and let $H$ be a subgroup of $G$. Then:

**(a)** There exists at least one left transversal of $H$ in $G$.

**(b)** There exists at least one right transversal of $H$ in $G$.

*Proof sketch.* **(a)** The left cosets of $H$ in $G$ are nonempty and disjoint. Thus, we can obtain a left transversal simply by choosing one element from each left coset.

**(b)** Analogous to part **(a)**. □

What is much less obvious is that (as we have already mentioned in Remark 3.7.7) there exists a subset of $G$ that is simultaneously a left transversal and a right transversal of $H$ in $G$, provided that $G$ is finite (a theorem of Miller); but we will not need it.

We will, however, use the following simple fact:

**Proposition 5.8.12.** Let $G$ be a group, and let $H$ be a subgroup of $G$. Let $L$ be a left transversal of $H$ in $G$. Then, the map

$$\begin{aligned} H \times L &\to G, \\ (h, x) &\mapsto xh \end{aligned}$$

is a bijection.

*Proof sketch.* The map is clearly well-defined. Furthermore:

- This map is injective.

  [*Proof:* Let $(h, x) \in H \times L$ and $(h', x') \in H \times L$ be two pairs that this map sends to the same value. Thus, $xh = x'h'$. Now,

  $$x \underbrace{H}_{\substack{=hH \\ (\text{since } h \in H)}} = \underbrace{xh}_{=x'h'} H = x' \underbrace{h'H}_{\substack{=H \\ (\text{since } h' \in H)}} = x'H.$$

  But $L$ is a left transversal of $H$; thus, the left coset $xH$ contains exactly one element of $L$. Therefore, any two elements of $L$ contained in $xH$ must be equal. Since $x$ and $x'$ are two elements of $L$ contained in $xH$ (because $x \in xH$ and $x' \in x'H = xH$), we thus conclude that $x$ and $x'$ are equal. In other words, $x = x'$. Now, $xh = \underbrace{x'}_{=x} h' = xh'$, so that $h = h'$ (by cancelling $x$). Combined with $x = x'$, this yields $(h, x) = (h', x')$.

  Forget that we fixed $(h, x)$ and $(h', x')$. We thus have shown that if $(h, x) \in H \times L$ and $(h', x') \in H \times L$ are two pairs that our map sends to the same value, then $(h, x) = (h', x')$. In other words, our map is injective.]

- This map is surjective.

  [*Proof:* Let $g \in G$. Then, $gH$ is a left coset of $H$ in $G$. But $L$ is a left transversal of $H$; thus, the left coset $gH$ contains exactly one element of $L$. Let $x$ be this

element. Thus, $x \in L$ and $x \in gH$, so that $x = gh$ for some $h \in H$. Consider this $h$. Since $H$ is a subgroup of $G$, we obtain $h^{-1} \in H$ (since $h \in H$), so that $\left(h^{-1}, x\right) \in H \times L$ (since $x \in L$). But $x = gh$ leads to $g = xh^{-1}$. Thus, our map sends the pair $\left(h^{-1}, x\right) \in H \times L$ to $xh^{-1} = g$. In particular, our map takes $g$ as a value.

Forget that we fixed $g$. We thus have shown that our map takes each $g \in G$ as a value. In other words, this map is surjective.]

Hence, this map is both injective and surjective. Thus, it is bijective, i.e., a bijection. This proves Proposition 5.8.12. □

Obviously, Proposition 5.8.12 has an analogue for right transversals (with an analogous proof):

**Proposition 5.8.13.** Let $G$ be a group, and let $H$ be a subgroup of $G$. Let $R$ be a right transversal of $H$ in $G$. Then, the map

$$\begin{aligned} H \times R &\to G, \\ (h, x) &\mapsto hx \end{aligned}$$

is a bijection.

When $G$ is a subgroup of a symmetric group, these propositions have some useful consequences in the group algebra:

**Corollary 5.8.14.** Let $G$ be a subgroup of $S_n$. Let $H$ be a subgroup of $G$. Let $L$ be a left transversal of $H$ in $G$. Then, in $\mathbf{k}\left[S_n\right]$, we have

$$\sum_{g \in G} g = \left(\sum_{x \in L} x\right)\left(\sum_{h \in H} h\right) \tag{202}$$

and

$$\sum_{g \in G} (-1)^g g = \left(\sum_{x \in L} (-1)^x x\right)\left(\sum_{h \in H} (-1)^h h\right). \tag{203}$$

*Proof.* Proposition 5.8.12 says that the map

$$\begin{aligned} H \times L &\to G, \\ (h, x) &\mapsto xh \end{aligned}$$

is a bijection. Hence, we can substitute $xh$ for $g$ in the sum $\sum_{g \in G} (-1)^g g$. Thus

we obtain

$$\sum_{g\in G}(-1)^g g = \underbrace{\sum_{(h,x)\in H\times L}}_{=\sum_{h\in H}\sum_{x\in L}} \underbrace{(-1)^{xh}}_{\substack{=(-1)^x(-1)^h \\ \text{(by the multiplicativity} \\ \text{of the sign)}}} xh = \sum_{h\in H}\sum_{x\in L}(-1)^x(-1)^h xh$$
$$= \left(\sum_{x\in L}(-1)^x x\right)\left(\sum_{h\in H}(-1)^h h\right).$$

This proves (203). An analogous argument (but without the signs) proves (202). Thus, Corollary 5.8.14 is verified. □

Of course, the analogue of Corollary 5.8.14 for right transversals is just as easy to state and to prove:

**Corollary 5.8.15.** Let $G$ be a subgroup of $S_n$. Let $H$ be a subgroup of $G$. Let $R$ be a right transversal of $H$ in $G$. Then, in $\mathbf{k}[S_n]$, we have

$$\sum_{g\in G} g = \left(\sum_{h\in H} h\right)\left(\sum_{x\in R} x\right) \tag{204}$$

and

$$\sum_{g\in G}(-1)^g g = \left(\sum_{h\in H}(-1)^h h\right)\left(\sum_{x\in R}(-1)^x x\right). \tag{205}$$

**Exercise 5.8.1.** [2] Explain how Lemma 3.7.10 is a particular case of (204).

The following simple property of left transversals will be used in what follows:

**Lemma 5.8.16.** Let $G$ be a group. Let $H$ and $K$ be two subgroups of $G$. Let $L$ be a left transversal of $H\cap K$ in $H$. (This makes sense, since $H\cap K$ is a subgroup of $H$.) Then:

**(a)** There is exactly one $u\in L$ that belongs to $K$.

**(b)** This $u$ satisfies $L\setminus\{u\} = L\setminus K$.

*Proof.* We have $L\subseteq H$ (since $L$ is a left transversal of $H\cap K$ in $H$), thus $H\cap L = L$.

**(a)** Consider the trivial left coset $1(H\cap K)$ of the subgroup $H\cap K$ in $H$. This coset contains exactly one element of $L$ (since $L$ is a left transversal of $H\cap K$ in $H$, and therefore each left coset of $H\cap K$ in $H$ contains exactly one element of $L$). In other words, there is exactly one $u\in 1(H\cap K)$ that belongs to $L$.

In other words, there is exactly one $u \in 1\left(H \cap K\right) \cap L$. Since $\underbrace{1\left(H \cap K\right)}_{=H \cap K} \cap L = H \cap K \cap L = \underbrace{H \cap L}_{=L} \cap K = L \cap K$, we can rewrite this as follows: There is exactly one $u \in L \cap K$. In other words, there is exactly one $u \in L$ that belongs to $K$. This proves Lemma 5.8.16 **(a)**.

**(b)** Let $u$ be the unique $u \in L$ that belongs to $K$. Thus, $u$ is the only element of $L$ that belongs to $K$. In other words, $u$ is the only element of $L \cap K$. Hence, $L \cap K = \{u\}$.

But $L \setminus \left(L \cap K\right) = L \setminus K$ (since any two sets $X$ and $Y$ satisfy $X \setminus \left(X \cap Y\right) = X \setminus Y$). In view of $L \cap K = \{u\}$, we can rewrite this as $L \setminus \{u\} = L \setminus K$. This proves Lemma 5.8.16 **(b)**. □

### 5.8.3. The Garnir relations

We are now ready to state the *Garnir relations*:

**Theorem 5.8.17** (Garnir relations)**.** Let $D$ be a diagram. Let $T$ be an $n$-tableau of shape $D$. Let $j$ and $k$ be two distinct integers. Let $Z$ be a subset of $\operatorname{Col}\left(j, T\right) \cup \operatorname{Col}\left(k, T\right)$ such that $\left|Z\right| > \left|D\left\langle j\right\rangle \cup D\left\langle k\right\rangle\right|$.

Define $S_{n,Z}$ as in Proposition 3.7.4. This set $S_{n,Z}$ is a subgroup of $S_n$ (by Proposition 3.7.4 **(a)**, applied to $X = Z$). Thus, both $\mathcal{C}\left(T\right)$ and $S_{n,Z}$ are subgroups of $S_n$. Hence, their intersection $S_{n,Z} \cap \mathcal{C}\left(T\right)$ is a subgroup of $S_{n,Z}$. Let $L$ be a left transversal of $S_{n,Z} \cap \mathcal{C}\left(T\right)$ in $S_{n,Z}$. Then, in the Specht module $\mathcal{S}^D$, we have

$$\sum_{x \in L} \left(-1\right)^x \mathbf{e}_{xT} = 0 \tag{206}$$

and

$$\mathbf{e}_T = -\sum_{x \in L \setminus \mathcal{C}(T)} \left(-1\right)^x \mathbf{e}_{xT}. \tag{207}$$

The equality (206) is known as a *Garnir relation* (after H. G. Garnir, who discovered it in 1950 [Garnir50, §11, Théorème III][71]). The equality (207) is (as we will soon see) just a restatement of (206), in which all but one of the addends have been moved to the right hand side.

Before we prove Theorem 5.8.17, let us illustrate it on an example.

**Example 5.8.18.** Let $n = 8$ and $D = Y\left(3,3,2\right)$ and $T = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 6 & 3 & 7 \\ \hline 8 & 5 \\ \cline{1-2}\end{array}$. (This particular $T$ is straight and column-standard, although neither of these prop-

[71] To be more precise, Garnir only proved this theorem in the case when $D$ is a Young diagram $Y\left(\lambda\right)$ (and he stated it in terms of the Specht–Vandermonde avatar, i.e., as a determinantal identity). I don't know who generalized it to general skew Young diagrams.

erties is needed for Theorem 5.8.17 to hold.) Set $j = 1$ and $k = 2$ and $Z = \{2,3,6,8\}$ (the set of the entries highlighted in green). This satisfies all requirements of Theorem 5.8.17 (since $Z \subseteq \operatorname{Col}(1,T) \cup \operatorname{Col}(2,T)$ and $|Z| = 4$ and $|D\langle 1\rangle \cup D\langle 2\rangle| = 3$).

In order to apply Theorem 5.8.17, we need to choose a left transversal $L$ of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$. Let us find a good way to choose such an $L$. Note that

$$S_{n,Z} = \{w \in S_8 \mid w \text{ fixes } 1,4,5,7\} \qquad \text{and}$$
$$S_{n,Z} \cap \mathcal{C}(T) = \{w \in S_8 \mid w \text{ fixes } 1,4,5,7 \text{ and permutes } \{6,8\}\}$$

(why?). Thus, two permutations $u$ and $v$ in $S_{n,Z}$ belong to the same left coset of $S_{n,Z} \cap \mathcal{C}(T)$ if and only if we can write $u$ as $u = vh$ for some $h \in S_8$ that fixes $1,4,5,7$ and permutes $\{6,8\}$. In other words, two permutations $u$ and $v$ in $S_{n,Z}$ belong to the same left coset of $S_{n,Z} \cap \mathcal{C}(T)$ if and only if $u(\{6,8\}) = v(\{6,8\})$ (why?). Thus, each left coset of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$ consists of all permutations $w \in S_{n,Z}$ that send the set $\{6,8\}$ to some specific 2-element subset of $Z$. Therefore, in order to choose a left transversal $L$ of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$, we can proceed as follows: For each 2-element subset $U$ of $Z$, we choose some permutation $x_U \in S_{n,Z}$ that sends $\{6,8\}$ to $U$. The set of all these chosen permutations $x_U$ will then be our left transversal $L$.

There are many choices for these permutations $x_U$, but here is perhaps the simplest one (we are also listing the corresponding tableaux $x_U T$):

| $U$ | $\{6,8\}$ | $\{2,8\}$ | $\{2,6\}$ | $\{3,8\}$ | $\{3,6\}$ | $\{2,3\}$ |
|---|---|---|---|---|---|---|
| $x_U$ | id | $t_{2,6}$ | $t_{2,8}$ | $t_{3,6}$ | $t_{3,8}$ | $t_{2,6}t_{3,8}$ |
| $x_U T$ | $\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 4 \\ \hline 6 & 3 & 7 \\ \hline 8 & 5 \\ \cline{1-2}\end{array}$ | $\begin{array}{\|c\|c\|c\|}\hline 1 & 6 & 4 \\ \hline 2 & 3 & 7 \\ \hline 8 & 5 \\ \cline{1-2}\end{array}$ | $\begin{array}{\|c\|c\|c\|}\hline 1 & 8 & 4 \\ \hline 6 & 3 & 7 \\ \hline 2 & 5 \\ \cline{1-2}\end{array}$ | $\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 4 \\ \hline 3 & 6 & 7 \\ \hline 8 & 5 \\ \cline{1-2}\end{array}$ | $\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 4 \\ \hline 6 & 8 & 7 \\ \hline 3 & 5 \\ \cline{1-2}\end{array}$ | $\begin{array}{\|c\|c\|c\|}\hline 1 & 6 & 4 \\ \hline 2 & 8 & 7 \\ \hline 3 & 5 \\ \cline{1-2}\end{array}$ |

.

So we obtain the left transversal $L = \{\mathrm{id},\ t_{2,6},\ t_{2,8},\ t_{3,6},\ t_{3,8},\ t_{2,6}t_{3,8}\}$. The Garnir relation (206) thus becomes

$$\mathbf{e}_{\mathrm{id}\,T} - \mathbf{e}_{t_{2,6}T} - \mathbf{e}_{t_{2,8}T} - \mathbf{e}_{t_{3,6}T} - \mathbf{e}_{t_{3,8}T} + \mathbf{e}_{t_{2,6}t_{3,8}T} = 0.$$

Solving this for $\mathbf{e}_{\mathrm{id}\,T} = \mathbf{e}_T$, we obtain

$$\mathbf{e}_T = \mathbf{e}_{t_{2,6}T} + \mathbf{e}_{t_{2,8}T} + \mathbf{e}_{t_{3,6}T} + \mathbf{e}_{t_{3,8}T} - \mathbf{e}_{t_{2,6}t_{3,8}T}. \tag{208}$$

This is also what we would have obtained from (207).

**Remark 5.8.19.** Let $D$, $T$, $j$, $k$ and $Z$ be as in Theorem 5.8.17. Set $X = Z \cap \operatorname{Col}(j,T)$ and $Y = Z \cap \operatorname{Col}(k,T)$. There are many possible choices for the left transversal $L$, but are there any "natural" ones? Yes, and in fact there are

several:

1. We can choose $L$ to be the set of all permutations $w \in S_{n,Z}$ that are increasing on $X$ and increasing on $Y$. (These permutations are known as *riffle shuffles*.)

2. For any $|X|$-element subset $U$ of $Z$, we define an involution $w_U \in S_{n,Z}$ by $w_U = t_{u_1,x_1} t_{u_2,x_2} \cdots t_{u_p,x_p}$, where $u_1, u_2, \ldots, u_p$ are the elements of $U \setminus X$ listed in increasing order, and where $x_1, x_2, \ldots, x_p$ are the elements of $X \setminus U$ listed in increasing order. (Thus, $w_U$ exchanges the elements of $U \setminus X$ with those of $X \setminus U$ while leaving all other elements undisturbed.) We then can choose $L$ to be the set of all these involutions $w_U$ for all possible $U$. (This is how the $L$ in Example 5.8.18 was chosen.)

3. We can obtain variants of these choices by replacing "increasing" by "decreasing".

All these choices lead to left transversals $L$, which in general are different. Of course, they all satisfy $|L| = \dbinom{|X|+|Y|}{|X|}$.

*Proof of Theorem 5.8.17.* We note that $|D| = n$ (since $T$ is an $n$-tableau of shape $D$).

In the Young module $\mathcal{M}^D$, we have

$$\sum_{x \in L} (-1)^x \underbrace{\mathbf{e}_{xT}}_{= x\mathbf{e}_T \text{ (by Lemma 5.4.6 \textbf{(a)})}} = \sum_{x \in L} (-1)^x x \mathbf{e}_T = \left( \sum_{x \in L} (-1)^x x \right) \underbrace{\mathbf{e}_T}_{= \nabla^-_{\operatorname{Col} T} \overline{T} \text{ (by Proposition 5.5.12)}}$$

$$= \left( \sum_{x \in L} (-1)^x x \right) \nabla^-_{\operatorname{Col} T} \overline{T}. \tag{209}$$

Now, recall that both $\mathcal{C}(T)$ and $S_{n,Z}$ are subgroups of $S_n$. Hence, their intersection $S_{n,Z} \cap \mathcal{C}(T)$ is a subgroup of $\mathcal{C}(T)$. Proposition 5.8.11 **(b)** (applied to $G = \mathcal{C}(T)$ and $H = S_{n,Z} \cap \mathcal{C}(T)$) thus shows that there exists at least one right transversal of $S_{n,Z} \cap \mathcal{C}(T)$ in $\mathcal{C}(T)$. Pick such a right transversal and call it $R$. Thus, (205) (applied to $G = \mathcal{C}(T)$ and $H = S_{n,Z} \cap \mathcal{C}(T)$) yields

$$\sum_{g \in \mathcal{C}(T)} (-1)^g g = \left( \sum_{h \in S_{n,Z} \cap \mathcal{C}(T)} (-1)^h h \right) \left( \sum_{x \in R} (-1)^x x \right).$$

But the definition of $\nabla^-_{\operatorname{Col} T}$ yields

$$\begin{aligned}\nabla^-_{\operatorname{Col} T} &= \sum_{w \in \mathcal{C}(T)} (-1)^w w = \sum_{g \in \mathcal{C}(T)} (-1)^g g \\ &= \left( \sum_{h \in S_{n,Z} \cap \mathcal{C}(T)} (-1)^h h \right) \left( \sum_{x \in R} (-1)^x x \right). \end{aligned} \tag{210}$$

On the other hand, Proposition 3.7.4 **(b)** (applied to $X = Z$) yields

$$\nabla_Z = \sum_{w \in S_{n,Z}} w \qquad \text{and} \qquad \nabla^-_Z = \sum_{w \in S_{n,Z}} (-1)^w w.$$

Furthermore, (203) (applied to $G = S_{n,Z}$ and $H = S_{n,Z} \cap \mathcal{C}(T)$) yields that

$$\sum_{g \in S_{n,Z}} (-1)^g g = \left( \sum_{x \in L} (-1)^x x \right) \left( \sum_{h \in S_{n,Z} \cap \mathcal{C}(T)} (-1)^h h \right)$$

(since $L$ is a left transversal of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$). Thus,

$$\begin{aligned}\nabla^-_Z &= \sum_{w \in S_{n,Z}} (-1)^w w = \sum_{g \in S_{n,Z}} (-1)^g g \\ &= \left( \sum_{x \in L} (-1)^x x \right) \left( \sum_{h \in S_{n,Z} \cap \mathcal{C}(T)} (-1)^h h \right). \end{aligned} \tag{211}$$

Now, (209) becomes

$$\begin{aligned}\sum_{x \in L} (-1)^x \mathbf{e}_{xT} &= \left( \sum_{x \in L} (-1)^x x \right) \nabla^-_{\operatorname{Col} T} \overline{T} \\ &= \underbrace{\left( \sum_{x \in L} (-1)^x x \right) \left( \sum_{h \in S_{n,Z} \cap \mathcal{C}(T)} (-1)^h h \right)}_{= \nabla^-_Z \text{ (by (211))}} \underbrace{\left( \sum_{x \in R} (-1)^x x \right) \overline{T}}_{= \sum_{x \in R} (-1)^x x \overline{T}} \\ &\qquad \text{(by (210))} \\ &= \nabla^-_Z \sum_{x \in R} (-1)^x x \overline{T} \\ &= \sum_{x \in R} (-1)^x \nabla^-_Z x \overline{T}. \end{aligned} \tag{212}$$

Now, let $x \in R$. Then, $x \in R \subseteq \mathcal{C}(T)$ (by the definition of $R$). Hence, Proposition 5.3.12 **(b)** (applied to $w = x$) shows that the tableau $x \rightharpoonup T$ is column-equivalent to $T$. Thus, Lemma 5.8.6 (applied to $S = x \rightharpoonup T$) shows that $\nabla^-_Z \overline{x \rightharpoonup T} = 0$. In other words, $\nabla^-_Z x \overline{T} = 0$ (since $\overline{x \rightharpoonup T} = \overline{x \overline{T}} = x \overline{T}$).

Forget that we fixed $x$. We thus have shown that $\nabla_Z^- x\overline{T} = 0$ for each $x \in R$. Thus, (212) rewrites as

$$\sum_{x \in L} (-1)^x \mathbf{e}_{xT} = \sum_{x \in R} (-1)^x \underbrace{\nabla_Z^- x\overline{T}}_{=0} = 0.$$

This proves (206). It remains to prove (207).

Lemma 5.8.16 **(a)** (applied to $G = S_n$ and $H = S_{n,Z}$ and $K = \mathcal{C}(T)$) shows that there is exactly one $u \in L$ that belongs to $\mathcal{C}(T)$ (since $L$ is a left transversal of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$). Consider this $u$. Thus, $u \in \mathcal{C}(T)$. Hence, Lemma 5.4.6 **(c)** yields $\mathbf{e}_{uT} = (-1)^u \mathbf{e}_T$. Thus,

$$(-1)^u \mathbf{e}_{uT} = \underbrace{(-1)^u (-1)^u}_{=\left((-1)^2\right)^u = 1} \mathbf{e}_T = \mathbf{e}_T.$$

Moreover, Lemma 5.8.16 **(b)** (applied to $G = S_n$ and $H = S_{n,Z}$ and $K = \mathcal{C}(T)$) shows that $L \setminus \{u\} = L \setminus \mathcal{C}(T)$.

Now, (206) yields

$$\begin{aligned} 0 = \sum_{x \in L} (-1)^x \mathbf{e}_{xT} &= \underbrace{(-1)^u \mathbf{e}_{uT}}_{=\mathbf{e}_T} + \underbrace{\sum_{\substack{x \in L; \\ x \neq u}} (-1)^x \mathbf{e}_{xT}}_{= \sum_{x \in L \setminus \{u\}}} \\ &\qquad \text{(here, we have split off the addend for } x = u \text{ from the sum)} \\ &= \mathbf{e}_T + \sum_{x \in L \setminus \{u\}} (-1)^x \mathbf{e}_{xT}. \end{aligned}$$

Solving this for $\mathbf{e}_T$, we find

$$\mathbf{e}_T = -\sum_{x \in L \setminus \{u\}} (-1)^x \mathbf{e}_{xT} = -\sum_{x \in L \setminus \mathcal{C}(T)} (-1)^x \mathbf{e}_{xT}$$

(since $L \setminus \{u\} = L \setminus \mathcal{C}(T)$). This proves (207). Thus, the proof of Theorem 5.8.17 is complete. □

Theorem 5.8.17 gives a linear dependence relation (206) between several polytabloids $\mathbf{e}_{xT}$; this relation allows us to express any of these polytabloids through the others. In particular, we can use it (in the form (207)) to express $\mathbf{e}_T$ as a $\mathbf{k}$-linear combination of other polytabloids. In the next section, we shall use this to express the non-standard polytabloids through the standard ones when $D$ is a skew Young diagram.

**Remark 5.8.20.** In Theorem 5.8.17, the left transversal $L$ can be chosen at will. However, the resulting Garnir relation will not significantly depend on the choice, because of the following:

**Exercise 5.8.2.** $\boxed{1}$ Prove that any addend $(-1)^x \mathbf{e}_{xT}$ in (206) depends only on the left coset $x(S_{n,Z} \cap \mathcal{C}(T))$, but not on $x$ itself.

[**Hint:** Lemma 5.4.6 **(c)**.]

**Remark 5.8.21.** Theorem 5.8.17 can be easily generalized to a "multi-column Garnir relation", in which the two integers $j$ and $k$ are replaced by several integers. We will have no need for this generality, however.

## 5.9. The standard basis theorem

### 5.9.1. The theorem

We can now describe the linear structure of $\mathcal{S}^D$ completely when $D$ is a skew Young diagram:

**Theorem 5.9.1** (standard basis theorem)**.** Let $D$ be a skew Young diagram. Then, the standard polytabloids $\mathbf{e}_T$ (that is, the $\mathbf{e}_T$ where $T$ ranges over all standard tableaux of shape $D$) form a basis of the $\mathbf{k}$-module $\mathcal{S}^D$.

This theorem was first proved by Garnir in 1950 for straight Young diagrams $D = Y(\lambda)$ (see [Garnir50, §12]).

Before we properly prove Theorem 5.9.1, let me explain the rough idea of the proof: We already know (from Theorem 5.7.4) that the standard polytabloids $\mathbf{e}_T$ are $\mathbf{k}$-linearly independent. It remains to show that they span $\mathcal{S}^D$. It suffices to show that any polytabloid $\mathbf{e}_T$ can be written as a $\mathbf{k}$-linear combination of standard polytabloids. Furthermore, it suffices to prove this in the case when $T$ is column-standard. Indeed, if an $n$-tableau $T$ is not column-standard, then we can find a column-standard $n$-tableau $S$ that is column-equivalent to $T$ (just by sorting each column of $T$ in increasing order), and then Lemma 5.4.6 **(d)** yields $\mathbf{e}_T = \pm \mathbf{e}_S$, so that any set of vectors that spans $\mathbf{e}_S$ will also span $\mathbf{e}_T$.

Thus, we only need to show that any column-standard polytabloid $\mathbf{e}_T$ can be written as a $\mathbf{k}$-linear combination of standard polytabloids.

We try to leverage the Garnir relations (207) to show this. Here is an example:

**Example 5.9.2.** Let $n = 8$ and $D = Y(3,3,2)$ and $T = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 6 & 3 & 7 \\ \hline 8 & 5 \\ \cline{1-2}\end{array}$. This tableau $T$ is column-standard, but not standard, so we need to expand $\mathbf{e}_T$ as a $\mathbf{k}$-linear combination of standard polytabloids. To this purpose, we consider the two entries $T(2,1) = 6$ and $T(2,2) = 3$ of $T$, which witness the non-standardness of $T$ (since $6 > 3$). We denote the respective two cells $(2,1)$ and $(2,2)$ by $(i,j)$ and $(i,k)$ (so we set $i = 2$ and $j = 1$ and $k = 2$), and we

apply the Garnir relation (207) to the set $Z = X \cup Y$, where

$$X := \left\{ T\left(i', j\right) \mid i' \in \mathbb{Z} \text{ with } i' \geq i \text{ and } \left(i', j\right) \in D \right\} = \{6, 8\} \qquad \text{and}$$
$$Y := \left\{ T\left(i', k\right) \mid i' \in \mathbb{Z} \text{ with } i' \leq i \text{ and } \left(i', k\right) \in D \right\} = \{2, 3\}.$$

The result is the equality (208) (since our set $Z$ is precisely the set $Z = \{2, 3, 6, 8\}$ we used in Example 5.8.18). Thus, we have expressed our non-standard polytabloid $\mathbf{e}_T$ as a $\mathbf{k}$-linear combination (and, in fact, a $\mathbb{Z}$-linear combination) of the polytabloids $\mathbf{e}_{t_{2,6}T}$, $\mathbf{e}_{t_{2,8}T}$, $\mathbf{e}_{t_{3,6}T}$, $\mathbf{e}_{t_{3,8}T}$, $\mathbf{e}_{t_{2,6}t_{3,8}T}$.

Alas, these latter polytabloids are not all standard. For example, the tableau $t_{2,6}T$ is neither row-standard ($6 > 4$ and $8 > 5$) nor column-standard ($6 > 3$). Thus, we have not yet expressed $\mathbf{e}_T$ as a $\mathbf{k}$-linear combination of standard polytabloids.

But have we at least moved closer to this goal?

As we will soon see, the answer is "yes". To make this precise, we need to define an appropriate notion of "closer", i.e., a partial order on the set of all $n$-tableaux such that all our new tableaux $t_{2,6}T$, $t_{2,8}T$, ... (more generally: all the tableaux appearing on the right hand side of (207)) are larger than $T$. More precisely, we will introduce a total order on the set of all "$n$-column-tabloids" (an analogue of $n$-tabloids relying on column-equivalence instead of row-equivalence).

### 5.9.2. Column-tabloids

First, we define $n$-column-tabloids. The definition is completely analogous to Definition 5.3.6:

**Definition 5.9.3. (a)** Let $D$ be a diagram. An *$n$-column-tabloid* of shape $D$ is an equivalence class of $n$-tableaux of shape $D$ with respect to column-equivalence. For example, for $D = Y(2,1)$ and $n = 3$, the set $\{12\backslash\backslash 3,\ 32\backslash\backslash 1\}$ (where we are using the shorthand notation from Convention 5.2.4) is an $n$-column-tabloid of shape $D$.

**(b)** We use the notation $\widetilde{T}$ for the $n$-column-tabloid corresponding to an $n$-tableau $T$ (that is, the equivalence class containing $T$). For instance, the 3-column-tabloid $\{12\backslash\backslash 3,\ 32\backslash\backslash 1\}$ can thus be written as $\widetilde{12\backslash\backslash 3}$ or (equivalently) as $\widetilde{32\backslash\backslash 1}$.

To draw an $n$-column-tabloid, we just draw one of the $n$-tableaux in it (often we pick the column-standard one), and we drop the horizontal subdividers between the entries of a column. For instance, the 3-column-tabloid $\widetilde{12\backslash\backslash 3}$ is thus drawn as $\begin{vmatrix} 1 & 2 \\ 3 & \end{vmatrix}$.

**(c)** If $\widetilde{T}$ is an $n$-column-tabloid, and if $j \in \mathbb{Z}$, then the *$j$-th column of $\widetilde{T}$* shall

mean the set of the entries of the $j$-th column of $T$. This is well-defined, since two column-equivalent tableaux will have the same entries in their $j$-th column. For example, the 2-nd column of the 5-column-tabloid $\left|\begin{array}{c|c|c|c|c} 2 & 4 & 3 \\ 1 & 5 & \end{array}\right|$ is the set $\{4,5\}$.

**Example 5.9.4.** There are six 4-column-tabloids of shape $Y(2,2)$, namely

$$\left|\begin{array}{c|c|} 1 & 3 \\ 2 & 4 \end{array}\right|, \quad \left|\begin{array}{c|c|} 1 & 2 \\ 3 & 4 \end{array}\right|, \quad \left|\begin{array}{c|c|} 1 & 2 \\ 4 & 3 \end{array}\right|, \quad \left|\begin{array}{c|c|} 2 & 1 \\ 3 & 4 \end{array}\right|, \quad \left|\begin{array}{c|c|} 2 & 1 \\ 4 & 3 \end{array}\right|, \quad \left|\begin{array}{c|c|} 3 & 1 \\ 4 & 2 \end{array}\right|.$$

While $n$-column-tabloids are equivalence classes, they also can be represented in a canonical way, namely by column-standard $n$-tableaux:

**Proposition 5.9.5.** Let $D$ be any diagram. Then, there is a bijection

$$\begin{aligned} &\text{from } \{\text{column-standard } n\text{-tableaux of shape } D\} \\ &\text{to } \{n\text{-column-tabloids of shape } D\} \end{aligned}$$

that sends each column-standard $n$-tableau $T$ to its column-tabloid $\widetilde{T}$.

*Proof.* Analogous to Proposition 5.3.8. □

We could furthermore define an $S_n$-action on the set $\{n\text{-column-tabloids of shape } D\}$ (for a given diagram $D$), similarly to the analogous action for tabloids (see Definition 5.3.10). However, we will have no need for this action, so we leave this to the interested reader.

### 5.9.3. The column last letter order

Let us now define the column last letter order: an analogue of the Young last letter order (see Definition 5.7.5 and the paragraphs afterwards) for $n$-column-tabloids instead of $n$-tabloids.

**Definition 5.9.6. (a)** If $\widetilde{T}$ is any $n$-column-tabloid (of any shape), and if $i \in [n]$, then we let $c_{\widetilde{T}}(i)$ be the number of the column of $T$ that contains $i$. (Clearly, this depends only on $\widetilde{T}$, not on $T$, since we could just as well have written "the column of $\widetilde{T}$ that contains $i$" instead of "the column of $T$ that contains $i$".)

**(b)** Let $D$ be a diagram. Define a binary relation $<$ on the set $\{n\text{-column-tabloids of shape } D\}$ as follows:

Let $\widetilde{T}$ and $\widetilde{S}$ be two $n$-column-tabloids of shape $D$. Then, we declare that $\widetilde{T} < \widetilde{S}$ if and only if

- there exists at least one $i \in [n]$ such that $c_{\widetilde{T}}(i) \neq c_{\widetilde{S}}(i)$, and
- the **largest** such $i$ satisfies $c_{\widetilde{T}}(i) < c_{\widetilde{S}}(i)$.

In other words, we say that $\widetilde{T} < \widetilde{S}$ if the **largest** number that appears in different columns in $\widetilde{T}$ and $\widetilde{S}$ appears further left in $\widetilde{T}$ than in $\widetilde{S}$.

**Example 5.9.7.** If $\widetilde{T} = \begin{array}{|c|c|c|}\hline 5 & 4 & \\ \hline 2 & 1 & \\ \hline 3 & & \\ \hline \end{array}$ and $\widetilde{S} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 5 & 4 \\ \hline 3 & \\ \hline \end{array}$, then $\widetilde{T} < \widetilde{S}$, since the largest number that appears in different columns in $\widetilde{T}$ and $\widetilde{S}$ is 2, and this number appears further left in $\widetilde{T}$ than in $\widetilde{S}$.

The following proposition is an analogue to Proposition 5.7.7, and its proof is virtually identical to the proof of the latter:

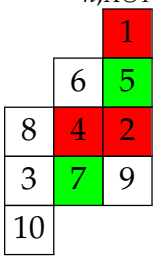
**Proposition 5.9.8.** Let $D$ be a diagram. Then, the relation $<$ introduced in Definition 5.9.6 **(b)** is the smaller relation of a total order on the set $\{n\text{-column-tabloids of shape } D\}$.

This total order $<$ is called the *column last letter order*. We define the relation $>$ on the set $\{n\text{-column-tabloids of shape } D\}$ accordingly (by setting $\widetilde{T} > \widetilde{S}$ if and only if $\widetilde{S} < \widetilde{T}$).

An important (to us) combinatorial property of the column last letter order will be the following lemma, which will help us show that the process illustrated in Example 5.9.2 is really bringing us closer to standardness:

**Lemma 5.9.9.** Let $T$ be an $n$-tableau (of any shape). Let $j$ and $k$ be two integers such that $j < k$. Let $X \subseteq \operatorname{Col}(j, T)$ and $Y \subseteq \operatorname{Col}(k, T)$ be two subsets such that each element of $X$ is larger than each element of $Y$. Let $w \in S_{n, X \cup Y} \setminus \mathcal{C}(T)$. Then, $\widetilde{wT} > \widetilde{T}$ (with respect to the column last letter order).

**Example 5.9.10.** Let $n = 10$ and $T = \begin{array}{|c|c|c|}\hline & & 1 \\ \hline & 6 & 4 \\ \hline 8 & 5 & 2 \\ \hline 3 & 7 & 9 \\ \hline 10 & & \\ \hline \end{array}$. Let $j = 2$ and $k = 3$ and $X = \{5, 7\}$ and $Y = \{1, 2, 4\}$. (We are highlighting the entries of $X$ in green and the entries of $Y$ in red.) Then, $j < k$ and $X \subseteq \operatorname{Col}(j, T)$ and $Y \subseteq \operatorname{Col}(k, T)$, and each element of $X$ is larger than each element of $Y$.

Hence, Lemma 5.9.9 yields that each $w \in S_{n,X\cup Y} \setminus \mathcal{C}(T)$ satisfies $\widetilde{wT} > \widetilde{T}$. For instance, for $w = t_{4,5}$, we have $wT =$  and thus visibly $\widetilde{wT} > \widetilde{T}$.



Lemma 5.9.9 is responsible for the fact that all the tableaux $t_{2,6}T$, $t_{2,8}T$, $t_{3,6}T$, $t_{3,8}T$ and $t_{2,6}t_{3,8}T$ in Example 5.9.2 are larger than $T$ (more precisely: that their $n$-column-tabloids are larger than $\widetilde{T}$). Hence, the formula (208), which expresses $\mathbf{e}_T$ as a linear combination of the polytabloids corresponding to these tableaux, is a step in the right direction.

*Proof of Lemma 5.9.9 (sketched).* Let us first give a slightly informal but short proof. A more formal but also more intricate argument will be given afterwards (and in fact will prove a more general version of the lemma).

From $w \in S_{n,X\cup Y} \setminus \mathcal{C}(T)$, we obtain $w \in S_{n,X\cup Y}$ and $w \notin \mathcal{C}(T)$. In particular, $w \notin \mathcal{C}(T)$ entails that the $n$-tableau $w \rightharpoonup T$ is not column-equivalent to $T$ (by Proposition 5.3.12 **(b)**), so that $\widetilde{w \rightharpoonup T} \neq \widetilde{T}$. In other words, $\widetilde{wT} \neq \widetilde{T}$ (since $wT = w \rightharpoonup T$).

From $w \in S_{n,X\cup Y}$, we see that the permutation $w$ fixes all elements of $[n]$ that don't belong to $X \cup Y$, while permuting the elements of $X \cup Y$ among themselves.

Thus, the $n$-column-tabloid $\widetilde{wT}$ can be obtained from $\widetilde{T}$ by moving some elements of $X$ from column $j$ to column $k$ and moving some elements of $Y$ from column $k$ to column $j$ (since $X \subseteq \operatorname{Col}(j,T)$ and $Y \subseteq \operatorname{Col}(k,T)$). [72] The # of entries moving east (i.e., from column $j$ to column $k$) must equal the # of entries moving west (i.e., from column $k$ to column $j$) (since the columns retain their sizes), and furthermore must be nonzero (since $\widetilde{wT} \neq \widetilde{T}$, so that at least one entry must change columns from $\widetilde{T}$ to $\widetilde{wT}$). But any entry moving east is an element of $X$, whereas any entry moving west is an element of $Y$. Hence, any entry moving east is larger than any entry moving west (because each element of $X$ is larger than each element of $Y$). Consequently, the largest entry to change columns from $\widetilde{T}$ to $\widetilde{wT}$ must be an entry moving east. In other words, the largest number that appears in different columns in $\widetilde{T}$ and $\widetilde{wT}$ appears further left in $\widetilde{T}$ than in $\widetilde{wT}$. In other words, $\widetilde{T} < \widetilde{wT}$. This proves Lemma 5.9.9. □

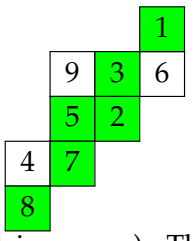
As promised, let me also give a more formal proof of Lemma 5.9.9. In fact, I will prove a more general result:

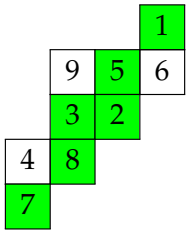
[72] Some elements of $X$ and of $Y$ also move vertically within their columns, but this does not affect the $n$-column-tab**loid**, only the $n$-tableau.

**Lemma 5.9.11.** Let $T$ be an $n$-tableau (of any shape). Let $Z$ be a subset of $[n]$. Assume that $Z$ has the following property:

$$\begin{aligned} &\text{If } a, b \in Z \text{ are two numbers satisfying } c_{\widetilde{T}}(a) < c_{\widetilde{T}}(b), \\ &\text{then } a > b. \end{aligned} \tag{213}$$

Let $w \in S_{n,Z} \setminus \mathcal{C}(T)$. Then, $\widetilde{wT} > \widetilde{T}$ (with respect to the column last letter order).

**Example 5.9.12.** Let $n = 9$ and $T =$ 

and $Z = \{1,2,3,5,7,8\}$. (The elements of $Z$ have been highlighted in green.) Then, $Z$ has the property (213). Thus, Lemma 5.9.11 says that $\widetilde{wT} > \widetilde{T}$ for each $w \in S_{n,Z} \setminus \mathcal{C}(T)$. For example, picking $w = t_{3,5}t_{7,8}$, we find $wT =$ 

and thus easily $\widetilde{wT} > \widetilde{T}$.

*Proof of Lemma 5.9.11.* From $w \in S_{n,Z} \setminus \mathcal{C}(T)$, we obtain $w \in S_{n,Z}$ and $w \notin \mathcal{C}(T)$.

From $w \in S_{n,Z}$, we easily obtain $w^{-1}(Z) \subseteq Z$ [73].

If the $n$-tableau $w \rightharpoonup T$ was column-equivalent to $T$, then we would have $w \in \mathcal{C}(T)$ (by Proposition 5.3.12 **(b)**), which would contradict $w \notin \mathcal{C}(T)$. Hence, $w \rightharpoonup T$ is not column-equivalent to $T$. In other words, $\widetilde{w \rightharpoonup T} \neq \widetilde{T}$ (since an $n$-column-tabloid is an equivalence class with respect to column-equivalence). In other words, $\widetilde{wT} \neq \widetilde{T}$ (since $w \rightharpoonup T = wT$).

We must prove that $\widetilde{wT} > \widetilde{T}$. Assume the contrary. Thus, $\widetilde{wT} \leq \widetilde{T}$ (since the column last letter order is a total order), so that $\widetilde{wT} < \widetilde{T}$ (since $\widetilde{wT} \neq \widetilde{T}$). By the definition of the relation $<$, this means that

- there exists at least one $i \in [n]$ such that $c_{\widetilde{wT}}(i) \neq c_{\widetilde{T}}(i)$, and
- the **largest** such $i$ satisfies $c_{\widetilde{wT}}(i) < c_{\widetilde{T}}(i)$.

[73] *Proof.* Let $k \in w^{-1}(Z)$. Thus, $w(k) \in Z$. We shall show that $k \in Z$.

Indeed, assume the contrary. Thus, $k \notin Z$.

However, $w \in S_{n,Z}$. This means that $w(i) = i$ for all $i \notin Z$ (by the definition of $S_{n,Z}$). Applying this to $i = k$, we obtain $w(k) = k$ (since $k \notin Z$). Hence, $w(k) = k \notin Z$, which contradicts $w(k) \in Z$. This contradiction shows that our assumption was false. Hence, $k \in Z$ is proved.

Forget that we fixed $k$. We thus have proved that $k \in Z$ for each $k \in w^{-1}(Z)$. In other words, $w^{-1}(Z) \subseteq Z$.

Let $k$ be this largest $i$. Thus,

$$c_{\widetilde{wT}}(i) = c_{\widetilde{T}}(i) \qquad \text{for all } i > k \tag{214}$$

(since $k$ is the **largest** $i \in [n]$ such that $c_{\widetilde{wT}}(i) \neq c_{\widetilde{T}}(i)$) and

$$c_{\widetilde{wT}}(k) < c_{\widetilde{T}}(k)$$

(since this largest $i$ satisfies $c_{\widetilde{wT}}(i) < c_{\widetilde{T}}(i)$).

It is easy to see that

$$c_{\widetilde{wT}}(w(m)) = c_{\widetilde{T}}(m) \qquad \text{for each } m \in [n] \tag{215}$$

[74]. Applying this to $m = w^{-1}(k)$, we find $c_{\widetilde{wT}}\left(w\left(w^{-1}(k)\right)\right) = c_{\widetilde{T}}\left(w^{-1}(k)\right)$. Since $w\left(w^{-1}(k)\right) = k$, we can rewrite this as $c_{\widetilde{wT}}(k) = c_{\widetilde{T}}\left(w^{-1}(k)\right)$. Hence, $c_{\widetilde{T}}\left(w^{-1}(k)\right) = c_{\widetilde{wT}}(k) < c_{\widetilde{T}}(k)$. Hence, $c_{\widetilde{T}}\left(w^{-1}(k)\right) \neq c_{\widetilde{T}}(k)$, so that $w^{-1}(k) \neq k$. Hence, $w(k) \neq k$.

Recall that $w \in S_{n,Z}$. This means that the equality $w(i) = i$ holds for all $i \notin Z$ (by the definition of $S_{n,Z}$). If we had $k \notin Z$, then we could apply this equality to $i = k$ and obtain $w(k) = k$, which would contradict $w(k) \neq k$. Hence, we cannot have $k \notin Z$. Thus, $k \in Z$. Therefore, $w^{-1}(k) \in w^{-1}(Z) \subseteq Z$.

Hence, (213) (applied to $a = w^{-1}(k)$ and $b = k$) yields $w^{-1}(k) > k$ (since $c_{\widetilde{T}}\left(w^{-1}(k)\right) < c_{\widetilde{T}}(k)$).

Now we claim the following:

> *Claim 1:* For each positive integer $i$, we have $w^{-i}(k) > k$ and $w^{-i}(k) \in Z$ and $c_{\widetilde{T}}\left(w^{-i}(k)\right) < c_{\widetilde{T}}(k)$.

*Proof of Claim 1.* We induct on $i$:

*Base case:* We have shown that $w^{-1}(k) > k$ and $w^{-1}(k) \in Z$ and $c_{\widetilde{T}}\left(w^{-1}(k)\right) < c_{\widetilde{T}}(k)$. In other words, Claim 1 holds for $i = 1$.

*Induction step:* Let $j > 1$ be an integer. Assume (as the induction hypothesis) that Claim 1 holds for $i = j - 1$. In other words, we have $w^{-(j-1)}(k) > k$ and $w^{-(j-1)}(k) \in Z$ and $c_{\widetilde{T}}\left(w^{-(j-1)}(k)\right) < c_{\widetilde{T}}(k)$. We must show that Claim 1 holds for $i = j$.

Now,

$$w\left(w^{-j}(k)\right) = w^{1+(-j)}(k) = w^{-(j-1)}(k) \qquad (\text{since } 1 + (-j) = -(j-1))$$
$$\in Z,$$

so that

$$w^{-j}(k) \in w^{-1}(Z) \subseteq Z.$$

[74] *Proof:* Let $m \in [n]$. Let $(i, j)$ be the cell of $T$ that contains the entry $m$. Thus, $T(i, j) = m$. Hence, the number $m$ appears in the $j$-th column of $T$ (namely, in the cell $(i, j)$). Hence, $c_{\widetilde{T}}(m) = j$ (by the definition of $c_{\widetilde{T}}(m)$).

However, the definition of $wT$ yields $wT = w \circ T$, so that $(wT)(i, j) = (w \circ T)(i, j) = w\left(\underbrace{T(i, j)}_{=m}\right) = w(m)$. Thus, the number $w(m)$ appears in the $j$-th column of $wT$ (namely, in the cell $(i, j)$). Hence, $c_{\widetilde{wT}}(w(m)) = j$ (by the definition of $c_{\widetilde{wT}}(w(m))$). In other words, $c_{\widetilde{wT}}(w(m)) = c_{\widetilde{T}}(m)$ (since $c_{\widetilde{T}}(m) = j$). This proves (215).

Furthermore, from $w^{-(j-1)}(k) > k$, we obtain

$$c_{\widetilde{w}\widetilde{T}}\left(w^{-(j-1)}(k)\right) = c_{\widetilde{T}}\left(w^{-(j-1)}(k)\right) \qquad \left(\text{by (214), applied to } i = w^{-(j-1)}(k)\right)$$
$$< c_{\widetilde{T}}(k).$$

However, (215) (applied to $m = w^{-j}(k)$) yields

$$c_{\widetilde{w}\widetilde{T}}\left(w\left(w^{-j}(k)\right)\right) = c_{\widetilde{T}}\left(w^{-j}(k)\right).$$

Therefore,

$$c_{\widetilde{T}}\left(w^{-j}(k)\right) = c_{\widetilde{w}\widetilde{T}}\left(\underbrace{w\left(w^{-j}(k)\right)}_{=w^{-(j-1)}(k)}\right) = c_{\widetilde{w}\widetilde{T}}\left(w^{-(j-1)}(k)\right) < c_{\widetilde{T}}(k).$$

Thus, (213) (applied to $a = w^{-j}(k)$ and $b = k$) yields $w^{-j}(k) > k$ (since $w^{-j}(k) \in Z$ and $k \in Z$).

Altogether, we have now proved that $w^{-j}(k) > k$ and $w^{-j}(k) \in Z$ and $c_{\widetilde{T}}\left(w^{-j}(k)\right) < c_{\widetilde{T}}(k)$. In other words, Claim 1 holds for $i = j$. This completes the induction step. Thus, Claim 1 is proven. □

Now, recall that $w$ is an element of the finite group $S_n$, and thus has a finite order (by Lagrange's theorem). In other words, there exists a positive integer $i$ such that $w^i = 1$. Consider this $i$. We find $w^{-i} = \left(\underbrace{w^i}_{=1}\right)^{-1} = 1^{-1} = 1 = \mathrm{id}$ and thus $w^{-i}(k) = k$. But Claim 1 yields $w^{-i}(k) > k$. This contradicts $w^{-i}(k) = k$. This contradiction shows that our assumption was wrong.

Thus, $\widetilde{w}\widetilde{T} > \widetilde{T}$ is proved. This completes the proof of Lemma 5.9.11. □

*Proof of Lemma 5.9.9.* Set $Z = X \cup Y$. Then, the set $Z$ has the property (213)[75]. Moreover,

[75] *Proof.* Let $a, b \in Z$ be two numbers satisfying $c_{\widetilde{T}}(a) < c_{\widetilde{T}}(b)$. We must prove that $a > b$.

We have $a \in Z = X \cup Y$. In other words, we have $a \in X$ or $a \in Y$. If $a \in X$, then the number $a$ lies in the $j$-th column of $T$ (since $a \in X \subseteq \mathrm{Col}(j, T)$) and thus satisfies $c_{\widetilde{T}}(a) = j$. If $a \in Y$, then we instead obtain $c_{\widetilde{T}}(a) = k$ (by a similar argument using $Y \subseteq \mathrm{Col}(k, T)$). Thus, we always have $c_{\widetilde{T}}(a) = j$ or $c_{\widetilde{T}}(a) = k$ (since we have $a \in X$ or $a \in Y$). In other words, $c_{\widetilde{T}}(a) \in \{j, k\}$. Similarly, $c_{\widetilde{T}}(b) \in \{j, k\}$.

From $c_{\widetilde{T}}(b) \in \{j, k\}$, we obtain $c_{\widetilde{T}}(b) \leq \max\{j, k\} = k$ (since $j < k$). From $c_{\widetilde{T}}(a) \in \{j, k\}$, we obtain $c_{\widetilde{T}}(a) \geq \min\{j, k\} = j$ (since $j < k$). However, we assumed that $c_{\widetilde{T}}(a) < c_{\widetilde{T}}(b)$. Thus, $c_{\widetilde{T}}(a) < c_{\widetilde{T}}(b) \leq k$, so that $c_{\widetilde{T}}(a) \neq k$. In other words, the number $a$ does not lie in the $k$-th column of $T$. In other words, $a \notin \mathrm{Col}(k, T)$. Hence, $a \notin Y$ (since $a \in Y$ would imply $a \in Y \subseteq \mathrm{Col}(k, T)$, which would contradict $a \notin \mathrm{Col}(k, T)$). Combining $a \in X \cup Y$ with $a \notin Y$, we obtain $a \in (X \cup Y) \setminus Y \subseteq X$.

Furthermore, from $c_{\widetilde{T}}(a) < c_{\widetilde{T}}(b)$, we obtain $c_{\widetilde{T}}(b) > c_{\widetilde{T}}(a) \geq j$, so that $c_{\widetilde{T}}(b) \neq j$. In other words, the number $b$ does not lie in the $j$-th column of $T$. In other words, $b \notin \mathrm{Col}(j, T)$. Hence, $b \notin X$ (since $b \in X$ would imply $b \in X \subseteq \mathrm{Col}(j, T)$, which would contradict $b \notin \mathrm{Col}(j, T)$). Combining $b \in Z = X \cup Y$ with $b \notin X$, we obtain $b \in (X \cup Y) \setminus X \subseteq Y$.

Now, recall that each element of $X$ is larger than each element of $Y$. In other words, each $x \in X$ and each $y \in Y$ satisfy $x > y$. Applying this to $x = a$ and $y = b$, we obtain $a > b$ (since $a \in X$ and $b \in Y$). Thus, we have shown that $Z$ satisfies (213).

we have $w \in S_{n,X\cup Y} \setminus \mathcal{C}(T) = S_{n,Z} \setminus \mathcal{C}(T)$ (since $X \cup Y = Z$). Therefore, Lemma 5.9.11 yields $\widetilde{wT} > \widetilde{T}$. This proves Lemma 5.9.9. □

### 5.9.4. Proof of the standard basis theorem

Now we can actually use the Garnir relations to prove the standard basis theorem. The main step is the following lemma:

**Lemma 5.9.13.** Let $D$ be a skew Young diagram. Let $T$ be an $n$-tableau of shape $D$ that is column-standard but not row-standard. Then, $\mathbf{e}_T$ can be written as a $\mathbb{Z}$-linear combination of polytabloids $\mathbf{e}_P$, where the subscripts $P$ are $n$-tableaux of shape $D$ satisfying $\widetilde{P} > \widetilde{T}$ (with respect to the column last letter order).

*Proof.* The tableau $T$ is not row-standard. Hence, $T$ has two entries in the same row that are in the wrong order (i.e., the entry further left is $\geq$ to the entry further right). In other words, there exist two entries $T(i,j)$ and $T(i,k)$ of $T$ with $j < k$ and $T(i,j) \geq T(i,k)$. Consider these $i$, $j$ and $k$. Define the two subsets

$$X := \left\{T\left(i',j\right) \mid i' \in \mathbb{Z} \text{ with } i' \geq i \text{ and } \left(i',j\right) \in D\right\} \qquad \text{and}$$
$$Y := \left\{T\left(i',k\right) \mid i' \in \mathbb{Z} \text{ with } i' \leq i \text{ and } \left(i',k\right) \in D\right\}$$

of $[n]$. Thus, $X$ is the set of entries of $T$ in the $j$-th column weakly below the cell $(i,j)$, while $Y$ is the set of entries of $T$ in the $k$-th column weakly above the cell $(i,k)$. In particular, $X \subseteq \mathrm{Col}(j,T)$ and $Y \subseteq \mathrm{Col}(k,T)$.

Note that $j \neq k$ (since $j < k$), and thus $(i,j) \neq (i,k)$. Since $T$ is an injective map (because $T$ is an $n$-tableau), we thus have $T(i,j) \neq T(i,k)$ as well. Combining this with $T(i,j) \geq T(i,k)$, we obtain $T(i,j) > T(i,k)$. Since $T$ is column-standard, we thus readily find that each element of $X$ is larger than each element of $Y$ [76]. Moreover, $D$ is a skew diagram, and thus Lemma 5.8.7 **(a)** entails

$$|X \cup Y| > |D\langle j\rangle \cup D\langle k\rangle| \tag{216}$$

[77]. Let us set $Z = X \cup Y$. Thus, we can rewrite (216) as $|Z| > |D\langle j\rangle \cup D\langle k\rangle|$. Furthermore, we have $Z = \underbrace{X}_{\subseteq \mathrm{Col}(j,T)} \cup \underbrace{Y}_{\subseteq \mathrm{Col}(k,T)} \subseteq \mathrm{Col}(j,T) \cup \mathrm{Col}(k,T)$. There-

[76] *Proof.* Let $x \in X$ and $y \in Y$. We must prove that $x > y$.

We have $x \in X = \{T(i',j) \mid i' \in \mathbb{Z} \text{ with } i' \geq i \text{ and } (i',j) \in D\}$. In other words, $x$ is an entry of $T$ in the $j$-th column weakly below the cell $(i,j)$. But $T$ is column-standard; thus, the entries in the $j$-th column of $T$ increase from top to bottom. Hence, $x \geq T(i,j)$ (since $x$ is an entry of $T$ in the $j$-th column weakly below the cell $(i,j)$).

We have $y \in Y = \{T(i',k) \mid i' \in \mathbb{Z} \text{ with } i' \leq i \text{ and } (i',k) \in D\}$. In other words, $y$ is an entry of $T$ in the $k$-th column weakly above the cell $(i,k)$. But $T$ is column-standard; thus, the entries in the $k$-th column of $T$ increase from top to bottom. Hence, $y \leq T(i,k)$ (since $y$ is an entry of $T$ in the $k$-th column weakly above the cell $(i,k)$). In other words, $T(i,k) \geq y$.

Now, $x \geq T(i,j) > T(i,k) \geq y$, qed.

[77] Even better: $|X \cup Y| = |D\langle j\rangle \cup D\langle k\rangle| + 1$. Do you see why? (But we don't need this.)

fore, the Garnir relations (Theorem 5.8.17) can be applied, once we pick a left transversal $L$ of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$. However, we can easily do this: We can easily see (as in Theorem 5.8.17) that $S_{n,Z} \cap \mathcal{C}(T)$ is a subgroup of $S_{n,Z}$. Thus, Proposition 5.8.11 **(a)** (applied to $G = S_{n,Z}$ and $H = S_{n,Z} \cap \mathcal{C}(T)$) shows that there exists at least one left transversal $L$ of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$. Consider this $L$. Then, (207) yields

$$\mathbf{e}_T = - \sum_{x \in L \setminus \mathcal{C}(T)} (-1)^x \mathbf{e}_{xT}. \tag{217}$$

However, for each $x \in L \setminus \mathcal{C}(T)$, we have $x \in \underbrace{L}_{\substack{\subseteq S_{n,Z} = S_{n,X \cup Y} \\ \text{(since } Z = X \cup Y\text{)}}} \setminus \mathcal{C}(T) \subseteq S_{n,X \cup Y} \setminus \mathcal{C}(T)$ and therefore $\widetilde{xT} > \widetilde{T}$ in the column last letter order (by Lemma 5.9.9 (applied to $w = x$), since each element of $X$ is larger than each element of $Y$). In other words, each $x$ on the right hand side of (217) satisfies $\widetilde{xT} > \widetilde{T}$. Hence, the equality (217) expresses $\mathbf{e}_T$ as a $\mathbb{Z}$-linear combination of polytabloids $\mathbf{e}_P$, where the subscripts $P$ are $n$-tableaux of shape $D$ satisfying $\widetilde{P} > \widetilde{T}$ (with respect to the column last letter order). This proves Lemma 5.9.13. □

**Proposition 5.9.14** (straightening law for standard polytabloids)**.** Let $D$ be a skew Young diagram. Let $T$ be an $n$-tableau of shape $D$. Then, $\mathbf{e}_T$ is a $\mathbb{Z}$-linear combination of standard polytabloids $\mathbf{e}_Q$, where $Q$ ranges over all standard tableaux of shape $D$.

*Proof.* Forget that we fixed $T$. We must show that each polytabloid $\mathbf{e}_T$ is a $\mathbb{Z}$-linear combination of standard polytabloids.

In this proof, we will not consider any diagrams other than $D$. Thus, the word "$n$-tableau" will always mean "$n$-tableau of shape $D$". Likewise for the word "$n$-column-tabloid".

Define a $\mathbb{Z}$-submodule $\overline{\mathcal{S}}^D$ of $\mathcal{S}^D$ by

$$\overline{\mathcal{S}}^D := \operatorname{span}_{\mathbb{Z}} \left\{ \mathbf{e}_Q \mid Q \text{ is a standard } n\text{-tableau} \right\}.$$

Thus, $\overline{\mathcal{S}}^D$ is the $\mathbb{Z}$-linear span of all standard polytabloids. We shall show that

$$\mathbf{e}_T \in \overline{\mathcal{S}}^D \qquad \text{for all } n\text{-tableaux } T. \tag{218}$$

Indeed, let us define the *depth* $\operatorname{dep} \widetilde{T}$ of an $n$-column-tabloid $\widetilde{T}$ to be the # of all $n$-column-tabloids $\widetilde{P}$ that satisfy $\widetilde{P} > \widetilde{T}$ (with respect to the column last letter order). This is clearly a well-defined nonnegative integer (since there are only finitely many $n$-column-tabloids $\widetilde{P}$). Moreover, if two $n$-column-tabloids $\widetilde{S}$ and $\widetilde{T}$ satisfy $\widetilde{S} > \widetilde{T}$ (with respect to the column last letter order), then

$$\operatorname{dep} \widetilde{S} < \operatorname{dep} \widetilde{T} \tag{219}$$

[78].

Now, we shall prove (218) by strong induction on $\operatorname{dep} \widetilde{T}$:

*Induction step:* Fix an $n$-tableau $U$, and assume (as the induction hypothesis) that (218) is already proved for all $n$-tableaux $T$ satisfying $\operatorname{dep} \widetilde{T} < \operatorname{dep} \widetilde{U}$. Our goal is now to prove (218) for $T = U$. In other words, our goal is to prove that $\mathbf{e}_U \in \overline{\mathcal{S}}^D$.

The tableau $U$ may or may not be column-standard. However, it is easy to find a column-standard $n$-tableau $S$ that is column-equivalent to $U$ (just let $S$ be the result of sorting the entries in each column of $U$ into increasing order). Consider this $S$. Then, $\widetilde{S} = \widetilde{U}$ (since $S$ is column-equivalent to $U$) and $\mathbf{e}_U = \pm \mathbf{e}_S$ (by Lemma 5.4.6 **(d)**, applied to $T = U$).

If the $n$-tableau $S$ is standard, then we have $\mathbf{e}_S \in \overline{\mathcal{S}}^D$ (by the definition of $\overline{\mathcal{S}}^D$, since $\mathbf{e}_S$ is a standard polytabloid), and thus $\mathbf{e}_U = \pm \underbrace{\mathbf{e}_S}_{\in \overline{\mathcal{S}}^D} \in \overline{\mathcal{S}}^D$ (since $\overline{\mathcal{S}}^D$ is a $\mathbb{Z}$-module), which immediately completes our induction step (since our goal is to prove that $\mathbf{e}_U \in \overline{\mathcal{S}}^D$).

Hence, for the rest of the induction step, we WLOG assume that $S$ is not standard. Since $S$ is column-standard, we thus conclude that $S$ is not row-standard (since otherwise, $S$ would be both row-standard and column-standard and therefore standard).

Hence, Lemma 5.9.13 (applied to $S$ instead of $T$) shows that $\mathbf{e}_S$ can be written as a $\mathbb{Z}$-linear combination of polytabloids $\mathbf{e}_P$, where the subscripts $P$ are $n$-tableaux of shape $D$ satisfying $\widetilde{P} > \widetilde{S}$ (with respect to the column last letter order). In other words,

$$\mathbf{e}_S \in \operatorname{span}_{\mathbb{Z}} \left\{ \mathbf{e}_P \mid P \text{ is an } n\text{-tableau satisfying } \widetilde{P} > \widetilde{S} \right\}. \tag{220}$$

Now, let $P$ be an $n$-tableau satisfying $\widetilde{P} > \widetilde{S}$. Then, $\widetilde{P} > \widetilde{S} = \widetilde{U}$ and therefore $\operatorname{dep} \widetilde{P} < \operatorname{dep} \widetilde{U}$ (by (219), applied to $P$ and $U$ instead of $S$ and $T$). But

[78] *Proof:* Let $\widetilde{S}$ and $\widetilde{T}$ be two $n$-column-tabloids that satisfy $\widetilde{S} > \widetilde{T}$ (with respect to the column last letter order). Define the two sets

$$X := \left\{ n\text{-column-tabloids } \widetilde{P} \text{ that satisfy } \widetilde{P} > \widetilde{S} \right\} \qquad \text{and}$$
$$Y := \left\{ n\text{-column-tabloids } \widetilde{P} \text{ that satisfy } \widetilde{P} > \widetilde{T} \right\}.$$

Then, $\operatorname{dep} \widetilde{T} = |Y|$ (since $\operatorname{dep} \widetilde{T}$ is defined as the # of all $n$-column-tabloids $\widetilde{P}$ that satisfy $\widetilde{P} > \widetilde{T}$) and $\operatorname{dep} \widetilde{S} = |X|$ (similarly). However, each element of $X$ is an $n$-column-tabloid $\widetilde{P}$ that satisfies $\widetilde{P} > \widetilde{S}$ (by the definition of $X$), and thus also satisfies $\widetilde{P} > \widetilde{T}$ (since $\widetilde{P} > \widetilde{S} > \widetilde{T}$), and therefore belongs to $Y$ as well (by the definition of $Y$). Thus, $X \subseteq Y$. Furthermore, $\widetilde{S}$ belongs to $Y$ (since $\widetilde{S} > \widetilde{T}$) but not to $X$ (since we don't have $\widetilde{S} > \widetilde{S}$). Hence, the sets $X$ and $Y$ differ in the element $\widetilde{S}$. Thus, $X \neq Y$. Combining this with $X \subseteq Y$, we see that $X$ is a proper subset of $Y$. Since the set $Y$ is finite, this entails $|X| < |Y|$. In other words, $\operatorname{dep} \widetilde{S} < \operatorname{dep} \widetilde{T}$ (since $\operatorname{dep} \widetilde{S} = |X|$ and $\operatorname{dep} \widetilde{T} = |Y|$). This proves (219).

our induction hypothesis says that (218) is already proved for all $n$-tableaux $T$ satisfying $\operatorname{dep} \widetilde{T} < \operatorname{dep} \widetilde{U}$. Applying this to $T = P$, we conclude that (218) is already proved for $T = P$ (since $\operatorname{dep} \widetilde{P} < \operatorname{dep} \widetilde{U}$). Hence, $\mathbf{e}_P \in \overline{\mathcal{S}}^D$.

Forget that we fixed $P$. We thus have shown that $\mathbf{e}_P \in \overline{\mathcal{S}}^D$ whenever $P$ is an $n$-tableau satisfying $\widetilde{P} > \widetilde{S}$. In other words,

$$\left\{ \mathbf{e}_P \mid P \text{ is an } n\text{-tableau satisfying } \widetilde{P} > \widetilde{S} \right\} \subseteq \overline{\mathcal{S}}^D.$$

Now, (220) becomes

$$\begin{aligned} \mathbf{e}_S &\in \operatorname{span}_{\mathbb{Z}} \underbrace{\left\{ \mathbf{e}_P \mid P \text{ is an } n\text{-tableau satisfying } \widetilde{P} > \widetilde{S} \right\}}_{\subseteq \overline{\mathcal{S}}^D} \\ &\subseteq \operatorname{span}_{\mathbb{Z}} \left( \overline{\mathcal{S}}^D \right) = \overline{\mathcal{S}}^D \qquad \left( \text{since } \overline{\mathcal{S}}^D \text{ is a } \mathbb{Z}\text{-module} \right). \end{aligned}$$

Hence, $\mathbf{e}_U = \pm \underbrace{\mathbf{e}_S}_{\in \overline{\mathcal{S}}^D} \in \overline{\mathcal{S}}^D$ (since $\overline{\mathcal{S}}^D$ is a $\mathbb{Z}$-module). But this is precisely what we intended to prove. Thus, we have proved (218) for $T = U$. This completes the induction step. Thus, (218) is proved for all $T$.

In other words, we have proved that each polytabloid $\mathbf{e}_T$ belongs to $\overline{\mathcal{S}}^D$. In other words, each polytabloid $\mathbf{e}_T$ belongs to the $\mathbb{Z}$-linear span of all standard polytabloids (since $\overline{\mathcal{S}}^D$ is the $\mathbb{Z}$-linear span of all standard polytabloids). In other words, each polytabloid $\mathbf{e}_T$ is a $\mathbb{Z}$-linear combination of standard polytabloids. This proves Proposition 5.9.14. □

The above proof of Proposition 5.9.14 is fully constructive: It provides a recursive algorithm for expressing any polytabloid as a $\mathbb{Z}$-linear combination of standard polytabloids. This expansion is called *straightening*. Note that the word "straightening" here is an archaism; a better word would be "standardization" (since it makes the tableau standard, not straight-shaped), but this word has a different meaning already.

We can now easily prove the standard basis theorem (Theorem 5.9.1):

*Proof of Theorem 5.9.1.* In this proof, we will not consider any diagrams other than $D$. Thus, the word "polytabloid" will always mean "polytabloid of shape $D$".

The definition of the Specht module $\mathcal{S}^D$ yields

$$\mathcal{S}^D = \operatorname{span}_{\mathbf{k}} \left\{ \text{all polytabloids } \mathbf{e}_T \right\}.$$

However, each polytabloid $\mathbf{e}_T$ is a $\mathbb{Z}$-linear combination of standard polytabloids (by Proposition 5.9.14), and thus is a $\mathbf{k}$-linear combination of standard

polytabloids as well (since any $\mathbb{Z}$-linear combination in a $\mathbf{k}$-module is automatically a $\mathbf{k}$-linear combination). In other words, each polytabloid $\mathbf{e}_T$ belongs to $\operatorname{span}_{\mathbf{k}} \{\text{standard polytabloids}\}$. In other words,

$$\{\text{all polytabloids } \mathbf{e}_T\} \subseteq \operatorname{span}_{\mathbf{k}} \{\text{standard polytabloids}\}.$$

Hence,

$$\begin{aligned} \mathcal{S}^D &= \operatorname{span}_{\mathbf{k}} \underbrace{\{\text{all polytabloids } \mathbf{e}_T\}}_{\subseteq \operatorname{span}_{\mathbf{k}} \{\text{standard polytabloids}\}} \\ &\subseteq \operatorname{span}_{\mathbf{k}} \left(\operatorname{span}_{\mathbf{k}} \{\text{standard polytabloids}\}\right) \\ &= \operatorname{span}_{\mathbf{k}} \{\text{standard polytabloids}\} \end{aligned}$$

(since $\operatorname{span}_{\mathbf{k}} \{\text{standard polytabloids}\}$ is a $\mathbf{k}$-module). In other words, the standard polytabloids span the $\mathbf{k}$-module $\mathcal{S}^D$. Since these standard polytabloids are furthermore $\mathbf{k}$-linearly independent (by Theorem 5.7.4), we thus conclude that they form a basis of $\mathcal{S}^D$. Thus, Theorem 5.9.1 is proved at last. □

The specific coefficients that appear when a given polytabloid $\mathbf{e}_T$ is expanded in the standard polytabloid basis are not easily described. (A recent paper by Hodges [Hodges23] describes them in the case of a straight Young diagram $D = Y(\lambda)$, but even that description is still essentially recursive, as it involves a sum over a set of sequences.[79] There are various other algorithms for computing these coefficients – such as the "turbo-straightening" of [CaLaLe92] and [Krob95, Exercice 2.2.19] – that compute them faster than a naive approach using Garnir relations, but are still a far cry from an explicit formula.)

**Remark 5.9.15.** Proposition 5.9.14 can be slightly strengthened:

Let $D$ be a skew Young diagram. Let $T$ be an $n$-tableau of shape $D$. Then, $\mathbf{e}_T$ is a $\mathbb{Z}$-linear combination of standard polytabloids $\mathbf{e}_Q$, where $Q$ ranges over all standard tableaux of shape $D$ that satisfy $\widetilde{Q} \geq \widetilde{T}$ in the column last letter order.

This can be proved using the same induction argument that we used to prove Proposition 5.9.14, with just a little more bookkeeping.

**Exercise 5.9.1.** [2] Let $D$ be a skew Young diagram. Prove that the Specht module $\mathcal{S}^D$ is a direct addend of $\mathcal{M}^D$ as a $\mathbf{k}$-module (although not necessarily as a representation of $S_n$).

[79] Hodges works in the somewhat more general setting of $\mathrm{GL}(m)$-representations and semistandard tableaux instead of $S_n$-representations and standard tableaux; but we can recover our setting by restricting ourselves to the "multilinear component", i.e., to the span of the $n$-tabloids coming from standard tabloids.

### 5.9.5. Remarks on bad shapes

What goes wrong if $D$ is a bad shape (i.e., not a skew Young diagram)? Theorem 5.9.1 does not hold in this case, as we already found in Remark 5.7.11; but let us see what goes wrong if we try to straighten a polytabloid using the Garnir relations:

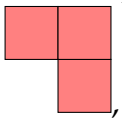
**Example 5.9.16.** Let $n$ and $D$ be as in Remark 5.7.11. Then, the $n$-tableau $\begin{array}{|c|c|}\hline 2 & 1 \\ \hline \multicolumn{1}{c}{} & \multicolumn{1}{c}{3} \\ \cline{2-2}\end{array} = 21\backslash\backslash \circ 3$ of shape $D$ is not standard. Can we simplify the polytabloid $\mathbf{e}_T$ using a Garnir relation? The only Garnir relation that can be applied is one that involves all three entries $1,2,3$ (that is, Theorem 5.8.17 for $j = 1$ and $k = 2$ and $Z = \{1,2,3\}$), since otherwise the inequality $|Z| > |D\langle j\rangle \cup D\langle k\rangle|$ would not be satisfied. This Garnir relation (specifically, (207)) gives

$$\mathbf{e}_{21\backslash\backslash\circ 3} = \mathbf{e}_{12\backslash\backslash\circ 3} + \mathbf{e}_{31\backslash\backslash\circ 2}.$$

But the tableau $31\backslash\backslash \circ 2$ is still not standard, and applying a further Garnir relation to this tableau yields

$$\mathbf{e}_{31\backslash\backslash\circ 2} = \mathbf{e}_{13\backslash\backslash\circ 2} + \mathbf{e}_{21\backslash\backslash\circ 3}.$$

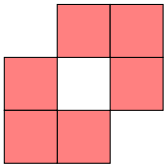
Note that the $\mathbf{e}_{21\backslash\backslash\circ 3}$ reappears again here. Thus, we have expressed $\mathbf{e}_{21\backslash\backslash\circ 3}$ through $\mathbf{e}_{31\backslash\backslash\circ 2}$, but then expressed $\mathbf{e}_{31\backslash\backslash\circ 2}$ back through $\mathbf{e}_{21\backslash\backslash\circ 3}$. In other words, we are running in circles. Hence, we cannot use Garnir relations to express $\mathbf{e}_{21\backslash\backslash\circ 3}$ entirely though standard polytabloids.

Of course, this example does not doom the Specht module $\mathcal{S}^D$ to be an inscrutable mystery. Indeed, when $D =$  , we can find a basis of $\mathcal{S}^D$ by reflecting $D$ vertically (or horizontally); the result will be a Young diagram (or a skew Young diagram), to which the standard basis theorem (Theorem 5.9.1) applies. (The reflection does not change $\mathcal{S}^D$, since $\mathcal{S}^D$ is defined in terms of all $n$-tableaux, without any standardness requirement.)

More generally, if you permute the rows or the columns of a diagram $D$, then its Young module $\mathcal{M}^D$ and its Specht module $\mathcal{S}^D$ remain unchanged (up to isomorphism). Thus, if a bad shape $D$ can be transformed into a skew Young diagram using such permutations, then we can still find a basis for $\mathcal{S}^D$. Many bad shapes (in particular, all diagrams $D$ with at most 5 cells) can be handled in this way.

However, there are some bad shapes $D$ that **cannot** be turned into skew Young diagrams using such permutations. For example:

**Exercise 5.9.2.** [2] Prove that

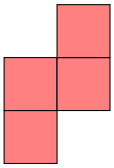


$= \left\{(i,j) \in [3]^2 \mid i \neq j\right\}$

is such a diagram.

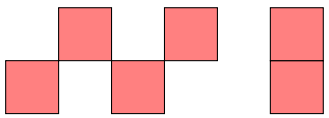
There have been works by James, Peel, Taylor, Reiner, Shimozono, Liu and others ([JamPee79, Theorem 2.4], [ReiShi95a], [ReiShi95b], [Taylor01], [Liu10], [Liu16], [BilPaw13, §3], etc.) dealing with these kinds of bad-shape Specht modules, but the theory is still mysterious and intricate and we don't know a basis in general. For instance, explicit matrix computations using SageMath show that when $D$ is the diagram from Exercise 5.9.2, the Specht module $\mathcal{S}^D$ is a free $\mathbf{k}$-module of rank 42 (no matter what $\mathbf{k}$ is)[80], but this brute-force approach quickly surpasses the abilities of modern computers as $D$ becomes larger. Thus, Question 5.7.1 remains unanswered in general. Some diagrams $D$ lend themselves to easy answers:

**Exercise 5.9.3.** **(a)** [1] Let

$D = \{(1,2),\ (2,1),\ (2,2),\ (3,1)\} =$  .

Prove that $\mathcal{S}^D$ is a free $\mathbf{k}$-module, and find its rank.

**(b)** [3] Let

$D = \{(1,2),\ (1,4),\ (1,6),\ (2,1),\ (2,3),\ (2,6)\} =$  .

Prove that $\mathcal{S}^D$ is a free $\mathbf{k}$-module, and find its rank.

**Exercise 5.9.4.** Let $D$ be a diagram. The following is a brute-force method for determining whether $\mathcal{S}^D$ is "well-behaved".

Let $k$ be the # of $n$-tabloids of shape $D$. Let $M \in \mathbb{Z}^{n! \times k}$ be the $n! \times k$-matrix whose rows are indexed by the $n$-tableaux $T$ of shape $D$, and whose columns are indexed by the $n$-tabloids $\overline{S}$ of shape $D$, and whose $\left(T, \overline{S}\right)$-th entry is the

[80] The method to compute this is explained in Exercise 5.9.4 below.

coefficient of $\overline{S}$ in the polytabloid $\mathbf{e}_T \in \mathcal{M}^D$ (expanded in the basis consisting of all $n$-tabloids). Note that all entries of $M$ are 0's, 1's and $-1$'s.

Let $M^{\mathrm{sm}}$ be the Smith normal form of the matrix $M$. Prove the following:

**(a)** $\boxed{1}$ If all diagonal entries of $M^{\mathrm{sm}}$ are 0's and 1's, then $\mathcal{S}^D$ is a direct addend of $\mathcal{M}^D$ as a $\mathbf{k}$-module for any $\mathbf{k}$.

**(b)** $\boxed{1}$ If all diagonal entries of $M^{\mathrm{sm}}$ are 0's and 1's, then $\mathcal{S}^D$ is a free $\mathbf{k}$-module of rank $r$ for any $\mathbf{k}$, where $r$ is the # of 1's on the diagonal of $M^{\mathrm{sm}}$.

**(c)** $\boxed{1}$ If **not** all diagonal entries of $M^{\mathrm{sm}}$ are 0's and 1's, then there exists a prime number $p$ such that the dimension of $\mathcal{S}^D$ for $\mathbf{k} = \mathbb{F}_p$ is smaller than the dimension of $\mathcal{S}^D$ for $\mathbf{k} = \mathbb{Q}$. (Namely, $p$ can be taken to be a prime divisor of some nonzero diagonal entry of $M^{\mathrm{sm}}$.)

**(d)** $\boxed{3}$ If **not** all diagonal entries of $M^{\mathrm{sm}}$ are 0's and 1's, then there exists a prime power $p^\ell$ such that $\mathcal{S}^D$ is not a free $\mathbf{k}$-module for $\mathbf{k} = \mathbb{Z}/p^\ell$.

**(e)** $\boxed{1}$ All the above claims remain true if we remove from the matrix $M$ all its rows except for those indexed by column-standard $n$-tableaux $T$. (This produces a much smaller matrix.)

[**Hint:** See [Elman22, §43–§44] for the Smith normal form and its uses.]

### 5.9.6. The rank of $\mathcal{S}^\lambda$

For future use, let us record a simple consequence of the standard basis theorem:

**Lemma 5.9.17.** Let $\lambda$ be a partition of $n$. Set $\mathcal{S}^\lambda := \mathcal{S}^{Y(\lambda)}$. Let $f^\lambda$ denote the # of standard tableaux of shape $Y(\lambda)$. Then, the Specht module $\mathcal{S}^\lambda$ is a free $\mathbf{k}$-module of rank $f^\lambda$.

*Proof.* The Young diagram $Y(\lambda)$ is a skew Young diagram (since $Y(\lambda) = Y(\lambda/\varnothing)$). Thus, the standard basis theorem (Theorem 5.9.1, applied to $D = Y(\lambda)$) shows that the standard polytabloids $\mathbf{e}_T$ (that is, the $\mathbf{e}_T$ where $T$ ranges over all standard tableaux of shape $Y(\lambda)$) form a basis of the $\mathbf{k}$-module $\mathcal{S}^{Y(\lambda)}$. Hence, $\mathcal{S}^{Y(\lambda)}$ is a free $\mathbf{k}$-module, and its rank is equal to the # of all standard tableaux of shape $Y(\lambda)$. Thus, its rank is $f^\lambda$ (since $f^\lambda$ is the # of standard tableaux of shape $Y(\lambda)$).

Thus we have shown that $\mathcal{S}^{Y(\lambda)}$ is a free $\mathbf{k}$-module of rank $f^\lambda$. In other words, $\mathcal{S}^\lambda$ is a free $\mathbf{k}$-module of rank $f^\lambda$ (since $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)}$). This proves Lemma 5.9.17. $\square$

## 5.10. The Young alternative

Let us next study the case when $D$ is a straight Young diagram $Y(\lambda)$ of a partition $\lambda$. In this case, the Specht module $\mathcal{S}^{Y(\lambda)}$ is also known as $\mathcal{S}^{\lambda}$, and such straight-shaped Specht modules $\mathcal{S}^{\lambda}$ play an important role: They are the irreducible representations of $S_n$ when $\mathbf{k}$ is a field of characteristic 0. The proof of this fact (Theorem 5.13.1) will take us some preparation.

### 5.10.1. A bit of combinatorics: the Young alternative

First, we need to lay some combinatorial groundwork. We begin with some properties of straight-shaped Young tableaux.

**Theorem 5.10.1** (Young alternative)**.** Let $\lambda$ and $\mu$ be two partitions of $n$.

Let $S$ be an $n$-tableau of shape $Y(\lambda)$. Let $T$ be an $n$-tableau of shape $Y(\mu)$.

Assume that there are no two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$. Then:

**(a)** There exists a permutation $c \in \mathcal{C}(T)$ such that each $i \geq 1$ satisfies the following property: Each entry of the $i$-th row of $S$ lies in one of the rows $1, 2, \ldots, i$ of $cT$.

**(b)** We have $\lambda_1 + \lambda_2 + \cdots + \lambda_i \leq \mu_1 + \mu_2 + \cdots + \mu_i$ for each $i \geq 1$.

**(c)** Now assume that $\lambda = \mu$. Then, there exist permutations $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$.

**Example 5.10.2. (a)** Let $n = 8$ and $\lambda = (3,2,1,1,1)$ and $\mu = (4,3,1)$ and

$$S = \begin{array}{|c|c|c|}\hline 6 & 1 & 5 \\ \hline 4 & 2 \\ \cline{1-2} 8 \\ \cline{1-1} 3 \\ \cline{1-1} 7 \\ \cline{1-1}\end{array} \qquad \text{and} \qquad T = \begin{array}{|c|c|c|c|}\hline 3 & 7 & 8 & 6 \\ \hline 1 & 5 & 2 \\ \cline{1-3} 4 \\ \cline{1-1}\end{array} .$$

Then, it is easy to check that there are no two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$. (For example, the entries $6, 1, 5$ of the first row of $S$ lie in distinct columns of $T$, and so do the entries 4 and 2 of the second row of $S$.) Thus, Theorem 5.10.1 **(a)** yields that there exists a permutation $c \in \mathcal{C}(T)$ such that each $i \geq 1$ satisfies the following property: Each entry of the $i$-th row of $S$ lies in one of the rows $1, 2, \ldots, i$ of $cT$. What is this permutation $c$ ? For example, we can take $c = \operatorname{cyc}_{4,3,1} t_{5,7}$, which produces the $n$-tableau $cT = \begin{array}{|c|c|c|c|}\hline 1 & 5 & 8 & 6 \\ \hline 4 & 7 & 2 \\ \cline{1-3} 3 \\ \cline{1-1}\end{array}$ .

Another valid choice is $c = \operatorname{cyc}_{4,3,1} t_{5,7} t_{2,8}$.

**(b)** Let $n = 6$ and $\lambda = (3,2,1)$ and $\mu = (2,2,2)$ and

$$S = \begin{array}{|c|c|c|}\hline 3 & 1 & 5 \\ \hline 4 & 2 \\ \cline{1-2} 6 \\ \cline{1-1}\end{array} \qquad \text{and} \qquad T = \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 5 \\ \hline 6 & 3 \\ \hline\end{array} .$$

Then, the claim of Theorem 5.10.1 **(b)** is not satisfied, since for $i = 2$ we have $\lambda_1 + \lambda_2 = 3 + 2 > 2 + 2 = \mu_1 + \mu_2$. Thus, the contrapositive of Theorem 5.10.1 **(b)** yields that there are two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$. And indeed, 3 and 5 are two such integers.

**(c)** Let $n = 6$ and $\lambda = (3,2,1)$ and $\mu = (3,2,1)$ and

$$S = \begin{array}{|c|c|c|}\hline 1 & 6 & 2 \\ \hline 4 & 3 \\ \cline{1-2} 5 \\ \cline{1-1}\end{array} \qquad \text{and} \qquad T = \begin{array}{|c|c|c|}\hline 5 & 3 & 6 \\ \hline 4 & 2 \\ \cline{1-2} 1 \\ \cline{1-1}\end{array} .$$

Then, $\lambda = \mu$. Moreover, it is again easy to check that there are no two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$. Hence, Theorem 5.10.1 **(c)** yields that there exist permutations $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$. And indeed, $r = t_{2,6}$ and $c = t_{1,5}t_{2,3}$ are such permutations, since they satisfy $rS = \begin{array}{|c|c|c|}\hline 1 & 2 & 6 \\ \hline 4 & 3 \\ \cline{1-2} 5 \\ \cline{1-1}\end{array} = cT$.

We will often use Theorem 5.10.1 in the equivalent form "**either** there are two distinct integers that lie in the same row of $S$ and in the same column of $T$, **or** the claims of Theorem 5.10.1 **(a)**, **(b)** and **(c)** hold". (This is a non-exclusive "or", at least as far as parts **(a)** and **(b)** are concerned.) Thus I call it the *Young alternative*.

*Proof of Theorem 5.10.1.* In Definition 5.7.5 **(a)**, we introduced the notation $r_{\overline{T}}(i)$ for the number of the row in which a given $n$-tableau $T$ contains a given number $i \in [n]$. This notation will again be useful in our current proof. (Of course, we will use it not only for our specific $n$-tableau $T$.)

Furthermore, if $i \in [n]$ is arbitrary, then the integer $r_{\overline{S}}(i)$ (which is simply the number of the row in which $S$ contains $i$) will be called the *$S$-depth* of $i$. For instance, if $S$ is as in Example 5.10.2 **(a)**, then $r_{\overline{S}}(4) = 2$ and $r_{\overline{S}}(7) = 5$. We make the following trivial observation:

*Claim 0:* For each $j \in [n]$, we have $r_{\overline{S}}(j) \geq 1$.

*Proof of Claim 0.* The tableau $S$ has shape $Y(\lambda)$, which is a subset of $\{1,2,3,\ldots\}^2$. Thus, each cell of $S$ belongs to one of the rows $1,2,3,\ldots$. Therefore, any of the numbers $1,2,\ldots,n$ appears in $S$ in one of the rows $1,2,3,\ldots$. In other words, $r_{\overline{S}}(j) \in \{1,2,3,\ldots\}$ for each $j \in [n]$ (since $r_{\overline{S}}(j)$ is defined to be the number of the row of $S$ that contains the entry $j$). Hence, for each $j \in [n]$, we have $r_{\overline{S}}(j) \in \{1,2,3,\ldots\}$ and therefore $r_{\overline{S}}(j) \geq 1$. This proves Claim 0. □

Let us now permute the entries of the $n$-tableau $T$ as follows: For each column of $T$, we sort the entries of this column in the order of increasing $S$-depth (so that entries with larger $S$-depth are moved south of entries with smaller $S$-depth that lie in the same column). The resulting $n$-tableau shall be called $R$. Thus, $R$ is an $n$-tableau of shape $Y(\mu)$ (since $T$ is an $n$-tableau of shape $Y(\mu)$), and is column-equivalent to $T$ (since it is obtained from $T$ by sorting the entries in each column).[81] Moreover, the very construction of $R$ shows that the entries in each column of $R$ are sorted in the order of increasing $S$-depth (from top to bottom). In other words, if $u$ and $v$ are two entries that lie in the same column of $R$, with $u$ lying further north than $v$, then

$$r_{\overline{S}}(u) \leq r_{\overline{S}}(v). \tag{221}$$

Now we claim the following:

*Claim 1:* For each $k \in [n]$, we have $r_{\overline{R}}(k) \leq r_{\overline{S}}(k)$.

*Proof of Claim 1.* Let $k \in [n]$. Let $(i,j)$ be the cell of $R$ that contains the entry $k$. Thus, $(i,j) \in Y(\mu)$ and $R(i,j) = k$, so that $r_{\overline{R}}(k) = i$ (since $R(i,j) = k$ shows that the number $k$ is contained in the $i$-th row of $R$).

We have $(i,j) \in Y(\mu)$. Hence, it can be easily shown that all the $i$ cells $(1,j)$, $(2,j)$, $\ldots$, $(i,j)$ (that is, all the $i$ cells $(\ell,j)$ for $\ell \in [i]$) belong to $Y(\mu)$ as well[82]. Hence, their entries $R(1,j)$, $R(2,j)$, $\ldots$, $R(i,j)$ are well-defined (since $R$ is an $n$-tableau of shape $Y(\mu)$). In other words, for each $\ell \in [i]$, the entry $R(\ell,j)$ is well-defined and belongs to $[n]$. Hence, let us set

$$r_\ell := r_{\overline{S}}(R(\ell,j)) \qquad \text{for each } \ell \in [i].$$

[81] For instance, if $S$ and $T$ are as in Example 5.10.2 **(a)**, then

$$R = \begin{array}{|c|c|c|c|}\hline 1 & 5 & 2 & 6 \\ \hline 4 & 7 & 8 \\ \cline{1-3} 3 \\ \cline{1-1}\end{array}\ .$$

[82] *Proof.* Let $\ell \in [i]$. Then, $\ell \leq i$. The cell $(\ell,j)$ belongs to $\{1,2,3,\ldots\}^2$ and lies weakly northwest of the cell $(i,j)$ (since $\ell \leq i$ and $j \leq j$). Therefore, Proposition 5.1.5 (applied to $\mu$, $(\ell,j)$ and $(i,j)$ instead of $\lambda$, $c$ and $d$) shows that $(\ell,j) \in Y(\mu)$ (since $(i,j) \in Y(\mu)$).

Forget that we fixed $\ell$. We thus have shown that $(\ell,j) \in Y(\mu)$ for each $\ell \in [i]$. In other words, all the $i$ cells $(1,j)$, $(2,j)$, $\ldots$, $(i,j)$ belong to $Y(\mu)$.

In particular, $r_i = r_{\overline{S}}(R(i,j)) = r_{\overline{S}}(k)$ (since $R(i,j) = k$). Thus, $r_{\overline{S}}(k) = r_i$.

Now, we shall show that $r_\ell - r_{\ell-1} \geq 1$ for each $\ell \in \{2,3,\ldots,i\}$.

Indeed, let $\ell \in \{2,3,\ldots,i\}$. Thus, $r_\ell = r_{\overline{S}}(R(\ell,j))$ and $r_{\ell-1} = r_{\overline{S}}(R(\ell-1,j))$ (by the definition of $r_{\ell-1}$). But $R(\ell-1,j)$ and $R(\ell,j)$ are two entries that lie in the same column of $R$ (namely, in the $j$-th column), with $R(\ell-1,j)$ lying further north than $R(\ell,j)$ (since the cell $(\ell-1,j)$ lies further north than $(\ell,j)$). Hence, (221) (applied to $u = R(\ell-1,j)$ and $v = R(\ell,j)$) shows that $r_{\overline{S}}(R(\ell-1,j)) \leq r_{\overline{S}}(R(\ell,j))$. In other words, $r_{\ell-1} \leq r_\ell$ (since $r_\ell = r_{\overline{S}}(R(\ell,j))$ and $r_{\ell-1} = r_{\overline{S}}(R(\ell-1,j))$).

But $r_{\overline{S}}(R(\ell,j))$ is defined to be the number of the row of $S$ that contains the entry $R(\ell,j)$. Hence, the number $R(\ell,j)$ lies in the $r_{\overline{S}}(R(\ell,j))$-th row of $S$. In other words, the number $R(\ell,j)$ lies in the $r_\ell$-th row of $S$ (since $r_\ell = r_{\overline{S}}(R(\ell,j))$). The same argument (applied to $\ell-1$ instead of $\ell$) shows that the number $R(\ell-1,j)$ lies in the $r_{\ell-1}$-th row of $S$.

Moreover, $R$ is an $n$-tableau, and thus injective. In other words, all entries of $R$ are distinct. Hence, $R(\ell-1,j) \neq R(\ell,j)$ (since $(\ell-1,j) \neq (\ell,j)$). Thus, $R(\ell-1,j)$ and $R(\ell,j)$ are two distinct integers.

As we know, the tableau $R$ is column-equivalent to $T$. Thus, the $j$-th column of $R$ contains the same entries as the $j$-th column of $T$.

The two integers $R(\ell-1,j)$ and $R(\ell,j)$ both lie in the $j$-th column of $R$ (obviously), and thus lie in the $j$-th column of $T$ (since the $j$-th column of $R$ contains the same entries as the $j$-th column of $T$). Thus, these two integers $R(\ell-1,j)$ and $R(\ell,j)$ lie in the same column of $T$.

Now, recall that there are no two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$. Hence, if two distinct integers lie in the same column of $T$, then they cannot lie in the same row of $S$. Applying this to the two integers $R(\ell-1,j)$ and $R(\ell,j)$ (which are distinct and lie in the same column of $T$), we thus conclude that these two integers $R(\ell-1,j)$ and $R(\ell,j)$ cannot lie in the same row of $S$. In other words, $r_{\ell-1} \neq r_\ell$ (since $R(\ell-1,j)$ lies in the $r_{\ell-1}$-th row of $S$, and since $R(\ell,j)$ lies in the $r_\ell$-th row of $S$).

Combining this with $r_{\ell-1} \leq r_\ell$, we obtain $r_{\ell-1} < r_\ell$. Hence, $r_{\ell-1} \leq r_\ell - 1$ (since $r_{\ell-1}$ and $r_\ell$ are integers). Thus, $r_\ell - r_{\ell-1} \geq 1$.

Now, forget that we fixed $\ell$. We thus have proved that $r_\ell - r_{\ell-1} \geq 1$ for each $\ell \in \{2,3,\ldots,i\}$. Adding these inequalities for all $\ell \in \{2,3,\ldots,i\}$ together, we obtain $\sum_{\ell=2}^{i} (r_\ell - r_{\ell-1}) \geq \sum_{\ell=2}^{i} 1 = i-1$. Thus,

$$i-1 \leq \sum_{\ell=2}^{i} (r_\ell - r_{\ell-1}) = r_i - r_1 \qquad \text{(by the telescope principle)}.$$

Next, recall that $r_1 = r_{\overline{S}}(R(1,j))$ (by the definition of $r_1$). But Claim 0 (applied to $R(1,j)$ instead of $j$) shows that $r_{\overline{S}}(R(1,j)) \geq 1$. In other words, $r_1 \geq 1$

(since $r_1 = r_{\overline{S}}(R(1,j))$). Now,

$$i - 1 \leq r_i - \underbrace{r_1}_{\geq 1} \leq r_i - 1,$$

so that $i \leq r_i$. In other words, $r_{\overline{R}}(k) \leq r_{\overline{S}}(k)$ (since $r_{\overline{R}}(k) = i$ and $r_{\overline{S}}(k) = r_i$). This proves Claim 1. □

Claim 1 easily yields the following:

> *Claim 2:* Let $i \geq 1$. Then, each entry of the $i$-th row of $S$ lies in one of the rows $1, 2, \ldots, i$ of $R$.

*Proof of Claim 2.* Let $k$ be an entry of the $i$-th row of $S$. We must prove that $k$ lies in one of the rows $1, 2, \ldots, i$ of $R$.

Indeed, the number $k$ lies in the $i$-th row of $S$ (by its definition). In other words, $r_{\overline{S}}(k) = i$ (by the definition of $r_{\overline{S}}(k)$).

Claim 0 (applied to $j = k$) yields $r_{\overline{S}}(k) \geq 1$. The same argument (but applied to $\mu$ and $R$ instead of $\lambda$ and $S$) yields $r_{\overline{R}}(k) \geq 1$.

However, Claim 1 yields $r_{\overline{R}}(k) \leq r_{\overline{S}}(k) = i$. Combining this with $r_{\overline{R}}(k) \geq 1$, we obtain $r_{\overline{R}}(k) \in \{1, 2, \ldots, i\}$ (since $r_{\overline{R}}(k)$ is an integer). In other words, $k$ lies in one of the rows $1, 2, \ldots, i$ of $R$ (since $r_{\overline{R}}(k)$ is the number of the row of $R$ in which $k$ lies). This completes our proof of Claim 2. □

Now, recall that the tableau $R$ is column-equivalent to $T$. In other words, the $n$-tableaux $T$ and $R$ are column-equivalent. Hence, Proposition 5.3.14 **(b)** (applied to $Y(\mu)$ and $R$ instead of $D$ and $S$) shows that there exists some $w \in \mathcal{C}(T)$ such that $R = w \rightharpoonup T$. Consider this $w$. Thus, $R = w \rightharpoonup T = wT$.

Claim 2 shows that that each $i \geq 1$ has the property that each entry of the $i$-th row of $S$ lies in one of the rows $1, 2, \ldots, i$ of $R$. In other words, each $i \geq 1$ has the property that each entry of the $i$-th row of $S$ lies in one of the rows $1, 2, \ldots, i$ of $wT$ (since $R = wT$). Thus, there exists a permutation $c \in \mathcal{C}(T)$ such that each $i \geq 1$ has the property that each entry of the $i$-th row of $S$ lies in one of the rows $1, 2, \ldots, i$ of $cT$ (namely, $c = w$). This proves Theorem 5.10.1 **(a)**.

**(b)** Let $i \geq 1$. Recall that $S$ is an $n$-tableau, thus injective. Hence, all entries of $S$ are distinct. Moreover, $S$ has shape $Y(\lambda)$, so that the $k$-th row of $S$ has $\lambda_k$ cells for each $k \geq 1$. Hence, rows $1, 2, \ldots, i$ of $S$ have $\lambda_1 + \lambda_2 + \cdots + \lambda_i$ cells in total. Obviously, these cells contain altogether $\lambda_1 + \lambda_2 + \cdots + \lambda_i$ many entries, which are all distinct (since all entries of $S$ are distinct). Thus, we have shown that rows $1, 2, \ldots, i$ of $S$ have altogether $\lambda_1 + \lambda_2 + \cdots + \lambda_i$ many distinct entries. In other words,

$$|\{\text{entries of rows } 1, 2, \ldots, i \text{ of } S\}| = \lambda_1 + \lambda_2 + \cdots + \lambda_i. \tag{222}$$

The same argument (applied to $R$ and $\mu$ instead of $S$ and $\lambda$) shows that

$$|\{\text{entries of rows } 1, 2, \ldots, i \text{ of } R\}| = \mu_1 + \mu_2 + \cdots + \mu_i \tag{223}$$

(since $R$ is an $n$-tableau of shape $Y(\mu)$).

However, Claim 1 easily shows that

$$\{\text{entries of rows } 1,2,\ldots,i \text{ of } S\} \subseteq \{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}$$

[83]. Hence,

$$|\{\text{entries of rows } 1,2,\ldots,i \text{ of } S\}| \leq |\{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}|.$$

In view of (222) and (223), we can rewrite this as

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i \leq \mu_1 + \mu_2 + \cdots + \mu_i.$$

So Theorem 5.10.1 **(b)** is proved.

**(c)** We know that $S$ is an $n$-tableau of shape $Y(\lambda)$, thus an $n$-tableau of shape $Y(\mu)$ (since $\lambda = \mu$). Thus, both $S$ and $R$ are $n$-tableaux of shape $Y(\mu)$.

We shall prove the following stronger version of Claim 1:

*Claim 3:* For each $k \in [n]$, we have $r_{\overline{R}}(k) = r_{\overline{S}}(k)$.

*Proof of Claim 3.* Let $k \in [n]$. We must prove that $r_{\overline{R}}(k) = r_{\overline{S}}(k)$.

Assume the contrary. Thus, $r_{\overline{R}}(k) \neq r_{\overline{S}}(k)$. But Claim 1 yields $r_{\overline{R}}(k) \leq r_{\overline{S}}(k)$. Hence, $r_{\overline{R}}(k) < r_{\overline{S}}(k)$ (since $r_{\overline{R}}(k) \neq r_{\overline{S}}(k)$).

Set $i := r_{\overline{R}}(k)$. Thus, $i = r_{\overline{R}}(k) < r_{\overline{S}}(k)$, so that $r_{\overline{S}}(k) > i$. Also, obviously, $i = r_{\overline{R}}(k) \geq 1$ (this is proved just as in the proof of Claim 2 above).

But $k$ is an entry of the $r_{\overline{R}}(k)$-th row of $R$ (by the definition of $r_{\overline{R}}(k)$). In other words, $k$ is an entry of the $i$-th row of $R$ (since $i = r_{\overline{R}}(k)$). Hence, $k$ is an entry of one of rows $1,2,\ldots,i$ of $R$ (namely, of the $i$-th row). In other words,

$$k \in \{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}. \tag{224}$$

On the other hand, if we had $k \in \{\text{entries of rows } 1,2,\ldots,i \text{ of } S\}$, then $k$ would be an entry of one of rows $1,2,\ldots,i$ of $S$; but this would mean that $r_{\overline{S}}(k) \in \{1,2,\ldots,i\}$ (since $r_{\overline{S}}(k)$ is defined as the number of the row of $S$ that contains $k$), which would contradict $r_{\overline{S}}(k) > i$. Hence, we cannot have $k \in \{\text{entries of rows } 1,2,\ldots,i \text{ of } S\}$. Thus, we have

$$k \notin \{\text{entries of rows } 1,2,\ldots,i \text{ of } S\}. \tag{225}$$

[83] *Proof.* Let $k \in \{\text{entries of rows } 1,2,\ldots,i \text{ of } S\}$. Thus, $k$ is an entry of one of the rows $1,2,\ldots,i$ of $S$. Thus, the number of the row of $S$ that contains $k$ is $\in \{1,2,\ldots,i\}$. In other words, $r_{\overline{S}}(k) \in \{1,2,\ldots,i\}$ (since $r_{\overline{S}}(k)$ is defined as the number of the row of $S$ that contains $k$). But Claim 1 yields $r_{\overline{R}}(k) \leq r_{\overline{S}}(k) \leq i$ (since $r_{\overline{S}}(k) \in \{1,2,\ldots,i\}$). Since we also have $r_{\overline{R}}(k) \geq 1$ (this is proved just like in the proof of Claim 2 above), we thus obtain $r_{\overline{R}}(k) \in \{1,2,\ldots,i\}$. In other words, the number of the row of $R$ that contains $k$ is $\in \{1,2,\ldots,i\}$ (since $r_{\overline{R}}(k)$ is defined as the number of the row of $R$ that contains $k$). In other words, $k$ is an entry of one of the rows $1,2,\ldots,i$ of $R$. In other words, $k \in \{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}$.

Forget that we fixed $k$. We thus have shown that $k \in \{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}$ for each $k \in \{\text{entries of rows } 1,2,\ldots,i \text{ of } S\}$. In other words, $\{\text{entries of rows } 1,2,\ldots,i \text{ of } S\} \subseteq \{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}$.

From (224) and (225), we see that the sets $\{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}$ and $\{\text{entries of rows } 1,2,\ldots,i \text{ of } S\}$ differ (indeed, the first one contains $k$, whereas the second does not). However, in our above proof of Theorem 5.10.1 **(b)**, we have shown that

$$\{\text{entries of rows } 1,2,\ldots,i \text{ of } S\} \subseteq \{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}.$$

Thus, the set $\{\text{entries of rows } 1,2,\ldots,i \text{ of } S\}$ is a **proper** subset of $\{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}$ (it is proper because the sets $\{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}$ and $\{\text{entries of rows } 1,2,\ldots,i \text{ of } S\}$ differ). Therefore,

$$|\{\text{entries of rows } 1,2,\ldots,i \text{ of } S\}| < |\{\text{entries of rows } 1,2,\ldots,i \text{ of } R\}|$$

(because if $X$ is a proper subset of a finite set $Y$, then $|X| < |Y|$). In view of (222) and (223), we can rewrite this as

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i < \mu_1 + \mu_2 + \cdots + \mu_i.$$

However, this contradicts the equality

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i = \mu_1 + \mu_2 + \cdots + \mu_i$$

(which holds because $\lambda = \mu$). This contradiction shows that our assumption was false. Thus, $r_{\overline{R}}(k) = r_{\overline{S}}(k)$ is proved. This proves Claim 3. □

We have $r_{\overline{R}}(i) = r_{\overline{S}}(i)$ for each $i \in [n]$ (by Claim 3, applied to $k = i$). Hence, Lemma 5.7.8 (applied to $P = R$ and $Q = S$) shows that $R$ and $S$ are row-equivalent. In other words, $S$ and $R$ are row-equivalent.

However, Proposition 5.3.14 **(a)** (applied to $Y(\mu)$, $S$ and $R$ instead of $D$, $T$ and $S$) says that the $n$-tableaux $S$ and $R$ are row-equivalent if and only if there exists some $w' \in \mathcal{R}(S)$ such that $R = w' \rightharpoonup S$ (note that we are using the letter $w'$ here for what was called $w$ in Proposition 5.3.14 **(a)**, since $w$ already has a different meaning in our context). Thus, there exists some $w' \in \mathcal{R}(S)$ such that $R = w' \rightharpoonup S$ (since $S$ and $R$ are row-equivalent). Consider this $w'$.

We have $R = w' \rightharpoonup S = w'S$. Comparing this with $R = wT$, we obtain $w'S = wT$. Thus, there exist permutations $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$ (namely, $r = w'$ and $c = w$). This proves Theorem 5.10.1 **(c)**. □

**Remark 5.10.3.** Theorem 5.10.1 can also be proved in several other ways:

- The most widely known proof of Theorem 5.10.1 **(a)** constructs the required permutation $c \in \mathcal{C}(T)$ iteratively, by successively raising those entries of $T$ that lie further south in $T$ than they do in $S$ so that they no longer do so. This proof appears, e.g., in [EGHetc11, proof of Lemma 5.13.1], [Fulton97, §7.1, proof of Lemma 1], [BrMaPe16, proof of Lemma 1.16], [Proces07, §9.2.3, proof of Lemma], [Morel19, proof of Proposition IV.2.7] and [JamKer81, proof

of Lemma 1.5.7] (usually implicitly, since these sources only prove Theorem 5.10.1 **(c)**, but the same argument can be used for part **(a)**). Several variants of this argument are possible, depending on how one chooses the order in which the entries are raised (most sources work in the order of increasing $S$-depth, but actually any order can be used). In a sense, our construction of $w$ in our above proof is doing the same thing, but it is doing so in one swoop rather than iteratively.

- Garsia proves parts **(b)** and **(c)** of Theorem 5.10.1 rather creatively in [GarEge20, §1.2], using an extra $n$-tableau $S \wedge T$ (which he calls $T_1 \wedge T_2$) that is defined as follows: Its shape is the set $D$ of all pairs $(i,j) \in \{1,2,3,\ldots\}^2$ such that the $i$-th row of $S$ has an entry in common with the $j$-th column of $T$. For any such cell $(i,j) \in D$, the entry of $S \wedge T$ in this cell is defined to be the unique common entry of the $i$-th row of $S$ and the $j$-th column of $T$ (the uniqueness here is guaranteed by the "there are no two distinct integers..." assumption of Theorem 5.10.1). Note that $S \wedge T$ is a bad-shaped tableau in general, but its very existence already yields Theorem 5.10.1 **(b)** quite easily, and with some more work yields Theorem 5.10.1 **(c)** as well.

- Another approach to Theorem 5.10.1 **(a)** is to construct the desired permutation $c \in \mathcal{C}(T)$ as a solution to a maximization problem: Namely, we define a number

$$\omega(c) := \sum_{k=1}^{n} r_{\overline{S}}(k)\, r_{\overline{cT}}(k) \in \mathbb{N} \qquad \text{(see Definition 5.7.5 \textbf{(a)} for the notation)}$$

for each $c \in \mathcal{C}(T)$, and we pick a permutation $c \in \mathcal{C}(T)$ that maximizes this number. We can then check that this $c$ has the property desired in Theorem 5.10.1 **(a)** (since otherwise, we can find two distinct entries $u$ and $v$ lying in the same column of $T$ such that $\omega(t_{u,v}c) > \omega(c)$).

**Exercise 5.10.1.** [1] Prove that Theorem 5.10.1 **(c)** has a converse: If there exist permutations $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$, then there are no two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$.

**Exercise 5.10.2.** [2] **(a)** Prove that the two permutations $r$ and $c$ in Theorem 5.10.1 **(c)** are unique.

[?] **(b)** Under what conditions is the $c$ in Theorem 5.10.1 **(a)** unique? Are there any good criteria that guarantee uniqueness?

**Exercise 5.10.3.** [1] In Theorem 5.10.1 **(a)**, can we replace $Y(\lambda)$ by a skew Young diagram? by an arbitrary diagram contained in $\{1,2,3,\ldots\}^2$ ?

Can we do this with $Y(\mu)$ ?

**Remark 5.10.4.** Parts of Theorem 5.10.1 can be generalized to partitions of different sizes. Indeed, let $m \in \mathbb{N}$ be a further nonnegative integer, and let $\mu$ be a partition of $m$ (instead of being a partition of $n$). Then, Theorem 5.10.1 remains valid if we additionally require $n \leq m$. Moreover, Theorem 5.10.1 **(a)** remains valid even if $n > m$ as long as we replace "Each entry of the $i$-th row of $S$" by "Each entry of the $i$-th row of $S$ that belongs to $[m]$". This all isn't too hard to show by adapting our above proof of Theorem 5.10.1. (For instance, we have to define $r_{\overline{S}}(k)$ to be some sufficiently large number when $k > n$, and likewise for $r_{\overline{T}}(k)$ when $k > m$.) We leave the details to the reader.

### 5.10.2. Algebraic consequences: the sandwich lemmas

The Young alternative has some useful corollaries, which we (following Adriano Garsia) call the *von Neumann sandwich lemmas*. Some of them seem appear already in Young's original works, but John von Neumann realized how they can be used to understand Specht modules.[84]

The simplest result of this type is a generalization of Theorem 3.7.11:

**Lemma 5.10.5** (symmetrizer cancellation lemma)**.** Let $S$ and $T$ be two $n$-tableaux of arbitrary shapes. Assume that there are two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$. Then,

$$\nabla^{-}_{\operatorname{Col} T} \nabla_{\operatorname{Row} S} = 0 \qquad \text{and} \qquad \nabla_{\operatorname{Row} S} \nabla^{-}_{\operatorname{Col} T} = 0.$$

*Proof.* We have assumed that there exist two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$. Let $i$ and $j$ be these integers.

In particular, $i$ and $j$ lie in the same column of $T$. Hence, Proposition 5.5.9 **(b)** shows that

$$\nabla^{-}_{\operatorname{Col} T} = \nabla^{\operatorname{even}}_{\operatorname{Col} T} \cdot \left(1 - t_{i,j}\right) = \left(1 - t_{i,j}\right) \cdot \nabla^{\operatorname{even}}_{\operatorname{Col} T},$$

where $\nabla^{\operatorname{even}}_{\operatorname{Col} T}$ is defined as in Proposition 5.5.9 **(b)**.

But $i$ and $j$ also lie in the same row of $S$. Hence, Proposition 5.5.8 **(b)** (applied to $S$ instead of $T$) shows that

$$\nabla_{\operatorname{Row} S} = \nabla^{\operatorname{even}}_{\operatorname{Row} S} \cdot \left(1 + t_{i,j}\right) = \left(1 + t_{i,j}\right) \cdot \nabla^{\operatorname{even}}_{\operatorname{Row} S},$$

where $\nabla^{\operatorname{even}}_{\operatorname{Row} S}$ is defined like the $\nabla^{\operatorname{even}}_{\operatorname{Row} T}$ in Proposition 5.5.8 **(b)** (but using $S$ instead of $T$). Thus,

$$\underbrace{\nabla^{-}_{\operatorname{Col} T}}_{=\nabla^{\operatorname{even}}_{\operatorname{Col} T} \cdot \left(1 - t_{i,j}\right)} \underbrace{\nabla_{\operatorname{Row} S}}_{=\left(1+t_{i,j}\right) \cdot \nabla^{\operatorname{even}}_{\operatorname{Row} S}} = \nabla^{\operatorname{even}}_{\operatorname{Col} T} \cdot \underbrace{\left(1 - t_{i,j}\right)\left(1 + t_{i,j}\right)}_{\substack{=1-t_{i,j}^2=0 \\ (\text{since } t_{i,j}^2 = \mathrm{id} = 1)}} \cdot \nabla^{\operatorname{even}}_{\operatorname{Row} S} = 0$$

[84] John von Neumann apparently never published on this topic, but his contributions appear in [Waerde91, volume 2, §14.7].

and

$$\underbrace{\nabla_{\operatorname{Row} S}}_{=\nabla^{\operatorname{even}}_{\operatorname{Row} S}\cdot\left(1+t_{i,j}\right)} \underbrace{\nabla^{-}_{\operatorname{Col} T}}_{=\left(1-t_{i,j}\right)\cdot\nabla^{\operatorname{even}}_{\operatorname{Col} T}} = \nabla^{\operatorname{even}}_{\operatorname{Row} S}\cdot\underbrace{\left(1+t_{i,j}\right)\left(1-t_{i,j}\right)}_{\substack{=1-t_{i,j}^2=0\\ (\text{since } t_{i,j}^2=\mathrm{id}=1)}}\cdot\nabla^{\operatorname{even}}_{\operatorname{Col} T} = 0.$$

Thus, Lemma 5.10.5 is proved. □

Now, we can state the first real "sandwich lemma":

**Theorem 5.10.6** (von Neumann sandwich lemma, version 1)**.** Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $Y(\lambda)$. Let $w \in S_n$. Then:

**(a)** If there exist two distinct integers that lie in the same row of $wT$ and simultaneously lie in the same column of $T$, then $\nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} = 0$.

**(b)** If not, then there exist two permutations $c \in \mathcal{C}(T)$ and $r \in \mathcal{R}(T)$ such that $w = cr$ and $\nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} = (-1)^c \nabla^{-}_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$.

The name "sandwich lemma" comes from viewing the product $\nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T}$ as "sandwiching" the $w$ between the two bread-pieces $\nabla^{-}_{\operatorname{Col} T}$ and $\nabla_{\operatorname{Row} T}$. Despite this name, we have chosen to declare it a theorem due to its significance.

*Proof of Theorem 5.10.6.* **(a)** Assume that there exist two distinct integers that lie in the same row of $wT$ and simultaneously lie in the same column of $T$. Then, Lemma 5.10.5 (applied to $S = wT$) yields

$$\nabla^{-}_{\operatorname{Col} T} \nabla_{\operatorname{Row}(wT)} = 0. \tag{226}$$

But Proposition 5.5.11 yields $\nabla_{\operatorname{Row}(w \rightharpoonup T)} = w \nabla_{\operatorname{Row} T} w^{-1}$. In other words, $\nabla_{\operatorname{Row}(wT)} = w \nabla_{\operatorname{Row} T} w^{-1}$ (since $w \rightharpoonup T = wT$). Hence, we can rewrite (226) as

$$\nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} w^{-1} = 0.$$

Multiplying this equality by $w$ on the right, we obtain $\nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} = 0$. Thus, Theorem 5.10.6 **(a)** is proved.

**(b)** Assume that there exist no two distinct integers that lie in the same row of $wT$ and simultaneously lie in the same column of $T$. Then, Theorem 5.10.1 **(c)** (applied to $S = wT$ and $\mu = \lambda$) thus yields that there exist permutations $r \in \mathcal{R}(wT)$ and $c \in \mathcal{C}(T)$ such that $rwT = cT$. Consider these $r$ and $c$, and let us denote them by $\widetilde{r}$ and $c$. Thus, $\widetilde{r} \in \mathcal{R}(wT)$ and $c \in \mathcal{C}(T)$ and $\widetilde{r} wT = cT$. (We are using the symbol $\widetilde{r}$ instead of $r$ because it is not the $r$ that we are looking for.)

It is easy to see that any two permutations $u, v \in S_n$ satisfying $uT = vT$ must necessarily be equal[85]. Applying this to $u = \widetilde{r}w$ and $v = c$, we obtain $\widetilde{r}w = c$ (since $\widetilde{r}wT = cT$).

Proposition 5.3.12 **(c)** yields $\mathcal{R}(w \rightharpoonup T) = w\mathcal{R}(T)w^{-1}$. In other words, $\mathcal{R}(wT) = w\mathcal{R}(T)w^{-1}$ (since $w \rightharpoonup T = wT$).

We have $\widetilde{r} \in \mathcal{R}(wT) = w\mathcal{R}(T)w^{-1}$. In other words, $\widetilde{r} = wr'w^{-1}$ for some $r' \in \mathcal{R}(T)$. Consider this $r'$. Then, $(r')^{-1} \in \mathcal{R}(T)$ (since $\mathcal{R}(T)$ is a group). Set $r := (r')^{-1}$. Then, $r = (r')^{-1} \in \mathcal{R}(T)$.

Hence, Proposition 5.5.5 **(a)** (applied to $r$ instead of $w$) yields $r\nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T} r = \nabla_{\operatorname{Row} T}$. Also, Proposition 5.5.5 **(b)** (applied to $c$ instead of $w$) yields $c\nabla^-_{\operatorname{Col} T} = \nabla^-_{\operatorname{Col} T} c = (-1)^c \nabla^-_{\operatorname{Col} T}$ (since $c \in \mathcal{C}(T)$).

We have $\widetilde{r}w = wr'$ (since $\widetilde{r} = wr'w^{-1}$). Comparing this with $\widetilde{r}w = c$, we obtain $wr' = c$, so that $w = c\underbrace{(r')^{-1}}_{=r} = cr$. Thus,

$$\nabla^-_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} = \underbrace{\nabla^-_{\operatorname{Col} T} c}_{=(-1)^c \nabla^-_{\operatorname{Col} T}} \underbrace{r\nabla_{\operatorname{Row} T}}_{=\nabla_{\operatorname{Row} T}} = (-1)^c \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}.$$

Thus, we have found two permutations $c \in \mathcal{C}(T)$ and $r \in \mathcal{R}(T)$ such that $w = cr$ and $\nabla^-_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} = (-1)^c \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$. This proves Theorem 5.10.6 **(b)**. □

**Corollary 5.10.7** (von Neumann sandwich lemma, version 2)**.** Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $Y(\lambda)$. Let $\mathbf{a} \in \mathbf{k}[S_n]$. Then,

$$\nabla^-_{\operatorname{Col} T} \mathbf{a} \nabla_{\operatorname{Row} T} = \kappa \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \qquad \text{for some scalar } \kappa \in \mathbf{k}.$$

(Of course, $\kappa$ depends on $\mathbf{a}$.)

*Proof.* Theorem 5.10.6 has the following consequence: For each permutation $w \in S_n$, there exists an integer $\kappa_w \in \mathbb{Z}$ such that

$$\nabla^-_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} = \kappa_w \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \tag{227}$$

[86]. Consider this integer $\kappa_w$.

[85] *Proof.* Let $u, v \in S_n$ be two permutations satisfying $uT = vT$. We must show that $u$ and $v$ are equal.

Recall that $T$ is an $n$-tableau. Thus, we can view $T$ as a bijection from $Y(\mu)$ to $[n]$. Hence, its inverse $T^{-1}$ is well-defined.

The definition of the action on entries (Definition 5.3.9) yields $uT = u \circ T$ and $vT = v \circ T$. Hence, $u \circ T = uT = vT = v \circ T$. Thus, $\underbrace{u \circ T}_{=v\circ T} \circ T^{-1} = v \circ \underbrace{T \circ T^{-1}}_{=\mathrm{id}} = v$. In view of $u \circ \underbrace{T \circ T^{-1}}_{=\mathrm{id}} = u$, we can rewrite this as $u = v$. In other words, $u$ and $v$ are equal, qed.

[86] *Proof.* If there exist two distinct integers that lie in the same row of $wT$ and simultaneously

Now, let us expand $\mathbf{a}$ in the standard basis of $\mathbf{k}[S_n]$. That is, let us write $\mathbf{a}$ in the form $\mathbf{a} = \sum_{w \in S_n} \alpha_w w$ for some scalars $\alpha_w \in \mathbf{k}$. Then,

$$\nabla^-_{\mathrm{Col}\, T} \mathbf{a} \nabla_{\mathrm{Row}\, T} = \nabla^-_{\mathrm{Col}\, T} \left( \sum_{w \in S_n} \alpha_w w \right) \nabla_{\mathrm{Row}\, T} = \sum_{w \in S_n} \alpha_w \underbrace{\nabla^-_{\mathrm{Col}\, T} w \nabla_{\mathrm{Row}\, T}}_{= \kappa_w \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T} \text{ (by (227))}}$$

$$= \sum_{w \in S_n} \alpha_w \kappa_w \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T} = \left( \sum_{w \in S_n} \alpha_w \kappa_w \right) \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}.$$

Thus, $\nabla^-_{\mathrm{Col}\, T} \mathbf{a} \nabla_{\mathrm{Row}\, T} = \kappa \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}$ for some scalar $\kappa \in \mathbf{k}$ (namely, for $\kappa = \sum_{w \in S_n} \alpha_w \kappa_w$). This proves Corollary 5.10.7. □

**Exercise 5.10.4.** [2] Generalize Corollary 5.10.7 to $\nabla^-_{\mathrm{Col}\, T} \mathbf{a} \nabla_{\mathrm{Row}\, S}$, where $T$ and $S$ are two $n$-tableaux of the same shape $Y(\lambda)$.

[**Hint:** The claim will have the form $\nabla^-_{\mathrm{Col}\, T} \mathbf{a} \nabla_{\mathrm{Row}\, S} = \kappa \mathbf{b}$ for some scalar $\kappa \in \mathbf{k}$ and some some fixed $\mathbf{b} \in \mathbf{k}[S_n]$ that is independent of $\mathbf{a}$. But $\mathbf{b}$ will not be $\nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, S}$ this time.]

**Exercise 5.10.5.** Prove that the scalar $\kappa$ in Corollary 5.10.7 can be computed by the following two explicit formulas:

**(a)** [2] If $\mathbf{a} = \sum_{w \in S_n} \alpha_w w$ for some scalars $\alpha_w \in \mathbf{k}$, then

$$\kappa = \sum_{c \in \mathcal{C}(T)} \sum_{r \in \mathcal{R}(T)} (-1)^c \alpha_{cr}.$$

**(b)** [2] If $\mathbf{a} \nabla_{\mathrm{Row}\, T} \nabla^-_{\mathrm{Col}\, T} = \sum_{w \in S_n} \beta_w w$ for some scalars $\beta_w \in \mathbf{k}$, then $\kappa = \beta_{\mathrm{id}}$.

We furthermore will use an analogue of the von Neumann sandwich lemma which involves two different shapes:

**Theorem 5.10.8** (von Neumann zero-sandwich lemma)**.** Let $\lambda$ and $\mu$ be two partitions of $n$. Assume that $\lambda_1 + \lambda_2 + \cdots + \lambda_i > \mu_1 + \mu_2 + \cdots + \mu_i$ for some $i \geq 1$.

Let $S$ be an $n$-tableau of shape $Y(\lambda)$. Let $T$ be an $n$-tableau of shape $Y(\mu)$. Then:

lie in the same column of $T$, then we can choose $\kappa_w = 0$, because Theorem 5.10.6 **(a)** shows that $\nabla^-_{\mathrm{Col}\, T} w \nabla_{\mathrm{Row}\, T} = 0$ in this case. In the contrary case, Theorem 5.10.6 **(b)** yields that $\nabla^-_{\mathrm{Col}\, T} w \nabla_{\mathrm{Row}\, T} = (-1)^c \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}$ for some $c \in \mathcal{C}(T)$ and some $r \in \mathcal{R}(T)$, and thus we can choose $\kappa_w = (-1)^c$. In either case, we have thus found an integer $\kappa_w \in \mathbb{Z}$ such that $\nabla^-_{\mathrm{Col}\, T} w \nabla_{\mathrm{Row}\, T} = \kappa_w \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}$.

**(a)** We have $\nabla^-_{\operatorname{Col} T} w \nabla_{\operatorname{Row} S} = 0$ for each $w \in S_n$.

**(b)** We have $\nabla^-_{\operatorname{Col} T} \mathbf{a} \nabla_{\operatorname{Row} S} = 0$ for each $\mathbf{a} \in \mathbf{k}[S_n]$.

**(c)** We have $\nabla_{\operatorname{Row} S} w \nabla^-_{\operatorname{Col} T} = 0$ for each $w \in S_n$.

**(d)** We have $\nabla_{\operatorname{Row} S} \mathbf{a} \nabla^-_{\operatorname{Col} T} = 0$ for each $\mathbf{a} \in \mathbf{k}[S_n]$.

*Proof.* **(a)** Let $w \in S_n$. Then, $wS$ is an $n$-tableau of shape $Y(\lambda)$ (since $S$ is a such). Hence, by Theorem 5.10.1 **(b)** (or, rather, by its contrapositive), we can see that there are two distinct integers that lie in the same row of $wS$ and simultaneously lie in the same column of $T$ [87]. Hence, Lemma 5.10.5 (applied to $wS$ instead of $S$) yields

$$\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row}(wS)} = 0 \qquad \text{and} \qquad \nabla_{\operatorname{Row}(wS)} \nabla^-_{\operatorname{Col} T} = 0.$$

But Proposition 5.5.11 (applied to $S$ instead of $T$) yields $\nabla_{\operatorname{Row}(w \rightharpoonup S)} = w \nabla_{\operatorname{Row} S} w^{-1}$. In other words, $\nabla_{\operatorname{Row}(wS)} = w \nabla_{\operatorname{Row} S} w^{-1}$ (since $w \rightharpoonup S = wS$). Hence, $\nabla_{\operatorname{Row}(wS)} w = w \nabla_{\operatorname{Row} S}$ and therefore

$$\nabla^-_{\operatorname{Col} T} \underbrace{w \nabla_{\operatorname{Row} S}}_{= \nabla_{\operatorname{Row}(wS)} w} = \underbrace{\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row}(wS)}}_{=0} w = 0.$$

This proves Theorem 5.10.8 **(a)**.

**(b)** This follows from part **(a)** by linearity.[88]

**(c)** This can be done similarly to part **(a)**, but using the $\nabla_{\operatorname{Row}(wS)} \nabla^-_{\operatorname{Col} T} = 0$ equality instead of $\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row}(wS)} = 0$. (More precisely, this way we can obtain $\nabla_{\operatorname{Row} S} w^{-1} \nabla^-_{\operatorname{Col} T} = 0$, but since $w$ is arbitrary, we can apply this to $w^{-1}$ instead of $w$ to recover the equality we are looking for).

Alternatively, we can derive Theorem 5.10.8 **(c)** from part **(a)** using the antipode. This technique is useful in various situations, so let me show this proof in some detail:

Let us rename the $n$-tableau $S$ as $R$ (in order to avoid confusing it with the antipode $S$ of $\mathbf{k}[S_n]$). Thus, we must prove that $\nabla_{\operatorname{Row} R} w \nabla^-_{\operatorname{Col} T} = 0$ for each $w \in S_n$.

[87] *Proof.* Assume the contrary. Thus, there are no two distinct integers that lie in the same row of $wS$ and simultaneously lie in the same column of $T$. Hence, Theorem 5.10.1 **(b)** (applied to $wS$ instead of $S$) shows that we have $\lambda_1 + \lambda_2 + \cdots + \lambda_i \leq \mu_1 + \mu_2 + \cdots + \mu_i$ for each $i \geq 1$. But this contradicts the fact that $\lambda_1 + \lambda_2 + \cdots + \lambda_i > \mu_1 + \mu_2 + \cdots + \mu_i$ for some $i \geq 1$. This contradiction shows that our assumption was false, qed.

[88] In more detail: We must prove the equality $\nabla^-_{\operatorname{Col} T} \mathbf{a} \nabla_{\operatorname{Row} S} = 0$ for each $\mathbf{a} \in \mathbf{k}[S_n]$. Since both sides of this equality are $\mathbf{k}$-linear in $\mathbf{a}$, it suffices to prove it when $\mathbf{a}$ is a standard basis vector of $\mathbf{k}[S_n]$ (since the standard basis vectors span the $\mathbf{k}$-module $\mathbf{k}[S_n]$). In other words, it suffices to prove it when $\mathbf{a} = w$ for some permutation $w \in S_n$. In other words, it suffices to prove that $\nabla^-_{\operatorname{Col} T} w \nabla_{\operatorname{Row} S} = 0$ for each $w \in S_n$. But this follows immediately from Theorem 5.10.8 **(a)**. Thus, Theorem 5.10.8 **(b)** is proved.

Fix $w \in S_n$. Then, Theorem 5.10.8 **(a)** (applied to $w^{-1}$ instead of $w$) yields $\nabla^-_{\operatorname{Col} T} w^{-1} \nabla_{\operatorname{Row} R} = 0$ (since we have renamed the $n$-tableau $S$ as $R$). Applying the antipode $S$ to both sides of this equality, we obtain $S\left(\nabla^-_{\operatorname{Col} T} w^{-1} \nabla_{\operatorname{Row} R}\right) = S(0) = 0$. Hence,

$$\begin{aligned}
0 &= S\left(\nabla^-_{\operatorname{Col} T} w^{-1} \nabla_{\operatorname{Row} R}\right) \\
&= \underbrace{S\left(\nabla_{\operatorname{Row} R}\right)}_{\substack{=\nabla_{\operatorname{Row} R} \\ \text{(by Proposition 5.5.10,} \\ \text{applied to } R \text{ instead of } T\text{)}}} \underbrace{S\left(w^{-1}\right)}_{\substack{=\left(w^{-1}\right)^{-1} \\ \text{(by the definition of } S\text{)}}} \underbrace{S\left(\nabla^-_{\operatorname{Col} T}\right)}_{\substack{=\nabla^-_{\operatorname{Col} T} \\ \text{(by Proposition 5.5.10)}}} \\
&\qquad \left(\text{since } S \text{ is a } \mathbf{k}\text{-algebra anti-automorphism (by Theorem 3.11.14 } \textbf{(a)}\right)) \\
&= \nabla_{\operatorname{Row} R} \underbrace{\left(w^{-1}\right)^{-1}}_{=w} \nabla^-_{\operatorname{Col} T} = \nabla_{\operatorname{Row} R} w \nabla^-_{\operatorname{Col} T}.
\end{aligned}$$

Thus, $\nabla_{\operatorname{Row} R} w \nabla^-_{\operatorname{Col} T} = 0$. This is precisely what we wanted to prove (since we renamed the $n$-tableau $S$ as $R$). Thus, Theorem 5.10.8 **(c)** is proved.

**(d)** This follows from part **(c)** just like part **(b)** follows from part **(a)**. □

Let us observe yet another easy consequence of Theorem 5.10.5, connecting the row-symmetrizers and the column-antisymmetrizers with the Young last letter order from Definition 5.7.5:

**Lemma 5.10.9.** Let $\lambda$ be a partition of $n$. Let $S$ and $T$ be two $n$-tableaux of shape $Y(\lambda)$ such that $T$ is column-standard. Assume that $\overline{T} < \overline{S}$ with respect to the Young last letter order (i.e., the binary relation $<$ defined in Definition 5.7.5 **(b)**). Then:

**(a)** There exist no permutations $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$.

**(b)** There are two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$.

**(c)** We have

$$\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} S} = 0 \qquad \text{and} \qquad \nabla_{\operatorname{Row} S} \nabla^-_{\operatorname{Col} T} = 0.$$

*Proof.* **(a)** Assume the contrary. Thus, there exist permutations $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$. Consider such $r$ and $c$. Consider the $n$-tabloids $\overline{S}$ and $\overline{T}$. We have $r \in \mathcal{R}(S) = \left\{w \in S_n \mid \overline{w \rightharpoonup S} = \overline{S}\right\}$ (by Proposition 5.3.13 **(a)**, applied to $S$ instead of $T$). In other words, $r \in S_n$ and $\overline{r \rightharpoonup S} = \overline{S}$. But $rS$ is shorthand for $r \rightharpoonup S$. Thus, $rS = r \rightharpoonup S$, so that $\overline{rS} = \overline{r \rightharpoonup S} = \overline{S}$. In other words, $\overline{cT} = \overline{S}$ (since $rS = cT$).

But we have $c \in \mathcal{C}(T)$; therefore, we obtain $\overline{cT} \leq \overline{T}$ as an easy consequence of Lemma 5.7.9 **(a)**[89]. In view of $\overline{cT} = \overline{S}$, we can rewrite this as $\overline{S} \leq \overline{T}$. But this contradicts $\overline{T} < \overline{S}$. This contradiction shows that our assumption was false. Thus, the proof of Lemma 5.10.9 **(a)** is complete.

**(b)** Assume the contrary. Thus, there are no two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$. Hence, Theorem 5.10.1 **(c)** (applied to $\mu = \lambda$) yields that there exist permutations $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$. But this contradicts Lemma 5.10.9 **(a)**. This contradiction shows that our assumption was false. Thus, the proof of Lemma 5.10.9 **(b)** is complete.

**(c)** Lemma 5.10.9 **(b)** shows that there are two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$. Hence, Theorem 5.10.5 yields

$$\nabla^{-}_{\operatorname{Col} T} \nabla_{\operatorname{Row} S} = 0 \qquad \text{and} \qquad \nabla_{\operatorname{Row} S} \nabla^{-}_{\operatorname{Col} T} = 0.$$

This proves Lemma 5.10.9 **(b)**. □

## 5.11. Young symmetrizers

### 5.11.1. Definitions

The sandwich lemmas will help us understand the Specht modules $\mathcal{S}^{Y(\lambda)}$ better. First, we introduce a shorthand for them:

**Definition 5.11.1.** **(a)** Given any partition $\lambda$ of size $n$, we set $\mathcal{S}^{\lambda} := \mathcal{S}^{Y(\lambda)}$ (the Specht module for shape $Y(\lambda)$).

**(b)** For any $n$-tableau $T$ (of any shape), we set

$$\mathbf{E}_T := \nabla^{-}_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \in \mathbf{k}[S_n].$$

The element $\mathbf{E}_T$ is commonly called a *Young symmetrizer*.

**(c)** Furthermore, we set $\mathcal{A} := \mathbf{k}[S_n]$ for the rest of this chapter.

The straight-shaped Specht modules $\mathcal{S}^{\lambda} = \mathcal{S}^{Y(\lambda)}$ are a particular case of the skew-shaped Specht modules $\mathcal{S}^{Y(\lambda/\mu)}$ (since $Y(\lambda) = Y(\lambda/\varnothing)$ for any partition $\lambda$). As such, they are fairly well-understood (in particular, Theorem 5.9.1 provides a basis for each of them). However, they have a special role to play: If $\mathbf{k}$ is a field of characteristic 0, then they are precisely the irreps (= irreducible representations) of $S_n$ over $\mathbf{k}$.

[89] *Proof:* We must show that $\overline{cT} \leq \overline{T}$. If $c \neq \operatorname{id}$, then Lemma 5.7.9 **(a)** (applied to $w = c$) yields that $\overline{cT} < \overline{T}$, and thus $\overline{cT} \leq \overline{T}$ follows. If $c = \operatorname{id}$, then $cT = \operatorname{id} T = T$ and thus $\overline{cT} = \overline{T}$, so that $\overline{cT} \leq \overline{T}$. Thus, in either case, we have shown that $\overline{cT} \leq \overline{T}$.

To prove this will take us some work. First, we recall the left ideal avatar of a Specht module (Theorem 5.5.13 **(b)**): If $D$ is any diagram, and if $T$ is an $n$-tableau of shape $D$, then

$$\begin{aligned}\mathcal{S}^D \cong & \underbrace{\mathbf{k}\left[S_n\right]}_{\substack{=\mathcal{A} \\ \text{(by Definition 5.11.1 (c))}}} \cdot \underbrace{\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}}_{\substack{=\mathbf{E}_T \\ \text{(by Definition 5.11.1 (b))}}} \qquad \text{(by Theorem 5.5.13 (b))} \\ = & \mathcal{A}\mathbf{E}_T \end{aligned} \tag{228}$$

as left $\mathbf{k}\left[S_n\right]$-modules (i.e., as $S_n$-representations). Thus, we can hope to understand $\mathcal{S}^D$ by understanding the Young symmetrizer $\mathbf{E}_T$. This is true for any diagram $D$, but will be particularly useful in the case when $D$ is a straight Young diagram $Y(\lambda)$.

**Example 5.11.2.** Let $n = 3$. Then, each $n$-tableau of straight shape has one of the three forms

$$\begin{array}{|c|c|c|}\hline i & j & k \\ \hline \end{array} \quad \text{and} \quad \begin{array}{|c|}\hline i \\ \hline j \\ \hline k \\ \hline \end{array} \quad \text{and} \quad \begin{array}{|c|c|}\hline i & j \\ \hline k \\ \cline{1-1} \end{array} \,,$$

where $i, j, k$ are the three elements of $[n]$ in some order. Using the shorthand notation from Convention 5.2.4, we can rewrite these three forms as $ijk$ and $i\backslash\backslash j\backslash\backslash k$ and $ij\backslash\backslash k$, respectively. Thus, the Young symmetrizers for straight-shaped $n$-tableaux are the elements

$$\mathbf{E}_{ijk} = \underbrace{\nabla^-_{\operatorname{Col}(ijk)}}_{=1} \underbrace{\nabla_{\operatorname{Row}(ijk)}}_{=\nabla} = \nabla = 1 + s_1 + s_2 + t_{1,3} + \operatorname{cyc}_{1,2,3} + \operatorname{cyc}_{1,3,2};$$

$$\mathbf{E}_{i\backslash\backslash j\backslash\backslash k} = \underbrace{\nabla^-_{\operatorname{Col}(i\backslash\backslash j\backslash\backslash k)}}_{=\nabla^-} \underbrace{\nabla_{\operatorname{Row}(i\backslash\backslash j\backslash\backslash k)}}_{=1} = \nabla^- = 1 - s_1 - s_2 - t_{1,3} + \operatorname{cyc}_{1,2,3} + \operatorname{cyc}_{1,3,2};$$

$$\mathbf{E}_{ij\backslash\backslash k} = \underbrace{\nabla^-_{\operatorname{Col}(ij\backslash\backslash k)}}_{=1-t_{i,k}} \underbrace{\nabla_{\operatorname{Row}(ij\backslash\backslash k)}}_{=1+t_{i,j}} = \left(1 - t_{i,k}\right)\left(1 + t_{i,j}\right) = 1 - t_{i,k} + t_{i,j} - \operatorname{cyc}_{i,j,k}$$

for all such orderings $i, j, k$. (You are reading right: Both $\mathbf{E}_{ijk}$ and $\mathbf{E}_{i\backslash\backslash j\backslash\backslash k}$ are independent on $i, j, k$.)

**Exercise 5.11.1. (a)** [1] Prove that $\mathbf{E}_{123\backslash\backslash 456} = \mathbf{E}_{465\backslash\backslash 132}$ for $n = 6$, without fully expanding either side of the equality.

**(b)** [4] Let $P$ and $Q$ be two $n$-tableaux of any shapes (not necessarily of the same shape!). Assume that $2 \neq 0$ in $\mathbf{k}$. Prove that $\mathbf{E}_P = \mathbf{E}_Q$ holds if and only if the equalities $\mathcal{R}(P) = \mathcal{R}(Q)$ and $\mathcal{C}(P) = \mathcal{C}(Q)$ hold.

[**Hint:** In part **(b)**, consider the coefficients with which transpositions occur in $\mathbf{E}_P$ and $\mathbf{E}_Q$.]

### 5.11.2. Quasi-idempotency

When $D$ is a Young diagram, the Young symmetrizer $\mathbf{E}_T$ is particularly nice – it is what is called a *quasi-idempotent*:

**Theorem 5.11.3** (Young symmetrizer theorem)**.** Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $Y(\lambda)$. Let $f^\lambda$ be the # of standard tableaux of shape $Y(\lambda)$. Then, $\dfrac{n!}{f^\lambda}$ is a positive integer, and we have

$$\mathbf{E}_T^2 = \frac{n!}{f^\lambda}\mathbf{E}_T.$$

(Recall that $\mathbf{E}_T = \nabla^-_{\operatorname{Col} T}\nabla_{\operatorname{Row} T}$, as we said in Definition 5.11.1 **(b)**.)

**Example 5.11.4.** Let $n = 3$ and $\lambda = (2,1)$ and $T = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}$. Then,

$$\mathbf{E}_T = \underbrace{\nabla^-_{\operatorname{Col} T}}_{=1-t_{1,3}}\ \underbrace{\nabla_{\operatorname{Row} T}}_{=1+t_{1,2}} = (1-t_{1,3})(1+t_{1,2}) = 1 - t_{1,3} + t_{1,2} - \operatorname{cyc}_{1,2,3}.$$

So $\mathbf{E}_T^2$ is a sum of 16 terms. Theorem 5.11.3 tells us that this sum simplifies to

$$\frac{n!}{f^\lambda}\mathbf{E}_T = \frac{3!}{2}\mathbf{E}_T = 3\mathbf{E}_T = 3\left(1 - t_{1,3} + t_{1,2} - \operatorname{cyc}_{1,2,3}\right).$$

Note that Theorem 5.11.3 is not (generally) true when $Y(\lambda)$ is replaced by a skew Young diagram; in this case, $\mathbf{E}_T^2$ is usually not a scalar multiple of $\mathbf{E}_T$. [90]

[90] The simplest counterexample is obtained when $\mathbf{k} = \mathbb{Q}$ and $n = 4$ and $T$ is a 4-tableau of shape $Y((3,2)/(1))$. In this case, $\mathbf{E}_T^2$ is not a scalar multiple of $\mathbf{E}_T$, and there is no nontrivial quadratic polynomial that vanishes at $\mathbf{E}_T$. Rather, the minimal polynomial of $\mathbf{E}_T$ (that is, the monic polynomial $P \in \mathbb{Q}[x]$ of smallest degree that satisfies $P(\mathbf{E}_T) = 0$) is $(x-6)(x-4)x$ in this case.

This example might suggest that in general, the minimal polynomial of $\mathbf{E}_T$ might at least be split over $\mathbb{Z}$ (that is, all its roots are integers). Alas, this pattern is not sustained either. For example, if $\mathbf{k} = \mathbb{Q}$ and $n = 7$ and if $T$ is a 7-tableau of shape $Y((5,3,2)/(2,1))$, then the minimal polynomial of $\mathbf{E}_T$ is

$$\underbrace{\left(x^2 - 62x + 864\right)}_{\text{irreducible over } \mathbb{Q}}(x-48)(x-38)(x-32)(x-14)x.$$

The proof of Theorem 5.11.3 is rather surprising: It is very easy (using the von Neumann sandwich lemma) to show that $\mathbf{E}_T^2 = \kappa \mathbf{E}_T$ for some $\kappa \in \mathbf{k}$, but it is much harder to identify this $\kappa$ as $\dfrac{n!}{f^\lambda}$. We will thus approach this proof slowly, using several lemmas that seem unrelated yet will all find their use. First, we handle the sandwich part, in fact proving a more general version thereof:

**Proposition 5.11.5** (double sandwich lemma)**.** Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $Y(\lambda)$. Let $\mathbf{a} \in \mathbf{k}[S_n]$. Then, $\mathbf{E}_T \mathbf{a} \mathbf{E}_T = \kappa \mathbf{E}_T$ for some scalar $\kappa \in \mathbf{k}$.

*Proof.* We have $\mathbf{E}_T = \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$. Hence,

$$\begin{aligned} \mathbf{E}_T \mathbf{a} \mathbf{E}_T &= \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \mathbf{a} \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} = \nabla^-_{\operatorname{Col} T} \left( \nabla_{\operatorname{Row} T} \mathbf{a} \nabla^-_{\operatorname{Col} T} \right) \nabla_{\operatorname{Row} T} \\ &= \kappa \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \qquad \text{for some scalar } \kappa \in \mathbf{k} \end{aligned}$$

(by Corollary 5.10.7, applied to $\nabla_{\operatorname{Row} T} \mathbf{a} \nabla^-_{\operatorname{Col} T}$ instead of $\mathbf{a}$). In other words, $\mathbf{E}_T \mathbf{a} \mathbf{E}_T = \kappa \mathbf{E}_T$ for some scalar $\kappa \in \mathbf{k}$ (since $\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} = \mathbf{E}_T$). This proves Proposition 5.11.5. □

Clearly, applying Proposition 5.11.5 to $\mathbf{a} = 1$ yields $\mathbf{E}_T 1 \mathbf{E}_T = \kappa \mathbf{E}_T$, that is, $\mathbf{E}_T^2 = \kappa \mathbf{E}_T$. Now, we "only" need to show that this $\kappa$ is $\dfrac{n!}{f^\lambda}$. As we said, this will require a significant amount of preparation. We begin by introducing a notation:

**Definition 5.11.6.** Let $G$ be a group. Let $w \in G$ and $\mathbf{a} \in \mathbf{k}[G]$. Then, $[w]\,\mathbf{a}$ denotes the coefficient of $w$ in $\mathbf{a}$ (when $\mathbf{a}$ is expanded as a linear combination of the standard basis vectors of $\mathbf{k}[G]$). That is, if $\mathbf{a} = \sum_{g \in G} \alpha_g g$ for some scalars $\alpha_g \in \mathbf{k}$, then $[w]\,\mathbf{a} := \alpha_w$. The scalar $[w]\,\mathbf{a}$ is also known as the *$w$-coordinate* of $\mathbf{a}$.

For instance, the third jucys–murphy $\mathbf{m}_3 \in \mathbf{k}[S_n]$ equals $t_{1,3} + s_2$, so that we have

$$[s_2]\,\mathbf{m}_3 = 1 \qquad \text{but} \qquad [s_1]\,\mathbf{m}_3 = 0.$$

Also, for all $w \in S_n$, we have $[w]\,\nabla = 1$ and $[w]\,\nabla^- = (-1)^w$.

We shall furthermore use some elementary linear algebra, in particular the notion of a *trace*. The trace is defined both for a square matrix and for an endomorphism of a free $\mathbf{k}$-module of finite rank. Let us recall the definitions:

- The *trace* of a square matrix $M$ is the sum of the diagonal entries of $M$. It is denoted by $\operatorname{Tr} M$. For example, $\operatorname{Tr} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + d$.

  The trace has several nice properties. In particular, any two conjugate matrices have the same trace.

- If $V$ is a free $\mathbf{k}$-module of finite rank (i.e., with a finite basis), and if $f$ is any endomorphism of $V$ (that is, any $\mathbf{k}$-linear map from $V$ to $V$), then the *trace* of $f$ is defined as the trace of the matrix that represents $f$ with respect to any chosen basis of $V$. (The choice of the basis does not affect the trace[91].) Again, the trace of $f$ is denoted by $\operatorname{Tr} f$. Note that it belongs to $\mathbf{k}$.

Certain maps have particularly simple traces:

**Proposition 5.11.7.** Let $G$ be a finite group. Let $\mathbf{a} \in \mathbf{k}[G]$. Let $f_{\mathbf{a}} : \mathbf{k}[G] \to \mathbf{k}[G]$ be the $\mathbf{k}$-linear map that sends each $\mathbf{b} \in \mathbf{k}[G]$ to $\mathbf{ba}$. (For obvious reasons, this map $f_{\mathbf{a}}$ is called "right multiplication by $\mathbf{a}$"). Then,

$$\operatorname{Tr}(f_{\mathbf{a}}) = |G| \cdot [1]\,\mathbf{a}$$

(where 1 stands for the neutral element of $G$).

*Proof.* First, we show a general formula for traces of $\mathbf{k}$-module endomorphisms of $\mathbf{k}[G]$:

*Claim 1:* Let $f$ be any $\mathbf{k}$-module endomorphism of $\mathbf{k}[G]$. Then,

$$\operatorname{Tr} f = \sum_{g \in G} [g]\,(f(g)).$$

*Proof of Claim 1.* Let $g_1, g_2, \ldots, g_m$ be the elements of $G$, listed without repetition (so that $m = |G|$). Then, $\mathbf{k}$-module $\mathbf{k}[G]$ is free, with basis $(g_1, g_2, \ldots, g_m)$. The matrix that represents the $\mathbf{k}$-linear map $f$ with respect to this basis is the $m \times m$-matrix whose $(i,j)$-th entry is $[g_i]\left(f\left(g_j\right)\right)$ for all $i \in [m]$ and $j \in [m]$. Hence, in particular, the diagonal entries of this matrix are the elements $[g_i]\,(f(g_i))$ for all $i \in [m]$. Therefore, the trace of this matrix equals $\sum_{i=1}^{m} [g_i]\,(f(g_i))$ (since the trace of a matrix is the sum of the diagonal entries of this matrix). In other words,

$$\operatorname{Tr} f = \sum_{i=1}^{m} [g_i]\,(f(g_i)) \tag{229}$$

(since $\operatorname{Tr} f$ is defined to be the trace of this matrix).

But $g_1, g_2, \ldots, g_m$ are the elements of $G$, listed without repetition. Thus, $\sum_{g \in G} [g]\,(f(g)) = \sum_{i=1}^{m} [g_i]\,(f(g_i))$. Comparing this with (229), we obtain $\operatorname{Tr} f = \sum_{g \in G} [g]\,(f(g))$. This proves Claim 1. ∎

[91] Indeed, any two matrices that each represent $f$ with respect to some basis of $V$ will necessarily be conjugate to each other, and thus will have the same trace (since any two conjugate matrices have the same trace).

Now, let us consider the $\mathbf{k}$-module endomorphism $f_{\mathbf{a}}$ in particular. Write the element $\mathbf{a} \in \mathbf{k}[G]$ as $\mathbf{a} = \sum_{g \in G} \alpha_g g$ for some scalars $\alpha_g \in \mathbf{k}$. Then, $[1]\,\mathbf{a} = \alpha_1$ (by Definition 5.11.6). For each $h \in G$, we have

$$\begin{aligned} f_{\mathbf{a}}(h) &= h\mathbf{a} \qquad \text{(by the definition of } f_{\mathbf{a}}\text{)} \\ &= h \sum_{g \in G} \alpha_g g \qquad \left( \text{since } \mathbf{a} = \sum_{g \in G} \alpha_g g \right) \\ &= \sum_{g \in G} \alpha_g hg = \sum_{g \in G} \alpha_{h^{-1}g} \underbrace{hh^{-1}}_{=1} g \\ &\qquad \left( \begin{array}{c} \text{here, we have substituted } h^{-1}g \text{ for } g \text{ in the sum,} \\ \text{since the map } G \to G,\ g \mapsto h^{-1}g \text{ is a bijection} \\ \text{(because } G \text{ is a group)} \end{array} \right) \\ &= \sum_{g \in G} \alpha_{h^{-1}g} g, \end{aligned} \tag{230}$$

and thus

$$[g]\,(f_{\mathbf{a}}(h)) = \alpha_{h^{-1}g} \qquad \text{for each } g \in G \tag{231}$$

(by Definition 5.11.6 again).

Now, Claim 1 (applied to $f = f_{\mathbf{a}}$) yields

$$\operatorname{Tr}(f_{\mathbf{a}}) = \sum_{g \in G} \underbrace{[g]\,(f_{\mathbf{a}}(g))}_{=\alpha_{g^{-1}g}} = \sum_{g \in G} \underbrace{\alpha_{g^{-1}g}}_{=\alpha_1} = \sum_{g \in G} \alpha_1 = |G| \cdot \alpha_1.$$

(by (231), applied to $h{=}g$) (since $g^{-1}g{=}1$)

We can rewrite this as $\operatorname{Tr}(f_{\mathbf{a}}) = |G| \cdot [1]\,\mathbf{a}$ (since $[1]\,\mathbf{a} = \alpha_1$). This proves Proposition 5.11.7. □

We next need some more general properties of traces. The first one is a basic and well-known property of traces of matrices ([Axler24, 8.49]):

**Lemma 5.11.8.** Let $k, d \in \mathbb{N}$. Let $A \in \mathbf{k}^{d \times k}$ and $B \in \mathbf{k}^{k \times d}$ be two matrices. Then, $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$.

Note that the two matrices $AB$ and $BA$ are both square, but of different sizes. We can easily translate Lemma 5.11.8 into the language of linear maps:

**Lemma 5.11.9.** Let $V$ and $W$ be two free $\mathbf{k}$-modules of finite ranks (not necessarily of equal ranks!). Let $f : V \to W$ and $g : W \to V$ be two $\mathbf{k}$-linear maps. Then, $\operatorname{Tr}(f \circ g) = \operatorname{Tr}(g \circ f)$.

*Proof.* Pick a basis $(v_1, v_2, \ldots, v_k)$ of $V$ and a basis $(w_1, w_2, \ldots, w_d)$ of $W$. Let $A \in \mathbf{k}^{d \times k}$ be the matrix that represents the map $f : V \to W$ with respect to these two bases, and let $B \in \mathbf{k}^{k \times d}$ be the matrix that represents the map

$g : W \to V$ with respect to these two bases. Then, $AB$ is the matrix that represents the map $f \circ g : V \to V$ with respect to the basis $(v_1, v_2, \ldots, v_k)$ of $V$ (since matrix multiplication is "the same as" composition of linear maps). Thus, $\operatorname{Tr}(f \circ g) = \operatorname{Tr}(AB)$ (by the definition of the trace of a linear map). Similarly, $\operatorname{Tr}(g \circ f) = \operatorname{Tr}(BA)$. However, Lemma 5.11.8 says that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. In other words, $\operatorname{Tr}(f \circ g) = \operatorname{Tr}(g \circ f)$ (since $\operatorname{Tr}(f \circ g) = \operatorname{Tr}(AB)$ and $\operatorname{Tr}(g \circ f) = \operatorname{Tr}(BA)$). This proves Lemma 5.11.9. □

Our next property of traces is even more obvious:

**Lemma 5.11.10.** Let $d \in \mathbb{N}$. Let $W$ be a free $\mathbf{k}$-module of rank $d$. Let $\kappa \in \mathbf{k}$ be a scalar. Then, the $\mathbf{k}$-module endomorphism $\kappa \operatorname{id}_W$ of $W$ has trace $\operatorname{Tr}(\kappa \operatorname{id}_W) = d\kappa$.

*Proof.* The $\mathbf{k}$-module $W$ is free of rank $d$. Thus, it has a basis $(w_1, w_2, \ldots, w_d)$ consisting of $d$ vectors. Consider this basis.

Now, consider the $\mathbf{k}$-module endomorphism $\operatorname{id}_W$ of $W$. The matrix that represents this endomorphism with respect to the basis $(w_1, w_2, \ldots, w_d)$ is the $d \times d$ identity matrix (since the matrix that represents an identity map with respect to a basis is always the respective identity matrix). Thus, the diagonal entries of matrix are $\underbrace{1_{\mathbf{k}}, 1_{\mathbf{k}}, \ldots, 1_{\mathbf{k}}}_{d \text{ times}}$ (where $1_{\mathbf{k}}$ denotes the unity of $\mathbf{k}$), and therefore its trace is $\underbrace{1_{\mathbf{k}} + 1_{\mathbf{k}} + \cdots + 1_{\mathbf{k}}}_{d \text{ times}} = d \cdot 1_{\mathbf{k}}$. Hence, the trace of the $\mathbf{k}$-module endomorphism $\operatorname{id}_W$ is $d \cdot 1_{\mathbf{k}}$ as well (since the trace of a $\mathbf{k}$-module endomorphism is defined as the trace of a matrix that represents it with respect to a basis). In other words, $\operatorname{Tr}(\operatorname{id}_W) = d \cdot 1_{\mathbf{k}}$.

Recall that the trace of a $\mathbf{k}$-module endomorphism depends $\mathbf{k}$-linearly on this endomorphism. Hence, $\operatorname{Tr}(\omega j) = \omega \operatorname{Tr} j$ for any scalar $\omega \in \mathbf{k}$ and any $\mathbf{k}$-module endomorphism $j$. Thus, $\operatorname{Tr}(\kappa \operatorname{id}_W) = \kappa \underbrace{\operatorname{Tr}(\operatorname{id}_W)}_{= d \cdot 1_{\mathbf{k}}} = \kappa \cdot d \cdot 1_{\mathbf{k}} = d\kappa$. This proves Lemma 5.11.10. □

The next lemma is technical but crucial:

**Lemma 5.11.11.** Let $V$ be a free $\mathbf{k}$-module of finite rank. Let $f : V \to V$ be an endomorphism of $V$. Assume that the image $f(V)$ of $f$ is a free $\mathbf{k}$-module of rank $d$ for some $d \in \mathbb{N}$. Assume furthermore that $f^2 = \kappa f$ for some scalar $\kappa \in \mathbf{k}$ (where, as usual, $f^2$ means $f \circ f$). Then, $\operatorname{Tr} f = d\kappa$.

Before we come to the proof of this lemma, let me say a few words about its significance. We shall apply it to the situation where $V$ is $\mathcal{A} = \mathbf{k}[S_n]$ and where $f$ is right multiplication by $\mathbf{E}_T$ for some $n$-tableau $T$ of shape $Y(\lambda)$. In this case, the image $f(V)$ is $\mathcal{A}\mathbf{E}_T$, which is isomorphic to the Specht module $\mathcal{S}^{Y(\lambda)}$ (by (228)) and therefore a free $\mathbf{k}$-module of rank $f^\lambda$ (by Lemma 5.9.17). The equality

$f^2 = \kappa f$ will follow easily from the equality $\mathbf{E}_T^2 = \kappa \mathbf{E}_T$ that we have obtained above as a consequence of Proposition 5.11.5. Thus, Lemma 5.11.11 will be applicable and allow us to compute $\kappa$. Indeed, Lemma 5.11.11 is tailor-made for this application, but has also found uses in other representation-theoretical questions (see the "Motivation" in [Ladisc11]).

First, however, we must prove Lemma 5.11.11:

*Proof of Lemma 5.11.11.* The following beautiful proof is due to Frieder Ladisch ([Ladisc11]).

Clearly, $f(v) \in f(V)$ for each $v \in V$. Thus, we can define a map $h : V \to f(V)$ that sends each $v \in V$ to $f(v)$. Consider this map $h$. (This map $h$ is just the map $f$, except that we have replaced its target by $f(V)$.) Note that the map $h$ is $\mathbf{k}$-linear (since $f$ is $\mathbf{k}$-linear).

Let $g : f(V) \to V$ be the inclusion map of the $\mathbf{k}$-submodule $f(V)$ into the $\mathbf{k}$-module $V$. This is the map sending each $w \in f(V)$ to $w$ itself. Clearly, this map $g$ is $\mathbf{k}$-linear.

From $f^2 = \kappa f$, we easily obtain $h \circ g = \kappa \operatorname{id}_{f(V)}$.

[*Proof:* Let $w \in f(V)$. Thus, $w = f(v)$ for some $v \in V$. Consider this $v$. But $g$ is the inclusion map, so that $g(w) = w = f(v)$. Now,

$$(h \circ g)(w) = h\left(\underbrace{g(w)}_{=f(v)}\right) = h(f(v)) = f(f(v)) \qquad \text{(by the definition of } h\text{)}$$
$$= \underbrace{f^2}_{=\kappa f}(v) = (\kappa f)(v) = \kappa \underbrace{f(v)}_{=w} = \kappa w.$$

Forget that we fixed $w$. We thus have proved that $(h \circ g)(w) = \kappa w$ for each $w \in f(V)$. In other words, $h \circ g = \kappa \operatorname{id}_{f(V)}$.]

On the other hand, we have $g \circ h = f$.

[*Proof:* For each $v \in V$, we have

$$(g \circ h)(v) = g(h(v)) = h(v) \qquad \text{(since } g \text{ is the inclusion map)}$$
$$= f(v) \qquad \text{(by the definition of } h\text{)}.$$

In other words, $g \circ h = f$.]

We assumed that the $\mathbf{k}$-module $f(V)$ is free of rank $d$. Thus, Lemma 5.11.10 (applied to $W = f(V)$) yields $\operatorname{Tr}\left(\kappa \operatorname{id}_{f(V)}\right) = d\kappa$. Since $h \circ g = \kappa \operatorname{id}_{f(V)}$, we can rewrite this as

$$\operatorname{Tr}(h \circ g) = d\kappa. \tag{232}$$

But Lemma 5.11.9 (applied to $f(V)$ and $h$ instead of $W$ and $f$) yields $\operatorname{Tr}(h \circ g) = \operatorname{Tr}(g \circ h) = \operatorname{Tr} f$ (since $g \circ h = f$). Comparing this with (232), we obtain $\operatorname{Tr} f = d\kappa$. This proves Lemma 5.11.11. $\square$

Combining this lemma with Proposition 5.11.7, we obtain the following consequence for group algebras:

**Lemma 5.11.12.** Let $G$ be a finite group. Let $\mathcal{A} = \mathbf{k}[G]$. (In this, we break with Definition 5.11.1 **(c)** in this lemma.)

Let $\mathbf{a} \in \mathcal{A}$ and $\kappa \in \mathbf{k}$ be two elements such that $\mathbf{a}^2 = \kappa\mathbf{a}$. Assume furthermore that the left ideal $\mathcal{A}\mathbf{a}$ of $\mathcal{A}$ is a free $\mathbf{k}$-module of rank $d$ for some $d \in \mathbb{N}$. Then, in $\mathbf{k}$, we have

$$d\kappa = |G| \cdot [1]\,\mathbf{a}$$

(where 1 means the neutral element of $G$).

*Proof.* We have $\mathbf{a} \in \mathcal{A} = \mathbf{k}[G]$. Let $f_{\mathbf{a}} : \mathbf{k}[G] \to \mathbf{k}[G]$ be the $\mathbf{k}$-linear map that sends each $\mathbf{b} \in \mathbf{k}[G]$ to $\mathbf{ba}$. Then, Proposition 5.11.7 yields

$$\operatorname{Tr}(f_{\mathbf{a}}) = |G| \cdot [1]\,\mathbf{a}. \tag{233}$$

However, the $\mathbf{k}$-linear map $f_{\mathbf{a}}$ satisfies $f_{\mathbf{a}}^2 = \kappa f_{\mathbf{a}}$ (since each $\mathbf{b} \in \mathbf{k}[G]$ satisfies

$$f_{\mathbf{a}}^2(\mathbf{b}) = f_{\mathbf{a}}(f_{\mathbf{a}}(\mathbf{b})) = \underbrace{f_{\mathbf{a}}(\mathbf{b})}_{\substack{=\mathbf{ba} \\ \text{(by the definition} \\ \text{of } f_{\mathbf{a}}\text{)}}} \cdot \mathbf{a} \qquad \text{(by the definition of } f_{\mathbf{a}}\text{)}$$

$$= \mathbf{b}\,\underbrace{\mathbf{a}\cdot\mathbf{a}}_{=\mathbf{a}^2=\kappa\mathbf{a}} = \mathbf{b}\kappa\mathbf{a} = \kappa\,\underbrace{\mathbf{ba}}_{=f_{\mathbf{a}}(\mathbf{b})} = \kappa f_{\mathbf{a}}(\mathbf{b})$$

). Moreover, its image is

$$f_{\mathbf{a}}(\mathcal{A}) = \left\{ \underbrace{f_{\mathbf{a}}(\mathbf{b})}_{=\mathbf{ba}} \mid \mathbf{b} \in \mathcal{A} \right\} = \{\mathbf{ba} \mid \mathbf{b} \in \mathcal{A}\} = \mathcal{A}\mathbf{a},$$

and thus is a free $\mathbf{k}$-module of rank $d$ (since we assumed that $\mathcal{A}\mathbf{a}$ is a free $\mathbf{k}$-module of rank $d$). Hence, Lemma 5.11.11 (applied to $V = \mathcal{A}$ and $f = f_{\mathbf{a}}$) yields

$$\operatorname{Tr}(f_{\mathbf{a}}) = d\kappa.$$

Comparing this with (233), we find

$$d\kappa = |G| \cdot [1]\,\mathbf{a}.$$

This proves Lemma 5.11.12. □

**Lemma 5.11.13.** Let $T$ be an $n$-tableau of any shape. Then, $[1]\,\mathbf{E}_T = 1$.

*Proof.* Recall that

$$\mathbf{E}_T = \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} = \left( \sum_{c \in \mathcal{C}(T)} (-1)^c c \right) \left( \sum_{r \in \mathcal{R}(T)} r \right)$$

$$\left( \begin{array}{c} \text{since Definition 5.5.1 yields } \nabla_{\operatorname{Row} T} = \sum\limits_{w \in \mathcal{R}(T)} w = \sum\limits_{r \in \mathcal{R}(T)} r \\ \text{and } \nabla^-_{\operatorname{Col} T} = \sum\limits_{w \in \mathcal{C}(T)} (-1)^w w = \sum\limits_{c \in \mathcal{C}(T)} (-1)^c c \end{array} \right)$$

$$= \sum_{c \in \mathcal{C}(T)} \sum_{r \in \mathcal{R}(T)} (-1)^c cr = \sum_{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T)} (-1)^c cr. \tag{234}$$

Thus,

$$[\mathrm{id}]\, \mathbf{E}_T = \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr = \mathrm{id}}} (-1)^c cr = (-1)^{\mathrm{id}}\, \mathrm{id} \cdot \mathrm{id},$$

because Proposition 5.3.5 **(b)** shows that the only pair $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$ satisfying $cr = \mathrm{id}$ is the pair $(\mathrm{id}, \mathrm{id})$. Since we denote id by 1, we can rewrite this as

$$[1]\, \mathbf{E}_T = \underbrace{(-1)^{\mathrm{id}}}_{=1} \underbrace{\mathrm{id}}_{=1} \cdot \underbrace{\mathrm{id}}_{=1} = 1.$$

This proves Lemma 5.11.13. □

*Proof of Theorem 5.11.3.* Let us first prove the theorem for $\mathbf{k} = \mathbb{Z}$.

So let us assume that $\mathbf{k} = \mathbb{Z}$. Proposition 5.11.5 (applied to $\mathbf{a} = 1$) yields $\mathbf{E}_T 1 \mathbf{E}_T = \kappa \mathbf{E}_T$ for some scalar $\kappa \in \mathbf{k}$. Consider this $\kappa$. Thus, $\kappa \in \mathbf{k} = \mathbb{Z}$. Of course, $\mathbf{E}_T 1 \mathbf{E}_T = \mathbf{E}_T \mathbf{E}_T = \mathbf{E}_T^2$, so that

$$\mathbf{E}_T^2 = \mathbf{E}_T 1 \mathbf{E}_T = \kappa \mathbf{E}_T. \tag{235}$$

It remains to show that $\kappa = \dfrac{n!}{f^\lambda}$.

Let $\mathcal{A} = \mathbf{k}[S_n]$. Lemma 5.11.13 yields $[1]\, \mathbf{E}_T = 1 \neq 0$. Hence, the element $\mathbf{E}_T$ is nonzero. Now, the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_T$ contains the element $1\mathbf{E}_T = \mathbf{E}_T$, which is nonzero (as we have just shown). Hence, $\mathcal{A}\mathbf{E}_T$ itself is nonzero.

But $T$ is an $n$-tableau of shape $Y(\lambda)$. Hence, (228) (applied to $D = Y(\lambda)$) yields $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$.

Let $\mathcal{S}^\lambda$ be the Specht module $\mathcal{S}^{Y(\lambda)}$. Then, $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$. But Lemma 5.9.17 shows that $\mathcal{S}^\lambda$ is a free $\mathbf{k}$-module of rank $f^\lambda$. Therefore, $\mathcal{A}\mathbf{E}_T$ is a free $\mathbf{k}$-module of rank $f^\lambda$ (since $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T$). Since the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_T$ is nonzero (as we showed above), this entails that its rank $f^\lambda$ is nonzero.

Now, recall that $\mathbf{E}_T^2 = \kappa \mathbf{E}_T$. Hence, Lemma 5.11.12 (applied to $G = S_n$ and $\mathbf{a} = \mathbf{E}_T$ and $d = f^\lambda$) yields

$$f^\lambda \kappa = \underbrace{|S_n|}_{=n!} \cdot \underbrace{[1]\, \mathbf{E}_T}_{=1} = n!.$$

Hence,

$$\kappa = \frac{n!}{f^\lambda} \qquad \left(\text{since } f^\lambda \text{ is nonzero}\right).$$

Thus, $\dfrac{n!}{f^\lambda} \in \mathbb{Z}$ (since $\kappa \in \mathbb{Z}$). Hence, $\dfrac{n!}{f^\lambda}$ is a positive integer (since its positivity is obvious). Moreover, in view of $\kappa = \dfrac{n!}{f^\lambda}$, we can rewrite (235) as

$$\mathbf{E}_T^2 = \frac{n!}{f^\lambda}\mathbf{E}_T. \tag{236}$$

Thus, we have proved Theorem 5.11.3 in the case when $\mathbf{k} = \mathbb{Z}$. It remains to prove the theorem in the general case, i.e., for an arbitrary commutative ring $\mathbf{k}$.

But this is just an easy application of base change: Drop our assumption that $\mathbf{k} = \mathbb{Z}$, and consider an arbitrary commutative ring $\mathbf{k}$ instead. Then, there is a canonical ring morphism $f : \mathbb{Z} \to \mathbf{k}$ (which sends each integer $m$ to $m \cdot 1_\mathbf{k}$). This ring morphism $f$ induces a base change morphism $f_* : \mathbb{Z}[S_n] \to \mathbf{k}[S_n]$ (as defined in Definition 3.12.22), and this latter morphism $f_*$ sends the Young symmetrizer $\mathbf{E}_T \in \mathbb{Z}[S_n]$ to the corresponding Young symmetrizer $\mathbf{E}_T \in \mathbf{k}[S_n]$ (since the definition of $\mathbf{E}_T$ is independent of the base ring[92]). Hence, applying this morphism $f_*$ to both sides of the equality (236), we obtain the analogous equality

$$\mathbf{E}_T^2 = \frac{n!}{f^\lambda}\mathbf{E}_T \qquad \text{in } \mathbf{k}[S_n]$$

(since $f_*$ is a ring morphism and thus respects squares and scaling). This completes the proof of Theorem 5.11.3 in the general case (since we have already shown that $\dfrac{n!}{f^\lambda}$ is a positive integer). □

We observe one more easy property of Young symmetrizers:

**Lemma 5.11.14.** Let $T$ be an $n$-tableau. Let $w \in S_n$. Then,

$$\mathbf{E}_{wT} = w\mathbf{E}_T w^{-1}.$$

[92]To make this more concrete: As we saw in the proof of Lemma 5.11.13, we have

$$\mathbf{E}_T = \left(\sum_{c \in \mathcal{C}(T)} (-1)^c c\right)\left(\sum_{r \in \mathcal{R}(T)} r\right).$$

This equality holds both in $\mathbb{Z}[S_n]$ and in $\mathbf{k}[S_n]$. But the base change morphism $f_* : \mathbb{Z}[S_n] \to \mathbf{k}[S_n]$ sends the element $\left(\sum_{c \in \mathcal{C}(T)} (-1)^c c\right)\left(\sum_{r \in \mathcal{R}(T)} r\right)$ of $\mathbb{Z}[S_n]$ to the analogous element of $\mathbf{k}[S_n]$. Thus, it sends the Young symmetrizer $\mathbf{E}_T \in \mathbb{Z}[S_n]$ to the corresponding Young symmetrizer $\mathbf{E}_T \in \mathbf{k}[S_n]$.

*Proof.* We have $wT = w \rightharpoonup T$. Thus,

$$\begin{aligned}\mathbf{E}_{wT} &= \mathbf{E}_{w\rightharpoonup T} \\ &= \underbrace{\nabla^{-}_{\mathrm{Col}(w\rightharpoonup T)}}_{\substack{=w\nabla^{-}_{\mathrm{Col}\,T}w^{-1} \\ \text{(by Proposition 5.5.11)}}} \underbrace{\nabla_{\mathrm{Row}(w\rightharpoonup T)}}_{\substack{=w\nabla_{\mathrm{Row}\,T}w^{-1} \\ \text{(by Proposition 5.5.11)}}} \qquad \text{(by the definition of } \mathbf{E}_{w\rightharpoonup T}) \\ &= w\nabla^{-}_{\mathrm{Col}\,T}\underbrace{w^{-1}w}_{=1}\nabla_{\mathrm{Row}\,T}w^{-1} = w \underbrace{\nabla^{-}_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}}_{\substack{=\mathbf{E}_T \\ \text{(by the definition of } \mathbf{E}_T)}} w^{-1} = w\mathbf{E}_T w^{-1}.\end{aligned}$$

Thus, Lemma 5.11.14 is proved. □

Lemma 5.11.14 shows that if you know the Young symmetrizer $\mathbf{E}_T$ for an $n$-tableau $T$ of a given shape, then you can easily obtain the $\mathbf{E}_S$'s for all other $n$-tableaux $S$ of the same shape.

### 5.11.3. Orthogonality

Now let us see what happens when we multiply two Young symmetrizers corresponding to different tableaux. First, we deal with the easy case where the tableaux have different shapes:

**Proposition 5.11.15** (double zero-sandwich lemma)**.** Let $\lambda$ and $\mu$ be two distinct partitions of $n$. Let $S$ be an $n$-tableau of shape $Y(\lambda)$. Let $T$ be an $n$-tableau of shape $Y(\mu)$. Let $\mathbf{a} \in \mathbf{k}[S_n]$. Then,

$$\mathbf{E}_S \mathbf{a} \mathbf{E}_T = 0.$$

*Proof.* By the definition of a Young symmetrizer, we have $\mathbf{E}_S = \nabla^{-}_{\mathrm{Col}\,S}\nabla_{\mathrm{Row}\,S}$ and $\mathbf{E}_T = \nabla^{-}_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}$. Thus,

$$\mathbf{E}_S \mathbf{a} \mathbf{E}_T = \nabla^{-}_{\mathrm{Col}\,S}\nabla_{\mathrm{Row}\,S}\mathbf{a}\nabla^{-}_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}. \tag{237}$$

But we have $\lambda \neq \mu$. Hence, there exists some $i \geq 1$ such that $\lambda_i \neq \mu_i$. Consider the **smallest** such $i$. Then, $\lambda_k = \mu_k$ for all $k \in [i-1]$ (since $i$ is the smallest). Summing up all these $i-1$ equalities, we obtain

$$\lambda_1 + \lambda_2 + \cdots + \lambda_{i-1} = \mu_1 + \mu_2 + \cdots + \mu_{i-1}. \tag{238}$$

However, $\lambda_i \neq \mu_i$. Thus, we are in one of the following two cases:
*Case 1:* We have $\lambda_i > \mu_i$.
*Case 2:* We have $\lambda_i < \mu_i$.
Let us consider Case 1 first. In this case, we have $\lambda_i > \mu_i$. Adding this inequality to the equality (238), we obtain

$$(\lambda_1 + \lambda_2 + \cdots + \lambda_{i-1}) + \lambda_i > (\mu_1 + \mu_2 + \cdots + \mu_{i-1}) + \mu_i.$$

In other words,

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i > \mu_1 + \mu_2 + \cdots + \mu_i.$$

Hence, Theorem 5.10.8 **(d)** yields $\nabla_{\operatorname{Row} S} \mathbf{a} \nabla^-_{\operatorname{Col} T} = 0$. And thus, (237) becomes

$$\mathbf{E}_S \mathbf{a} \mathbf{E}_T = \nabla^-_{\operatorname{Col} S} \underbrace{\nabla_{\operatorname{Row} S} \mathbf{a} \nabla^-_{\operatorname{Col} T}}_{=0} \nabla_{\operatorname{Row} T} = 0.$$

This proves Proposition 5.11.15 in Case 1.

Let us now consider Case 2. In this case, we have $\lambda_i < \mu_i$. Adding this inequality to the equality (238), we obtain

$$(\lambda_1 + \lambda_2 + \cdots + \lambda_{i-1}) + \lambda_i < (\mu_1 + \mu_2 + \cdots + \mu_{i-1}) + \mu_i.$$

In other words,

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i < \mu_1 + \mu_2 + \cdots + \mu_i.$$

In other words,

$$\mu_1 + \mu_2 + \cdots + \mu_i > \lambda_1 + \lambda_2 + \cdots + \lambda_i.$$

Thus, Theorem 5.10.8 **(b)** (applied to $\mu$, $\lambda$, $T$, $S$ and $\nabla_{\operatorname{Row} S} \mathbf{a} \nabla^-_{\operatorname{Col} T}$ instead of $\lambda$, $\mu$, $S$, $T$ and $\mathbf{a}$) yields

$$\nabla^-_{\operatorname{Col} S} \nabla_{\operatorname{Row} S} \mathbf{a} \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} = 0.$$

In view of (237), we can rewrite this as $\mathbf{E}_S \mathbf{a} \mathbf{E}_T = 0$. This proves Proposition 5.11.15 in Case 2.

Hence, Proposition 5.11.15 is proved in both Cases 1 and 2. □

What happens when you multiply two Young symmetrizers $\mathbf{E}_S$ and $\mathbf{E}_T$ of the same shape (but possibly different tableaux)? The following proposition gives an answer:

**Proposition 5.11.16.** Let $\lambda$ be a partition of $n$. Let $S$ and $T$ be two $n$-tableaux of shape $Y(\lambda)$. Let $f^\lambda$ be the # of standard tableaux of shape $Y(\lambda)$. Then:

**(a)** If there exist no $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$, then $\mathbf{E}_S \mathbf{E}_T = 0$.

**(b)** If there exist $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$, then these $r$ and $c$ satisfy $\mathbf{E}_S \mathbf{E}_T = h^\lambda (-1)^c r^{-1} \mathbf{E}_U c$, where we set $U := rS = cT$ and $h^\lambda := \dfrac{n!}{f^\lambda}$.

*Proof.* **(a)** Assume that there exist no $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$.

If there were no two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$, then Theorem 5.10.1 **(c)** (applied to $\mu = \lambda$) would show that there exist permutations $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$; but this would contradict the preceding sentence. Thus, there are two distinct integers that lie in the same row of $S$ and simultaneously lie in the same column of $T$. Hence, Lemma 5.10.5 yields that $\nabla_{\operatorname{Row} S} \nabla^{-}_{\operatorname{Col} T} = 0$. But

$$\underbrace{\mathbf{E}_S}_{\substack{=\nabla^{-}_{\operatorname{Col} S}\nabla_{\operatorname{Row} S} \\ \text{(by the definition} \\ \text{of } \mathbf{E}_S)}} \underbrace{\mathbf{E}_T}_{\substack{=\nabla^{-}_{\operatorname{Col} T}\nabla_{\operatorname{Row} T} \\ \text{(by the definition} \\ \text{of } \mathbf{E}_T)}} = \nabla^{-}_{\operatorname{Col} S} \underbrace{\nabla_{\operatorname{Row} S} \nabla^{-}_{\operatorname{Col} T}}_{=0} \nabla_{\operatorname{Row} T} = 0.$$

This proves Proposition 5.11.16 **(a)**.

**(b)** Let $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ be such that $rS = cT$. Set $U := rS = cT$ and $h^\lambda := \dfrac{n!}{f^\lambda}$. We must prove that $\mathbf{E}_S \mathbf{E}_T = h^\lambda (-1)^c r^{-1} \mathbf{E}_U c$.

Theorem 5.11.3 (applied to $U$ instead of $T$) yields

$$\mathbf{E}_U^2 = \underbrace{\frac{n!}{f^\lambda}}_{=h^\lambda} \mathbf{E}_U = h^\lambda \mathbf{E}_U. \tag{239}$$

Proposition 5.5.5 **(a)** (applied to $S$ and $r$ instead of $T$ and $w$) yields $r \nabla_{\operatorname{Row} S} = \nabla_{\operatorname{Row} S} r = \nabla_{\operatorname{Row} S}$ (since $r \in \mathcal{R}(S)$). Hence,

$$\underbrace{\mathbf{E}_S}_{\substack{=\nabla^{-}_{\operatorname{Col} S}\nabla_{\operatorname{Row} S} \\ \text{(by the definition} \\ \text{of } \mathbf{E}_S)}} r = \nabla^{-}_{\operatorname{Col} S} \underbrace{\nabla_{\operatorname{Row} S} r}_{=\nabla_{\operatorname{Row} S}} = \nabla^{-}_{\operatorname{Col} S} \nabla_{\operatorname{Row} S} = \mathbf{E}_S.$$

Proposition 5.5.5 **(b)** (applied to $w = c$) yields $c \nabla^{-}_{\operatorname{Col} T} = \nabla^{-}_{\operatorname{Col} T} c = (-1)^c \nabla^{-}_{\operatorname{Col} T}$ (since $c \in \mathcal{C}(T)$). Hence,

$$c \underbrace{\mathbf{E}_T}_{\substack{=\nabla^{-}_{\operatorname{Col} T}\nabla_{\operatorname{Row} T} \\ \text{(by the definition} \\ \text{of } \mathbf{E}_T)}} = \underbrace{c \nabla^{-}_{\operatorname{Col} T}}_{=(-1)^c \nabla^{-}_{\operatorname{Col} T}} \nabla_{\operatorname{Row} T} = (-1)^c \underbrace{\nabla^{-}_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}}_{=\mathbf{E}_T} = (-1)^c \mathbf{E}_T.$$

But $U = rS$ and thus

$$\begin{aligned} \mathbf{E}_U = \mathbf{E}_{rS} &= r \underbrace{\mathbf{E}_S r^{-1}}_{\substack{=\mathbf{E}_S \\ \text{(since } \mathbf{E}_S r = \mathbf{E}_S\text{)}}} \qquad \text{(by Lemma 5.11.14)} \\ &= r \mathbf{E}_S, \end{aligned}$$

so that $\mathbf{E}_S = r^{-1}\mathbf{E}_U$. Furthermore, $U = cT$ and thus

$$\begin{aligned}\mathbf{E}_U = \mathbf{E}_{cT} &= \underbrace{c\mathbf{E}_T}_{=(-1)^c\mathbf{E}_T}\, c^{-1} \qquad \text{(by Lemma 5.11.14)} \\ &= (-1)^c\, \mathbf{E}_T c^{-1},\end{aligned}$$

so that $\mathbf{E}_T = \underbrace{\frac{1}{(-1)^c}}_{=(-1)^c}\, \mathbf{E}_U c = (-1)^c\, \mathbf{E}_U c$. Thus,

$$\underbrace{\mathbf{E}_S}_{=r^{-1}\mathbf{E}_U}\ \underbrace{\mathbf{E}_T}_{=(-1)^c\mathbf{E}_U c} = r^{-1}\mathbf{E}_U\,(-1)^c\,\mathbf{E}_U c = (-1)^c\, r^{-1}\ \underbrace{\mathbf{E}_U^2}_{\substack{=h^\lambda \mathbf{E}_U \\ \text{(by (239))}}}\ c = h^\lambda\,(-1)^c\, r^{-1}\mathbf{E}_U c.$$

This proves Proposition 5.11.16 **(b)**. □

A formula for $\mathbf{E}_S \mathbf{a} \mathbf{E}_T$ (for arbitrary $\mathbf{a} \in \mathbf{k}[S_n]$) can also be found, but it is less pleasant and no stronger than Proposition 5.11.16 (as it follows from the latter proposition using $\mathbf{E}_{wT} = w\mathbf{E}_T w^{-1}$).

**Example 5.11.17.** **(a)** Let us take $n = 4$ and $\lambda = (2,2)$ and $S = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \hline \end{array}$ and $T = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array}$. Then, there exist no $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$. Hence, Proposition 5.11.16 **(a)** yields $\mathbf{E}_S \mathbf{E}_T = 0$. A similar argument shows that $\mathbf{E}_T \mathbf{E}_S = 0$.

**(b)** Let us instead take $n = 5$ and $\lambda = (3,2)$ and $S = \begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 \\ \cline{1-2} \end{array}$ and $T = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2} \end{array}$. Then, there exist $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$ (namely, $r = t_{3,5}t_{2,4}$ and $c = t_{2,5}$). Hence, Proposition 5.11.16 **(b)** yields $\mathbf{E}_S \mathbf{E}_T = h^\lambda\,(-1)^c\, r^{-1}\mathbf{E}_U c$, where $U := rS = cT = \begin{array}{|c|c|c|}\hline 1 & 5 & 3 \\ \hline 4 & 2 \\ \cline{1-2} \end{array}$ and $h^\lambda := \dfrac{n!}{f^\lambda} = \dfrac{5!}{5} = 24$.

**Exercise 5.11.2.** **(a)** [4] Let $\lambda$ be a hook partition (see Definition 5.1.9). Let $S$ and $T$ be two distinct standard $n$-tableaux of shape $Y(\lambda)$. Prove that $\mathbf{E}_S \mathbf{E}_T = 0$.

**(b)** $\boxed{2}$ One might wonder what other partitions $\lambda$ have the property stated in part **(a)**. For example, $\lambda = (2,2)$ does (as we saw in Example 5.11.17 **(a)**), but $\lambda = (3,2)$ does not (as we saw in Example 5.11.17 **(b)**).

Show that if $\mathbf{k}$ is a field of characteristic 0, then the only partitions $\lambda$ that have the property stated in part **(a)** are the hook partitions and the partitions $()$ and $(2,2)$.

Proposition 5.11.16 is nice (note how $S$ and $T$ appear in symmetric roles in it), but we will eventually need a less symmetric (and more "$T$-biased") form as well:

**Proposition 5.11.18.** Let $\lambda$ be a partition of $n$. Let $S$ and $T$ be two $n$-tableaux of shape $Y(\lambda)$. Let $f^\lambda$ be the # of standard tableaux of shape $Y(\lambda)$. Then:

**(a)** If there exist no $\widetilde{r} \in \mathcal{R}(T)$ and $\widetilde{c} \in \mathcal{C}(T)$ such that $S = \widetilde{c}\widetilde{r}T$, then $\mathbf{E}_S \mathbf{E}_T = 0$.

**(b)** If there exist $\widetilde{r} \in \mathcal{R}(T)$ and $\widetilde{c} \in \mathcal{C}(T)$ such that $S = \widetilde{c}\widetilde{r}T$, then these $\widetilde{r}$ and $\widetilde{c}$ satisfy $\mathbf{E}_S \mathbf{E}_T = h^\lambda (-1)^{\widetilde{c}} \widetilde{c}\widetilde{r}\mathbf{E}_T$, where we set $h^\lambda := \dfrac{n!}{f^\lambda}$.

*Proof.* We shall derive this from Proposition 5.11.16.

First, we observe that each permutation $c \in S_n$ satisfies $cT = c \rightharpoonup T$ and thus

$$\mathcal{R}(cT) = \mathcal{R}(c \rightharpoonup T) = c\mathcal{R}(T)c^{-1} \tag{240}$$

(by Proposition 5.3.12 **(c)**, applied to $w = c$).

**(a)** Assume that there exist no $\widetilde{r} \in \mathcal{R}(T)$ and $\widetilde{c} \in \mathcal{C}(T)$ such that $S = \widetilde{c}\widetilde{r}T$. Then, it is not hard to see that there exist no $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$.

[*Proof:* Assume the contrary. Thus, there exist some $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$. Consider these $r$ and $c$. The $n$-tableau $cT = rS = r \rightharpoonup S$ is row-equivalent to $S$ (by Proposition 5.3.12 **(a)**, since $r \in \mathcal{R}(S)$), and thus we have $\mathcal{R}(cT) = \mathcal{R}(S)$ (by Proposition 5.3.4 **(a)**). Hence, $\mathcal{R}(S) = \mathcal{R}(cT) = c\mathcal{R}(T)c^{-1}$ (by (240)). Thus, $r \in \mathcal{R}(S) = c\mathcal{R}(T)c^{-1}$. In other words, $r = cr'c^{-1}$ for some $r' \in \mathcal{R}(T)$. Consider this $r'$. From $r = cr'c^{-1}$, we obtain $r^{-1} = \left(cr'c^{-1}\right)^{-1} = c(r')^{-1}c^{-1}$ and therefore $c^{-1}r^{-1}c = (r')^{-1} \in \mathcal{R}(T)$ (since $r' \in \mathcal{R}(T)$, but $\mathcal{R}(T)$ is a group).

However $rS = cT$ entails $S = r^{-1}cT = c\underbrace{c^{-1}r^{-1}c}_{=(r')^{-1}}T = c(r')^{-1}T$. Thus, there exist $\widetilde{r} \in \mathcal{R}(T)$ and $\widetilde{c} \in \mathcal{C}(T)$ such that $S = \widetilde{c}\widetilde{r}T$ (namely, $\widetilde{c} = c$ and $\widetilde{r} = (r')^{-1}$). But we assumed that there exist no such $\widetilde{r}$ and $\widetilde{c}$. Contradiction! This completes our proof.]

So we have shown that there exist no $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$. Thus, Proposition 5.11.16 **(a)** yields $\mathbf{E}_S \mathbf{E}_T = 0$. This proves Proposition 5.11.18 **(a)**.

**(b)** Assume that there exist $\widetilde{r} \in \mathcal{R}(T)$ and $\widetilde{c} \in \mathcal{C}(T)$ such that $S = \widetilde{c}\widetilde{r}T$. Consider these $\widetilde{r}$ and $\widetilde{c}$. Set $h^\lambda := \dfrac{n!}{f^\lambda}$. We must show that $\mathbf{E}_S \mathbf{E}_T = h^\lambda (-1)^{\widetilde{c}} \widetilde{c}\widetilde{r}^{-1} \mathbf{E}_T$.

Set $c := \widetilde{c}$ and $r := \widetilde{c}\widetilde{r}^{-1}\widetilde{c}^{-1}$. We have $\widetilde{r} \in \mathcal{R}(T)$ and therefore $\widetilde{r}^{-1} \in \mathcal{R}(T)$ (since $\mathcal{R}(T)$ is a group). Furthermore,

$$r = \underbrace{\widetilde{c}}_{=c} \widetilde{r}^{-1} \underbrace{\widetilde{c}^{-1}}_{\substack{=c^{-1} \\ (\text{since } \widetilde{c}=c)}} = c \underbrace{\widetilde{r}^{-1}}_{\in \mathcal{R}(T)} c^{-1} \in c\mathcal{R}(T) c^{-1} = \mathcal{R}(cT) \qquad (\text{by (240)}).$$

Thus, $r^{-1} \in \mathcal{R}(cT)$ (since $\mathcal{R}(cT)$ is a group). Also,

$$\underbrace{r}_{=\widetilde{c}\widetilde{r}^{-1}\widetilde{c}^{-1}} S = \underbrace{\widetilde{c}}_{=c} \widetilde{r}^{-1} \underbrace{\widetilde{c}^{-1} S}_{\substack{=\widetilde{r}T \\ (\text{since } S=\widetilde{c}\widetilde{r}T)}} = c \underbrace{\widetilde{r}^{-1}\widetilde{r}}_{=1} T = cT.$$

Hence, $S = r^{-1}cT = r^{-1} \rightharpoonup cT$. But the $n$-tableau $r^{-1} \rightharpoonup cT$ is row-equivalent to $cT$ (by Proposition 5.3.12 **(a)**, since $r^{-1} \in \mathcal{R}(cT)$). In other words, $S$ is row-equivalent to $cT$ (since $S = r^{-1} \rightharpoonup cT$). Thus, $\mathcal{R}(S) = \mathcal{R}(cT)$ (by Proposition 5.3.4 **(a)**). Hence, $r \in \mathcal{R}(S)$ (since $r \in \mathcal{R}(cT)$).

Altogether, we have now shown that $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ (since $c = \widetilde{c} \in \mathcal{C}(T)$) and $rS = cT$. Hence, Proposition 5.11.16 **(b)** shows that $\mathbf{E}_S \mathbf{E}_T = h^\lambda (-1)^c r^{-1} \mathbf{E}_U c$, where we set $U := rS = cT$.

From $U = cT$, we obtain $\mathbf{E}_U = \mathbf{E}_{cT} = c\mathbf{E}_T c^{-1}$ (by Lemma 5.11.14), so that $\mathbf{E}_U c = c\mathbf{E}_T$. Hence,

$$\begin{aligned}
\mathbf{E}_S \mathbf{E}_T &= h^\lambda (-1)^c \underbrace{r^{-1}}_{\substack{=\left(c\widetilde{r}^{-1}c^{-1}\right)^{-1} \\ (\text{since } r=c\widetilde{r}^{-1}c^{-1})}} \underbrace{\mathbf{E}_U c}_{=c\mathbf{E}_T} = h^\lambda (-1)^c \underbrace{\left(c\widetilde{r}^{-1}c^{-1}\right)^{-1}}_{=c\widetilde{r}c^{-1}} c\mathbf{E}_T \\
&= h^\lambda (-1)^c c\widetilde{r} \underbrace{c^{-1}c}_{=1} \mathbf{E}_T = h^\lambda (-1)^c c\widetilde{r}\mathbf{E}_T = h^\lambda (-1)^{\widetilde{c}} \widetilde{c}\widetilde{r}\mathbf{E}_T \qquad (\text{since } c = \widetilde{c}).
\end{aligned}$$

This proves Proposition 5.11.18 **(b)**. □

The formulas shown in this subsection answer the question "what is the product of two Young symmetrizers". Roughly speaking, the answer is "0 in most cases, or some modified version of a Young symmetrizer in the remaining cases".

### 5.11.4. The opposite Young symmetrizer

The Young symmetrizer $\mathbf{E}_T$ was defined to be the product $\nabla^-_{\mathrm{Col}\,T} \nabla_{\mathrm{Row}\,T}$. What can be said about the opposite product, $\nabla_{\mathrm{Row}\,T} \nabla^-_{\mathrm{Col}\,T}$ ? A few things, as it turns out:

**Proposition 5.11.19.** Let $\lambda$ be a partition. Let $T$ be an $n$-tableau of shape $\lambda$. Let

$$\mathbf{F}_T := \nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T} \in \mathcal{A}.$$

Then:

**(a)** The antipode $S$ of $\mathbf{k}[S_n]$ (which was defined in Definition 3.11.12) satisfies $S(\mathbf{E}_T) = \mathbf{F}_T$.

**(b)** We have $\mathbf{F}_T^2 = \dfrac{n!}{f^\lambda} \mathbf{F}_T$.

**(c)** Let $f^\lambda$ denote the # of standard tableaux of shape $Y(\lambda)$. Let $h^\lambda := \dfrac{n!}{f^\lambda}$. If $h^\lambda$ is invertible in $\mathbf{k}$, then $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T \cong \mathcal{A}\mathbf{F}_T$ as left $\mathcal{A}$-modules.

*Proof.* **(a)** The antipode $S$ is a $\mathbf{k}$-algebra anti-automorphism (by Theorem 3.11.14 **(a)**). Thus,

$$S\left(\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}\right) = \underbrace{S\left(\nabla_{\operatorname{Row} T}\right)}_{\substack{=\nabla_{\operatorname{Row} T} \\ \text{(by Proposition 5.5.10)}}} \underbrace{S\left(\nabla^-_{\operatorname{Col} T}\right)}_{\substack{=\nabla^-_{\operatorname{Col} T} \\ \text{(by Proposition 5.5.10)}}} = \nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T} = \mathbf{F}_T$$

(by the definition of $\mathbf{F}_T$). In view of $\mathbf{E}_T = \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$, we can rewrite this as $S(\mathbf{E}_T) = \mathbf{F}_T$. This proves Proposition 5.11.19 **(a)**.

**(b)** The antipode $S$ is a $\mathbf{k}$-algebra anti-automorphism (by Theorem 3.11.14 **(a)**). Thus, $S(\mathbf{E}_T \mathbf{E}_T) = S(\mathbf{E}_T) S(\mathbf{E}_T) = (S(\mathbf{E}_T))^2$. In other words, $S\left(\mathbf{E}_T^2\right) = (S(\mathbf{E}_T))^2$ (since $\mathbf{E}_T \mathbf{E}_T = \mathbf{E}_T^2$).

But Theorem 5.11.3 yields $\mathbf{E}_T^2 = \dfrac{n!}{f^\lambda} \mathbf{E}_T$. Applying the map $S$ to both sides of this equality, we obtain

$$S\left(\mathbf{E}_T^2\right) = S\left(\frac{n!}{f^\lambda} \mathbf{E}_T\right) = \frac{n!}{f^\lambda} S(\mathbf{E}_T) \qquad \text{(since the map } S \text{ is } \mathbf{k}\text{-linear)}.$$

Comparing this with $S\left(\mathbf{E}_T^2\right) = (S(\mathbf{E}_T))^2$, we find

$$(S(\mathbf{E}_T))^2 = \frac{n!}{f^\lambda} S(\mathbf{E}_T).$$

Since $S(\mathbf{E}_T) = \mathbf{F}_T$ (by Proposition 5.11.19 **(a)**), we can rewrite this as $\mathbf{F}_T^2 = \dfrac{n!}{f^\lambda} \mathbf{F}_T$. This proves Proposition 5.11.19 **(b)**.

**(c)** Assume that $h^\lambda$ is invertible in $\mathbf{k}$. The following is easy:

*Claim 1:* For each $\mathbf{b} \in \mathcal{A}\mathbf{E}_T$, we have $\mathbf{b} \nabla^-_{\operatorname{Col} T} \in \mathcal{A}\mathbf{F}_T$.

*Proof of Claim 1.* Let $\mathbf{b} \in \mathcal{A}\mathbf{E}_T$. Thus, $\mathbf{b} = \mathbf{a}\mathbf{E}_T$ for some $\mathbf{a} \in \mathcal{A}$. Consider this $\mathbf{a}$. Then, $\mathbf{b} = \mathbf{a}\mathbf{E}_T = \mathbf{a}\nabla^-_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}$ (since $\mathbf{E}_T = \nabla^-_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}$), so that

$$\mathbf{b}\nabla^-_{\mathrm{Col}\,T} = \mathbf{a}\nabla^-_{\mathrm{Col}\,T}\underbrace{\nabla_{\mathrm{Row}\,T}\nabla^-_{\mathrm{Col}\,T}}_{=\mathbf{F}_T} = \underbrace{\mathbf{a}\nabla^-_{\mathrm{Col}\,T}}_{\in\mathcal{A}}\mathbf{F}_T \in \mathcal{A}\mathbf{F}_T.$$

This proves Claim 1. □

*Claim 2:* For each $\mathbf{b} \in \mathcal{A}\mathbf{F}_T$, we have $\mathbf{b}\nabla_{\mathrm{Row}\,T} \in \mathcal{A}\mathbf{E}_T$.

*Proof of Claim 2.* This is analogous to Claim 1, but now the roles of $\nabla^-_{\mathrm{Col}\,T}$ and $\nabla_{\mathrm{Row}\,T}$ are interchanged (and so are the roles of $\mathbf{E}_T$ and $\mathbf{F}_T$). □

*Claim 3:* For each $\mathbf{b} \in \mathcal{A}\mathbf{E}_T$, we have $\mathbf{b}\mathbf{E}_T = h^\lambda \cdot \mathbf{b}$.

*Proof of Claim 3.* Let $\mathbf{b} \in \mathcal{A}\mathbf{E}_T$. Thus, $\mathbf{b} = \mathbf{a}\mathbf{E}_T$ for some $\mathbf{a} \in \mathcal{A}$. Consider this $\mathbf{a}$.

Theorem 5.11.3 yields $\mathbf{E}_T^2 = \dfrac{n!}{f^\lambda}\mathbf{E}_T$. In other words, $\mathbf{E}_T^2 = h^\lambda\mathbf{E}_T$ (since $\dfrac{n!}{f^\lambda} = h^\lambda$). But

$$\underbrace{\mathbf{b}}_{=\mathbf{a}\mathbf{E}_T}\mathbf{E}_T = \mathbf{a}\ \underbrace{\mathbf{E}_T\mathbf{E}_T}_{=\mathbf{E}_T^2=h^\lambda\mathbf{E}_T} = h^\lambda \cdot \underbrace{\mathbf{a}\mathbf{E}_T}_{=\mathbf{b}} = h^\lambda \cdot \mathbf{b}.$$

This proves Claim 3. □

*Claim 4:* For each $\mathbf{b} \in \mathcal{A}\mathbf{F}_T$, we have $\mathbf{b}\mathbf{F}_T = h^\lambda \cdot \mathbf{b}$.

*Proof of Claim 4.* This is analogous to Claim 3, but now the roles of $\nabla^-_{\mathrm{Col}\,T}$ and $\nabla_{\mathrm{Row}\,T}$ are interchanged (and so are the roles of $\mathbf{E}_T$ and $\mathbf{F}_T$). This time, we have to use Proposition 5.11.19 **(b)** instead of Theorem 5.11.3. □

Each $\mathbf{b} \in \mathcal{A}\mathbf{E}_T$ satisfies $\mathbf{b}\nabla^-_{\mathrm{Col}\,T} \in \mathcal{A}\mathbf{F}_T$ (by Claim 1). Thus, we can define the map

$$\begin{aligned} f : \mathcal{A}\mathbf{E}_T &\to \mathcal{A}\mathbf{F}_T, \\ \mathbf{b} &\mapsto \mathbf{b}\nabla^-_{\mathrm{Col}\,T}. \end{aligned}$$

Furthermore, each $\mathbf{b} \in \mathcal{A}\mathbf{F}_T$ satisfies $\mathbf{b}\nabla_{\mathrm{Row}\,T} \in \mathcal{A}\mathbf{E}_T$ (by Claim 2) and therefore $\mathbf{b} \cdot \dfrac{1}{h^\lambda}\nabla_{\mathrm{Row}\,T} = \dfrac{1}{h^\lambda}\underbrace{\mathbf{b}\nabla_{\mathrm{Row}\,T}}_{\in\mathcal{A}\mathbf{E}_T} \in \mathcal{A}\mathbf{E}_T$ (since $\mathcal{A}\mathbf{E}_T$ is a $\mathbf{k}$-module). Thus, we can define the map

$$\begin{aligned} g : \mathcal{A}\mathbf{F}_T &\to \mathcal{A}\mathbf{E}_T, \\ \mathbf{b} &\mapsto \mathbf{b} \cdot \frac{1}{h^\lambda}\nabla_{\mathrm{Row}\,T}. \end{aligned}$$

Consider these two maps $f$ and $g$. Both maps $f$ and $g$ are left $\mathcal{A}$-linear (this is clear from their definitions), i.e., are left $\mathcal{A}$-module morphisms. Moreover, we have $f \circ g = \mathrm{id}$, since each $\mathbf{b} \in \mathcal{A}\mathbf{F}_T$ satisfies

$$(f \circ g)(\mathbf{b}) = f(g(\mathbf{b})) = \underbrace{g(\mathbf{b})}_{\substack{= \mathbf{b} \cdot \frac{1}{h^\lambda} \nabla_{\mathrm{Row}\, T} \\ \text{(by the definition of } g\text{)}}} \cdot \nabla^-_{\mathrm{Col}\, T} \qquad \text{(by the definition of } f\text{)}$$

$$= \mathbf{b} \cdot \frac{1}{h^\lambda} \nabla_{\mathrm{Row}\, T} \nabla^-_{\mathrm{Col}\, T} = \frac{1}{h^\lambda} \mathbf{b} \underbrace{\nabla_{\mathrm{Row}\, T} \nabla^-_{\mathrm{Col}\, T}}_{= \mathbf{F}_T} = \frac{1}{h^\lambda} \underbrace{\mathbf{b}\mathbf{F}_T}_{\substack{= h^\lambda \cdot \mathbf{b} \\ \text{(by Claim 4)}}}$$

$$= \frac{1}{h^\lambda} h^\lambda \cdot \mathbf{b} = \mathbf{b} = \mathrm{id}(\mathbf{b}).$$

Furthermore, we have $g \circ f = \mathrm{id}$, since each $\mathbf{b} \in \mathcal{A}\mathbf{E}_T$ satisfies

$$(g \circ f)(\mathbf{b}) = g(f(\mathbf{b})) = \underbrace{f(\mathbf{b})}_{\substack{= \mathbf{b} \nabla^-_{\mathrm{Col}\, T} \\ \text{(by the definition of } f\text{)}}} \cdot \frac{1}{h^\lambda} \nabla_{\mathrm{Row}\, T} \qquad \text{(by the definition of } g\text{)}$$

$$= \mathbf{b} \nabla^-_{\mathrm{Col}\, T} \cdot \frac{1}{h^\lambda} \nabla_{\mathrm{Row}\, T} = \frac{1}{h^\lambda} \mathbf{b} \underbrace{\nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}}_{= \mathbf{E}_T} = \frac{1}{h^\lambda} \underbrace{\mathbf{b}\mathbf{E}_T}_{\substack{= h^\lambda \cdot \mathbf{b} \\ \text{(by Claim 3)}}}$$

$$= \frac{1}{h^\lambda} h^\lambda \cdot \mathbf{b} = \mathbf{b} = \mathrm{id}(\mathbf{b}).$$

The maps $f$ and $g$ are mutually inverse (since $f \circ g = \mathrm{id}$ and $g \circ f = \mathrm{id}$), and thus are left $\mathcal{A}$-module isomorphisms (since they are left $\mathcal{A}$-module morphisms). Thus, $\mathcal{A}\mathbf{E}_T \cong \mathcal{A}\mathbf{F}_T$ as left $\mathcal{A}$-modules.

But $T$ is an $n$-tableau of shape $\lambda$, that is, an $n$-tableau of shape $Y(\lambda)$. Hence, (228) (applied to $D = Y(\lambda)$) shows that $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules. In other words, $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules (since $\mathcal{S}^\lambda$ is a shorthand for $\mathcal{S}^{Y(\lambda)}$). Thus, $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T \cong \mathcal{A}\mathbf{F}_T$ as left $\mathcal{A}$-modules. Proposition 5.11.19 **(c)** is thus proved. □

Proposition 5.11.19 **(c)** would not be true if the requirement that $h^\lambda$ be invertible was dropped. An example is given in the following exercise:

**Exercise 5.11.3.** Let $n \geq 2$ and $\lambda = (n-1, 1)$. Pick any $n$-tableau $T$ of shape $\lambda$. Let $\mathbf{F}_T$ be defined as in Proposition 5.11.19. Show that:

**(a)** [1] The $\mathcal{A}$-module $\mathcal{A}\mathbf{E}_T$ is isomorphic to the zero-sum representation $R(\mathbf{k}^n)$ (see Example 5.4.2).

**(b)** [4] The $\mathcal{A}$-module $\mathcal{A}\mathbf{F}_T$ is isomorphic to the quotient representation $\mathbf{k}^n / D(\mathbf{k}^n)$ (see Theorem 4.2.30).

**(c)** $\boxed{1}$ Assume that $n \geq 3$, and that $\mathbf{k}$ is a field of characteristic $p$ with $p \mid n$. Then, we don't have $\mathcal{A}\mathbf{E}_T \cong \mathcal{A}\mathbf{F}_T$ as left $\mathcal{A}$-modules.

However, $\mathcal{A}\mathbf{E}_T$ and $\mathcal{A}\mathbf{F}_T$ are always isomorphic as $\mathbf{k}$-modules (see Exercise 5.15.4 below).

## 5.12. Centralized Young symmetrizers

Having studied the Young symmetrizers $\mathbf{E}_T$ corresponding to straight-shaped $n$-tableaux $T$, we shall now explore certain "centralized" versions thereof, which are the sums of the Young symmetrizers $\mathbf{E}_T$ over all $n$-tableaux $T$ of a given shape $Y(\lambda)$. We could just as well call them "symmetrized" or "averaged", since they make up for the asymmetric placement of the numbers in a given $n$-tableau by summing over **all** $n$-tableaux of that shape. However, we will soon see why "centralized" is the most adequate word.

We will keep using Definition 5.11.1 throughout this section. Also, recall that if $\lambda$ is a partition, then a tableau of shape $\lambda$ means a tableau of shape $Y(\lambda)$.

### 5.12.1. Definition and centrality

**Definition 5.12.1.** For any partition $\lambda$ of $n$, we define the *centralized Young symmetrizer*

$$\mathbf{E}_\lambda := \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T. \tag{241}$$

Note that this sum has $n!$ addends, since it ranges over all the $n!$ many $n$-tableaux of shape $\lambda$ (not just the standard ones).

**Example 5.12.2.** Let $n = 3$. Then, using the formulas from Example 5.11.2, we can compute the centralized Young symmetrizers $\mathbf{E}_\lambda$ corresponding to all three partitions $(3)$, $(1,1,1)$ and $(2,1)$ of 3. Namely, (where each $\sum$ sign means a sum over all the six orderings $(i,j,k)$ of $[3]$) we have

$$\begin{aligned}
\mathbf{E}_{(3)} &= \sum \mathbf{E}_{ijk} = \sum \nabla = 6\nabla; \\
\mathbf{E}_{(1,1,1)} &= \sum \mathbf{E}_{i\backslash\backslash j\backslash\backslash k} = \sum \nabla^- = 6\nabla^-; \\
\mathbf{E}_{(2,1)} &= \sum \mathbf{E}_{ij\backslash\backslash k} = \sum \left(1 - t_{i,k} + t_{i,j} - \mathrm{cyc}_{i,j,k}\right) \\
&= 6 - \sum \mathrm{cyc}_{i,j,k} = 6 - 3\left(\mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2}\right).
\end{aligned}$$

The word "centralized" in the name of the $\mathbf{E}_\lambda$ is well-deserved:

**Proposition 5.12.3.** Let $\lambda$ be a partition of $n$. Then, $\mathbf{E}_\lambda$ belongs to the center $Z(\mathbf{k}[S_n])$ of $\mathbf{k}[S_n]$.

*Proof.* Let $w \in S_n$. Then, the action of $w \in S_n$ on the set $\{n\text{-tableaux of shape } \lambda\}$ is a permutation of this set. Thus, the map

$$\begin{aligned} \{n\text{-tableaux of shape } \lambda\} &\to \{n\text{-tableaux of shape } \lambda\}, \\ T &\mapsto wT \end{aligned} \tag{242}$$

is a bijection (and its inverse is given by $T \mapsto w^{-1}T$). Now, we have

$$\begin{aligned} w\mathbf{E}_\lambda w^{-1} &= w\left(\sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T\right) w^{-1} \qquad \text{(by (241))} \\ &= \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \underbrace{w\mathbf{E}_T w^{-1}}_{\substack{=\mathbf{E}_{wT} \\ \text{(by Lemma 5.11.14)}}} \\ &= \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_{wT} = \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T \\ &\qquad \left(\begin{array}{c}\text{here, we have substituted } T \text{ for } wT \text{ in the sum,} \\ \text{since the map (242) is a bijection}\end{array}\right) \\ &= \mathbf{E}_\lambda \qquad \text{(by (241))}. \end{aligned}$$

Forget that we fixed $w$. We thus have shown that each $w \in S_n$ satisfies $w\mathbf{E}_\lambda w^{-1} = \mathbf{E}_\lambda$. Thus, Lemma 3.9.3 (applied to $G = S_n$ and $\mathbf{a} = \mathbf{E}_\lambda$) shows that $\mathbf{E}_\lambda \in Z(\mathbf{k}[S_n])$. This proves Proposition 5.12.3. □

If this centrality was the only property of the $\mathbf{E}_\lambda$, then it would be worth an exercise at best. But in fact, the $\mathbf{E}_\lambda$ play an important role in the center of $\mathbf{k}[S_n]$:

**Theorem 5.12.4.** Assume that $\mathbf{k}$ is a field of characteristic 0. Then:

**(a)** The family

$$(\mathbf{E}_\lambda)_{\lambda \text{ is a partition of } n}$$

is a basis of the $\mathbf{k}$-vector space $Z(\mathbf{k}[S_n])$.

**(b)** Each $\mathbf{E}_\lambda$ is quasi-idempotent, with $\mathbf{E}_\lambda^2 = c_\lambda \mathbf{E}_\lambda$ for a certain nonzero $c_\lambda \in \mathbb{Q}$.

**(c)** The $\mathbf{E}_\lambda$ are orthogonal, in the sense that $\mathbf{E}_\lambda \mathbf{E}_\mu = 0$ whenever $\lambda \neq \mu$.

We will prove this theorem soon, along with some slightly stronger claims. Once it is proved, it will follow easily that $Z(\mathbf{k}[S_n])$ is isomorphic to $\mathbf{k} \times \mathbf{k} \times \cdots \times \mathbf{k}$ as a $\mathbf{k}$-algebra when $\mathbf{k}$ is a field of characteristic 0. From here, it won't be far to Theorem 3.15.1 **(b)**.

First, however, let us rewrite a centralized Young symmetrizer $\mathbf{E}_\lambda$:

**Proposition 5.12.5.** Let $\lambda$ be a partition of $n$. Let $P$ be an $n$-tableau of shape $Y(\lambda)$. Then,

$$\mathbf{E}_\lambda = \sum_{w \in S_n} w \mathbf{E}_P w^{-1}.$$

*Proof.* Proposition 5.3.13 **(d)** (applied to $D = Y(\lambda)$ and $T = P$) shows that the map

$$\begin{aligned} f : S_n &\to \{n\text{-tableaux of shape } Y(\lambda)\}, \\ w &\mapsto w \rightharpoonup P \end{aligned}$$

is a bijection. Consider this bijection $f$.

But (241) shows that

$$\begin{aligned} \mathbf{E}_\lambda &= \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T \\ &= \sum_{T \text{ is an } n\text{-tableau of shape } Y(\lambda)} \mathbf{E}_T \qquad \left( \begin{array}{c} \text{since an } n\text{-tableau of shape } \lambda \\ \text{is the same thing as} \\ \text{an } n\text{-tableau of shape } Y(\lambda) \end{array} \right) \\ &= \sum_{T \in \{n\text{-tableaux of shape } Y(\lambda)\}} \mathbf{E}_T \\ &= \sum_{w \in S_n} \underbrace{\mathbf{E}_{f(w)}}_{= \mathbf{E}_{wP}} \\ &\qquad \begin{array}{c} \text{(since the definition} \\ \text{of } f \text{ yields } f(w) = w \rightharpoonup P = wP \\ \text{(since } wP \text{ is shorthand for } w \rightharpoonup P)) \end{array} \\ &\qquad \left( \begin{array}{c} \text{here, we have substituted } f(w) \text{ for } T \text{ in the sum,} \\ \text{since the map } f : S_n \to \{n\text{-tableaux of shape } Y(\lambda)\} \\ \text{is a bijection} \end{array} \right) \\ &= \sum_{w \in S_n} \underbrace{\mathbf{E}_{wP}}_{= w \mathbf{E}_P w^{-1}} = \sum_{w \in S_n} w \mathbf{E}_P w^{-1}. \\ &\qquad \begin{array}{c} \text{(by Lemma 5.11.14,} \\ \text{applied to } T = P) \end{array} \end{aligned}$$

This proves Proposition 5.12.5. □

### 5.12.2. Proof of the basic claims

We shall prove Theorem 5.12.4 step by step. We begin with the orthogonality part:

**Proposition 5.12.6.** Let $\lambda$ and $\mu$ be two distinct partitions of $n$. Then, $\mathbf{E}_\lambda \mathbf{E}_\mu = 0$.

*Proof.* The double zero-sandwich lemma (Proposition 5.11.15) says that

$$\mathbf{E}_S \mathbf{a} \mathbf{E}_T = 0 \tag{243}$$

for all $n$-tableaux $S$ of shape $Y(\lambda)$ and all $n$-tableaux $T$ of shape $Y(\mu)$ and all $\mathbf{a} \in \mathbf{k}[S_n]$. But the definition of $\mathbf{E}_\lambda$ yields

$$\begin{aligned}\mathbf{E}_\lambda &= \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T = \sum_{S \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_S \\ &= \sum_{S \text{ is an } n\text{-tableau of shape } Y(\lambda)} \mathbf{E}_S\end{aligned}$$

(since "shape $\lambda$" means "shape $Y(\lambda)$"), and similarly we find

$$\mathbf{E}_\mu = \sum_{T \text{ is an } n\text{-tableau of shape } Y(\mu)} \mathbf{E}_T.$$

Multiplying these two equalities, we obtain

$$\begin{aligned}\mathbf{E}_\lambda \mathbf{E}_\mu &= \left( \sum_{S \text{ is an } n\text{-tableau of shape } Y(\lambda)} \mathbf{E}_S \right) \left( \sum_{T \text{ is an } n\text{-tableau of shape } Y(\mu)} \mathbf{E}_T \right) \\ &= \sum_{S \text{ is an } n\text{-tableau of shape } Y(\lambda)} \ \sum_{T \text{ is an } n\text{-tableau of shape } Y(\mu)} \underbrace{\mathbf{E}_S \mathbf{E}_T}_{\substack{= \mathbf{E}_S 1 \mathbf{E}_T \\ = 0 \\ \text{(by (243),} \\ \text{applied to } \mathbf{a}=1)}} = 0.\end{aligned}$$

This proves Proposition 5.12.6. □

Thus, Theorem 5.12.4 **(c)** is proven.

The following lemma is an easy consequence of Lemma 5.11.13:

**Lemma 5.12.7.** Let $\lambda$ be a partition of $n$. Then, $[1]\, \mathbf{E}_\lambda = n!$.

*Proof.* The $n$-tableaux of shape $\lambda$ are just the bijections from $Y(\lambda)$ to $[n]$. Since both $Y(\lambda)$ and $[n]$ are $n$-element sets (because $\lambda$ is a partition of $n$), there are exactly $n!$ such bijections. In other words, there are $n!$ many $n$-tableaux of shape $\lambda$.

From (241), we obtain

$$\begin{aligned}[1]\, \mathbf{E}_\lambda &= [1] \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T \\ &= \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \underbrace{[1]\, \mathbf{E}_T}_{\substack{=1 \\ \text{(by Lemma 5.11.13)}}} \\ &= \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} 1 = n!\end{aligned}$$

(since there are $n!$ many $n$-tableaux of shape $\lambda$). This proves Lemma 5.12.7. □

The next lemma looks hardly like progress, but it will serve a crucial role in proving the rest of the theorem:

**Lemma 5.12.8.** Let $\lambda$ be a partition of $n$. Then, $\mathbf{E}_\lambda^2 \neq 0$ if $\mathbf{k}$ is a field of characteristic 0.

*Proof.* Assume that $\mathbf{k}$ is a field of characteristic 0. Then, the left ideal $\mathcal{A}\mathbf{a}$ of $\mathcal{A}$ is a finite-dimensional $\mathbf{k}$-vector space (since it is a subspace of the finite-dimensional $\mathbf{k}$-vector space $\mathcal{A}$), and thus is a free $\mathbf{k}$-module of rank $d$ for some $d \in \mathbb{N}$ (namely, for $d = \dim(\mathcal{A}\mathbf{a})$). Consider this $d$.

We must prove that $\mathbf{E}_\lambda^2 \neq 0$. Assume the contrary. Thus, $\mathbf{E}_\lambda^2 = 0 = 0\mathbf{E}_\lambda$. Hence, Lemma 5.11.12 (applied to $G = S_n$ and $\mathbf{a} = \mathbf{E}_\lambda$ and $\kappa = 0$) yields

$$d \cdot 0 = \underbrace{|S_n|}_{=n!} \cdot \underbrace{[1]\,\mathbf{E}_\lambda}_{\substack{=n! \\ \text{(by Lemma 5.12.7)}}} = n! \cdot n! = n!^2 \neq 0 \qquad \text{in } \mathbf{k}$$

(since $\mathbf{k}$ has characteristic 0). This contradicts $d \cdot 0 = 0$. This contradiction shows that our assumption was false, so that $\mathbf{E}_\lambda^2 \neq 0$ is proved. Thus, Lemma 5.12.8 is proved. □

We can now prove Theorem 5.12.4 **(a)**:

**Theorem 5.12.9.** Assume that $\mathbf{k}$ is a field of characteristic 0. Then, the family $(\mathbf{E}_\lambda)_{\lambda \text{ is a partition of } n}$ is a basis of the $\mathbf{k}$-vector space $Z(\mathbf{k}[S_n])$.

*Proof.* Let $P$ be the set of all partitions of $n$. Thus, we must show that the family $(\mathbf{E}_\lambda)_{\lambda \in P}$ is a basis of the $\mathbf{k}$-vector space $Z(\mathbf{k}[S_n])$. First of all, we observe that this family really consists of vectors in $Z(\mathbf{k}[S_n])$, since Proposition 5.12.3 shows that each $\lambda \in P$ satisfies $\mathbf{E}_\lambda \in Z(\mathbf{k}[S_n])$.

Note that $P$ is a finite set, and its size is $|P| = (\#\text{ of partitions of } n)$ (since $P$ is the set of all partitions of $n$).

Theorem 3.9.5 (applied to $G = S_n$) shows that the conjugacy class sums of $S_n$ form a basis of the $\mathbf{k}$-module $Z(\mathbf{k}[S_n])$. Hence, the dimension of $Z(\mathbf{k}[S_n])$ as a $\mathbf{k}$-vector space is

$$\begin{aligned}
&(\#\text{ of finite conjugacy classes of } S_n) \\
&\qquad\qquad \left(\begin{array}{c}\text{since the conjugacy class sums are} \\ \text{indexed by the finite conjugacy classes}\end{array}\right) \\
&= (\#\text{ of conjugacy classes of } S_n) \\
&\qquad\qquad (\text{since all conjugacy classes of } S_n \text{ are finite}) \\
&= (\#\text{ of partitions of } |[n]|) \\
&\qquad\qquad (\text{by Corollary 3.9.10 } \textbf{(c)}, \text{ applied to } X = [n]) \\
&= (\#\text{ of partitions of } n) \qquad (\text{since } |[n]| = n) \\
&= |P| \qquad (\text{since the set of all partitions of } n \text{ is } P).
\end{aligned}$$

In other words, the $\mathbf{k}$-vector space $Z(\mathbf{k}[S_n])$ is $|P|$-dimensional. Thus, our family $(\mathbf{E}_\lambda)_{\lambda\in P}$ has the right size to be a basis of $Z(\mathbf{k}[S_n])$. Therefore, in order to show that it is a basis of $Z(\mathbf{k}[S_n])$, it suffices to prove that it is $\mathbf{k}$-linearly independent (by a basic fact from linear algebra, since we are working over a field[93]). Let us do this:

*Claim 1:* The family $(\mathbf{E}_\lambda)_{\lambda\in P}$ is $\mathbf{k}$-linearly independent.

*Proof of Claim 1.* Let $(\alpha_\mu)_{\mu\in P} \in \mathbf{k}^P$ be a family of scalars such that

$$\sum_{\mu\in P} \alpha_\mu \mathbf{E}_\mu = 0. \tag{244}$$

We shall show that all the coefficients $\alpha_\mu$ are zero. Indeed, fix any $\lambda \in P$. Then, $\lambda$ is a partition of $n$. However, from (244), we obtain $\mathbf{E}_\lambda \sum_{\mu\in P} \alpha_\mu \mathbf{E}_\mu = \mathbf{E}_\lambda \cdot 0 = 0$. Hence,

$$\begin{aligned}
0 &= \mathbf{E}_\lambda \sum_{\mu\in P} \alpha_\mu \mathbf{E}_\mu = \sum_{\mu\in P} \alpha_\mu \mathbf{E}_\lambda \mathbf{E}_\mu \\
&= \alpha_\lambda \underbrace{\mathbf{E}_\lambda \mathbf{E}_\lambda}_{=\mathbf{E}_\lambda^2} + \sum_{\substack{\mu\in P;\\ \mu\neq\lambda}} \alpha_\mu \underbrace{\mathbf{E}_\lambda \mathbf{E}_\mu}_{\substack{=0 \\ \text{(by Proposition 5.12.6,} \\ \text{since } \lambda \text{ and } \mu \text{ are two} \\ \text{distinct partitions of } n\text{)}}} \\
&\qquad \left( \begin{array}{c} \text{here, we have split off the} \\ \text{addend for } \mu = \lambda \text{ from the sum} \end{array} \right) \\
&= \alpha_\lambda \mathbf{E}_\lambda^2 + \underbrace{\sum_{\substack{\mu\in P;\\ \mu\neq\lambda}} \alpha_\mu 0}_{=0} = \alpha_\lambda \mathbf{E}_\lambda^2.
\end{aligned}$$

But Lemma 5.12.8 yields $\mathbf{E}_\lambda^2 \neq 0$. If $\alpha_\lambda$ was nonzero, then we could divide the equality $0 = \alpha_\lambda \mathbf{E}_\lambda^2$ by $\alpha_\lambda$ (since $\mathbf{k}$ is a field!) and obtain $0 = \mathbf{E}_\lambda^2$, which would contradict $\mathbf{E}_\lambda^2 \neq 0$. Thus, $\alpha_\lambda$ cannot be nonzero. Hence, $\alpha_\lambda = 0$.

Forget that we fixed $\lambda$. We thus have shown that each $\lambda \in P$ satisfies $\alpha_\lambda = 0$. In other words, each $\mu \in P$ satisfies $\alpha_\mu = 0$.

Forget that we fixed $(\alpha_\mu)_{\mu\in P}$. We thus have shown that if $(\alpha_\mu)_{\mu\in P} \in \mathbf{k}^P$ is a family of scalars such that $\sum_{\mu\in P} \alpha_\mu \mathbf{E}_\mu = 0$, then each $\mu \in P$ satisfies $\alpha_\mu = 0$. In other words, the family $(\mathbf{E}_\mu)_{\mu\in P}$ is $\mathbf{k}$-linearly independent. In other words, the family $(\mathbf{E}_\lambda)_{\lambda\in P}$ is $\mathbf{k}$-linearly independent. This proves Claim 1. □

[93] We will explain this argument in more detail at the end of this proof.

However, $\mathbf{k}$ is a field. Thus, a classical fact from linear algebra (see, e.g., [Axler24, 2.38] or [Conrad, "The dimension of a vector space", Theorem 2.10, last sentence]) says that if $V$ is a $d$-dimensional $\mathbf{k}$-vector space, then any $\mathbf{k}$-linearly independent family of $d$ vectors in $V$ must be a basis of $V$. Applying this to $V = Z(\mathbf{k}[S_n])$ and $d = |P|$ and to the $\mathbf{k}$-linearly independent family $(\mathbf{E}_\lambda)_{\lambda \in P}$, we thus conclude that the family $(\mathbf{E}_\lambda)_{\lambda \in P}$ must be a basis of $Z(\mathbf{k}[S_n])$. This proves Theorem 5.12.9. □

Next let us prove Theorem 5.12.4 **(b)**:

**Proposition 5.12.10.** Assume that $\mathbf{k}$ is a commutative $\mathbb{Q}$-algebra. Let $\lambda$ be a partition of $n$. Then, $\mathbf{E}_\lambda^2 = c_\lambda \mathbf{E}_\lambda$ for a certain nonzero $c_\lambda \in \mathbb{Q}$.

*Proof.* If we can prove Proposition 5.12.10 for $\mathbf{k} = \mathbb{Q}$, then it will automatically follow that this proposition holds for any commutative $\mathbb{Q}$-algebra $\mathbf{k}$ (since there is a canonical ring morphism $\mathbb{Q} \to \mathbf{k}$, which induces a base change morphism $\mathbb{Q}[S_n] \to \mathbf{k}[S_n]$, and we can apply the latter base change morphism to the equality $\mathbf{E}_\lambda^2 = c_\lambda \mathbf{E}_\lambda$ in $\mathbb{Q}[S_n]$ to obtain the same equality in $\mathbf{k}[S_n]$ [94]).

Thus, we can WLOG assume that $\mathbf{k} = \mathbb{Q}$. Assume this. Hence, $\mathbf{k}$ is a field of characteristic 0.

Theorem 5.12.9 tells us that the family $\left(\mathbf{E}_\mu\right)_{\mu \text{ is a partition of } n}$ is a basis of the $\mathbf{k}$-vector space $Z(\mathbf{k}[S_n])$. Thus, in particular, this family spans $Z(\mathbf{k}[S_n])$. Hence, the vector 1 (which clearly belongs to $Z(\mathbf{k}[S_n])$) is a linear combination of this family. In other words, we can write 1 in the form

$$1 = \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\mu \qquad \text{for some scalars } \alpha_\mu \in \mathbf{k}.$$

Consider these scalars $\alpha_\mu$. Multiplying the equality $1 = \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\mu$ by $\mathbf{E}_\lambda$ from the left, we get

$$\begin{aligned}
\mathbf{E}_\lambda \cdot 1 &= \mathbf{E}_\lambda \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\mu = \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\lambda \mathbf{E}_\mu \\
&= \alpha_\lambda \underbrace{\mathbf{E}_\lambda \mathbf{E}_\lambda}_{=\mathbf{E}_\lambda^2} + \sum_{\substack{\mu \text{ is a partition of } n; \\ \mu \neq \lambda}} \alpha_\mu \underbrace{\mathbf{E}_\lambda \mathbf{E}_\mu}_{\substack{=0 \\ \text{(by Proposition 5.12.6)}}} \\
&\qquad \left( \begin{array}{c} \text{here, we have split off the} \\ \text{addend for } \mu = \lambda \text{ from the sum} \end{array} \right) \\
&= \alpha_\lambda \mathbf{E}_\lambda^2 + \underbrace{\sum_{\substack{\mu \text{ is a partition of } n; \\ \mu \neq \lambda}} \alpha_\mu 0}_{=0} = \alpha_\lambda \mathbf{E}_\lambda^2.
\end{aligned}$$

[94] This is analogous to the base change argument that we did at the end of our proof of Theorem 5.11.3.

Hence, $\alpha_\lambda \mathbf{E}_\lambda^2 = \mathbf{E}_\lambda \cdot 1 = \mathbf{E}_\lambda$.

But Lemma 5.12.8 yields $\mathbf{E}_\lambda^2 \neq 0$, and thus $\mathbf{E}_\lambda \neq 0$. Hence, $\alpha_\lambda \mathbf{E}_\lambda^2 = \mathbf{E}_\lambda \neq 0$, so that $\alpha_\lambda \neq 0$. Since $\alpha_\lambda \in \mathbf{k} = \mathbb{Q}$, the inverse $\dfrac{1}{\alpha_\lambda}$ is a well-defined nonzero element of $\mathbb{Q}$, and we can divide the equality $\alpha_\lambda \mathbf{E}_\lambda^2 = \mathbf{E}_\lambda$ by $\alpha_\lambda$ and obtain $\mathbf{E}_\lambda^2 = \dfrac{1}{\alpha_\lambda} \mathbf{E}_\lambda$. Hence, $\mathbf{E}_\lambda^2 = c_\lambda \mathbf{E}_\lambda$ for a certain nonzero $c_\lambda \in \mathbb{Q}$ (namely, for $c_\lambda = \dfrac{1}{\alpha_\lambda}$). This proves Proposition 5.12.10. □

### 5.12.3. A stronger quasi-idempotence

We have thus proved Theorem 5.12.4 completely. This is a good beginning, but we can improve the theorem in several directions, and prove some further properties of centralized Young symmetrizers. In doing so, we will use the following notations:

**Definition 5.12.11.** Let $\lambda$ be a partition of $n$. Then:

**(a)** We let $f^\lambda$ be the # of standard tableaux of shape $Y(\lambda)$.

**(b)** We set $h^\lambda := \dfrac{n!}{f^\lambda}$. This is a positive integer (by Theorem 5.11.3).

These notations have been already used above, but we shall now use them without explicit mention throughout the rest of this chapter.

**Example 5.12.12.** Here are lists of all partitions of $n$ for $n \leq 5$ with the corresponding numbers $f^\lambda$ and $h^\lambda$:

| $\lambda$ | $()$ | $(1)$ | $(2)$ | $(1,1)$ | $(3)$ | $(2,1)$ | $(1,1,1)$ |
|---|---|---|---|---|---|---|---|
| $f^\lambda$ | 1 | 1 | 1 | 1 | 1 | 2 | 1 |
| $h^\lambda$ | 1 | 1 | 2 | 2 | 6 | 3 | 6 |

| $\lambda$ | $(4)$ | $(3,1)$ | $(2,2)$ | $(2,1,1)$ | $(1,1,1,1)$ |
|---|---|---|---|---|---|
| $f^\lambda$ | 1 | 3 | 2 | 3 | 1 |
| $h^\lambda$ | 24 | 8 | 12 | 8 | 24 |

| $\lambda$ | $(5)$ | $(4,1)$ | $(3,2)$ | $(3,1,1)$ | $(2,2,1)$ | $(2,1,1,1)$ | $(1,1,1,1,1)$ |
|---|---|---|---|---|---|---|---|
| $f^\lambda$ | 1 | 4 | 5 | 6 | 5 | 4 | 1 |
| $h^\lambda$ | 120 | 30 | 24 | 20 | 24 | 30 | 120 |

**Remark 5.12.13.** Let $\lambda$ be a partition of $n$. Then, the hook length formula (Theorem 5.2.8) shows that the number $h^\lambda = \dfrac{n!}{f^\lambda}$ equals the product $\prod_{c \in Y(\lambda)} |H_\lambda(c)|$ of the sizes of the hooks of all cells $c \in Y(\lambda)$. We shall not need this fact, but we are mentioning it here since it gives the easiest way to compute $h^\lambda$.

Our first improvement on Theorem 5.12.4 is an explicit formula for the number $c_\lambda$ in Theorem 5.12.4 **(b)** (or in Proposition 5.12.10). Namely, we claim that $c_\lambda = \left(h^\lambda\right)^2$. In other words, we claim the following (using the notations of Definition 5.12.11):

**Theorem 5.12.14.** Let $\lambda$ be a partition of $n$. Then, $\mathbf{E}_\lambda^2 = \left(h^\lambda\right)^2 \mathbf{E}_\lambda$.

To prove this, we need another fact (which is also useful in its own right):

**Proposition 5.12.15.** Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $Y(\lambda)$. Then, $\mathbf{E}_\lambda \mathbf{E}_T = \left(h^\lambda\right)^2 \mathbf{E}_T$.

*Proof of Proposition 5.12.15.* We have $\mathbf{E}_\lambda = \sum_{S \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_S$ (by (241), with the index $T$ renamed as $S$). Thus,

$$\begin{aligned}\mathbf{E}_\lambda \mathbf{E}_T &= \left( \sum_{S \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_S \right) \mathbf{E}_T = \sum_{S \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_S \mathbf{E}_T \\ &= \sum_{\substack{S \text{ is an } n\text{-tableau of shape } \lambda; \\ \text{there exist } \widetilde{r} \in \mathcal{R}(T) \text{ and } \widetilde{c} \in \mathcal{C}(T) \\ \text{such that } S = \widetilde{c}\widetilde{r}T}} \mathbf{E}_S \mathbf{E}_T + \sum_{\substack{S \text{ is an } n\text{-tableau of shape } \lambda; \\ \text{there exist no } \widetilde{r} \in \mathcal{R}(T) \text{ and } \widetilde{c} \in \mathcal{C}(T) \\ \text{such that } S = \widetilde{c}\widetilde{r}T}} \mathbf{E}_S \mathbf{E}_T\end{aligned}$$

(here, we have split the sum according to the existence or non-existence of $\widetilde{r} \in \mathcal{R}(T)$ and $\widetilde{c} \in \mathcal{C}(T)$ such that $S = \widetilde{c}\widetilde{r}T$). Now, let us consider the two sums on the right hand side here. The second sum is easy to simplify:

$$\begin{aligned}&\sum_{\substack{S \text{ is an } n\text{-tableau of shape } \lambda; \\ \text{there exist no } \widetilde{r} \in \mathcal{R}(T) \text{ and } \widetilde{c} \in \mathcal{C}(T) \\ \text{such that } S = \widetilde{c}\widetilde{r}T}} \underbrace{\mathbf{E}_S \mathbf{E}_T}_{\substack{=0 \\ \text{(by Proposition 5.11.18 (a))}}} \\ &= \sum_{\substack{S \text{ is an } n\text{-tableau of shape } \lambda; \\ \text{there exist no } \widetilde{r} \in \mathcal{R}(T) \text{ and } \widetilde{c} \in \mathcal{C}(T) \\ \text{such that } S = \widetilde{c}\widetilde{r}T}} 0 = 0. \end{aligned} \tag{245}$$

It remains to simplify the left hand side. Here it helps to prove the following simple claim:

*Claim 1:* Let $S$ be an $n$-tableau of shape $\lambda$. If there exists a pair $(\widetilde{r}, \widetilde{c}) \in \mathcal{R}(T) \times \mathcal{C}(T)$ such that $S = \widetilde{c}\widetilde{r}T$, then this pair $(\widetilde{r}, \widetilde{c})$ is unique.

*Proof of Claim 1.* Assume that $(\widetilde{r}_1, \widetilde{c}_1)$ and $(\widetilde{r}_2, \widetilde{c}_2)$ are two such pairs. We must show that these two pairs are equal – i.e., that $(\widetilde{r}_1, \widetilde{c}_1) = (\widetilde{r}_2, \widetilde{c}_2)$.

We have assumed that $(\widetilde{r}_1, \widetilde{c}_1)$ is a pair $(\widetilde{r}, \widetilde{c}) \in \mathcal{R}(T) \times \mathcal{C}(T)$ such that $S = \widetilde{c}\widetilde{r}T$. In other words, $(\widetilde{r}_1, \widetilde{c}_1) \in \mathcal{R}(T) \times \mathcal{C}(T)$ and $S = \widetilde{c}_1\widetilde{r}_1 T$. Similarly, $(\widetilde{r}_2, \widetilde{c}_2) \in \mathcal{R}(T) \times \mathcal{C}(T)$ and $S = \widetilde{c}_2\widetilde{r}_2 T$. Comparing $S = \widetilde{c}_1\widetilde{r}_1 T$ with $S = \widetilde{c}_2\widetilde{r}_2 T$, we obtain $\widetilde{c}_1\widetilde{r}_1 T = \widetilde{c}_2\widetilde{r}_2 T$.

Moreover, from $(\widetilde{r}_1, \widetilde{c}_1) \in \mathcal{R}(T) \times \mathcal{C}(T)$, we obtain $\widetilde{r}_1 \in \mathcal{R}(T)$ and $\widetilde{c}_1 \in \mathcal{C}(T)$. Similarly, $\widetilde{r}_2 \in \mathcal{R}(T)$ and $\widetilde{c}_2 \in \mathcal{C}(T)$. From $\widetilde{c}_1 \in \mathcal{C}(T)$ and $\widetilde{c}_2 \in \mathcal{C}(T)$, we obtain $\widetilde{c}_1^{-1}\widetilde{c}_2 \in \mathcal{C}(T)$ (since $\mathcal{C}(T)$ is a group). Similarly, $\widetilde{r}_1\widetilde{r}_2^{-1} \in \mathcal{R}(T)$.

However, it is easy to see that any two permutations $u, v \in S_n$ satisfying $uT = vT$ must necessarily be equal[95]. Applying this to $u = \widetilde{c}_1\widetilde{r}_1$ and $v = \widetilde{c}_2\widetilde{r}_2$, we see that the two permutations $\widetilde{c}_1\widetilde{r}_1$ and $\widetilde{c}_2\widetilde{r}_2$ must be equal (since $\widetilde{c}_1\widetilde{r}_1 T = \widetilde{c}_2\widetilde{r}_2 T$). In other words, $\widetilde{c}_1\widetilde{r}_1 = \widetilde{c}_2\widetilde{r}_2$. Hence, $\widetilde{r}_1 = \widetilde{c}_1^{-1}\widetilde{c}_2\widetilde{r}_2$, so that $\widetilde{r}_1\widetilde{r}_2^{-1} = \widetilde{c}_1^{-1}\widetilde{c}_2 \in \mathcal{C}(T)$. Combining this with $\widetilde{r}_1\widetilde{r}_2^{-1} \in \mathcal{R}(T)$, we obtain $\widetilde{r}_1\widetilde{r}_2^{-1} \in \mathcal{R}(T) \cap \mathcal{C}(T) = \{\mathrm{id}\}$ (by Proposition 5.3.5 **(a)**). In other words, $\widetilde{r}_1\widetilde{r}_2^{-1} = \mathrm{id}$. Hence, $\widetilde{r}_1 = \widetilde{r}_2$. Furthermore, comparing $\widetilde{r}_1\widetilde{r}_2^{-1} = \mathrm{id}$ with $\widetilde{r}_1\widetilde{r}_2^{-1} = \widetilde{c}_1^{-1}\widetilde{c}_2$, we find $\widetilde{c}_1^{-1}\widetilde{c}_2 = \mathrm{id}$, and thus $\widetilde{c}_1 = \widetilde{c}_2$. Combining $\widetilde{r}_1 = \widetilde{r}_2$ with $\widetilde{c}_1 = \widetilde{c}_2$, we obtain $(\widetilde{r}_1, \widetilde{c}_1) = (\widetilde{r}_2, \widetilde{c}_2)$. This proves Claim 1. □

Now, the map

$$\begin{aligned}&\mathcal{R}(T) \times \mathcal{C}(T)\\ &\to \{n\text{-tableaux } S \text{ of shape } \lambda \mid \text{there exist } \widetilde{r} \in \mathcal{R}(T) \text{ and } \widetilde{c} \in \mathcal{C}(T)\\ &\qquad\qquad \text{such that } S = \widetilde{c}\widetilde{r}T\}\end{aligned}$$

that sends each pair $(\widetilde{r}, \widetilde{c}) \in \mathcal{R}(T) \times \mathcal{C}(T)$ to the $n$-tableau $\widetilde{c}\widetilde{r}T$ is clearly well-defined and surjective (indeed, its target has been handpicked to contain precisely all its values), and furthermore (by Claim 1) is injective. Hence, this map is a bijection. Thus, we can substitute $\widetilde{c}\widetilde{r}T$ for $S$ in the sum

$$\sum_{\substack{S \text{ is an } n\text{-tableau of shape } \lambda;\\ \text{there exist } \widetilde{r}\in\mathcal{R}(T) \text{ and } \widetilde{c}\in\mathcal{C}(T)\\ \text{such that } S=\widetilde{c}\widetilde{r}T}} \mathbf{E}_S \mathbf{E}_T.$$

[95] This was shown during the proof of Theorem 5.10.6 **(b)**.

We thus obtain

$$\begin{aligned}
\sum_{\substack{S \text{ is an } n\text{-tableau of shape } \lambda; \\ \text{there exist } \widetilde{r}\in\mathcal{R}(T) \text{ and } \widetilde{c}\in\mathcal{C}(T) \\ \text{such that } S=\widetilde{c}\widetilde{r}T}} \mathbf{E}_S\mathbf{E}_T &= \underbrace{\sum_{(\widetilde{r},\widetilde{c})\in\mathcal{R}(T)\times\mathcal{C}(T)}}_{=\sum\limits_{\widetilde{r}\in\mathcal{R}(T)}\sum\limits_{\widetilde{c}\in\mathcal{C}(T)}} \mathbf{E}_{\widetilde{c}\widetilde{r}T}\mathbf{E}_T \\
&= \sum_{\widetilde{r}\in\mathcal{R}(T)} \sum_{\widetilde{c}\in\mathcal{C}(T)} \underbrace{\mathbf{E}_{\widetilde{c}\widetilde{r}T}\mathbf{E}_T}_{\substack{=h^\lambda(-1)^{\widetilde{c}}\widetilde{c}\widetilde{r}\mathbf{E}_T \\ \text{(by Proposition 5.11.18 \textbf{(b)},} \\ \text{applied to } S=\widetilde{c}\widetilde{r}T)}} \\
&= \sum_{\widetilde{r}\in\mathcal{R}(T)} \sum_{\widetilde{c}\in\mathcal{C}(T)} h^\lambda\,(-1)^{\widetilde{c}}\,\widetilde{c}\widetilde{r}\mathbf{E}_T \\
&= h^\lambda \underbrace{\left(\sum_{\widetilde{c}\in\mathcal{C}(T)} (-1)^{\widetilde{c}}\widetilde{c}\right)}_{\substack{=\sum\limits_{w\in\mathcal{C}(T)}(-1)^w w \\ =\nabla^-_{\mathrm{Col}\,T} \\ \text{(by the definition} \\ \text{of } \nabla^-_{\mathrm{Col}\,T})}} \underbrace{\left(\sum_{\widetilde{r}\in\mathcal{R}(T)} \widetilde{r}\right)}_{\substack{=\sum\limits_{w\in\mathcal{R}(T)} w \\ =\nabla_{\mathrm{Row}\,T} \\ \text{(by the definition} \\ \text{of } \nabla_{\mathrm{Row}\,T})}} \mathbf{E}_T \\
&= h^\lambda \underbrace{\nabla^-_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}}_{\substack{=\mathbf{E}_T \\ \text{(by the definition of } \mathbf{E}_T)}} \mathbf{E}_T = h^\lambda \underbrace{\mathbf{E}_T\mathbf{E}_T}_{=\mathbf{E}_T^2} = h^\lambda \mathbf{E}_T^2.
\end{aligned}$$

Moreover, Theorem 5.11.3 yields $\mathbf{E}_T^2 = \dfrac{n!}{f^\lambda}\mathbf{E}_T = h^\lambda\mathbf{E}_T$ (since $\dfrac{n!}{f^\lambda} = h^\lambda$). Thus,

$$\begin{aligned}
\sum_{\substack{S \text{ is an } n\text{-tableau of shape } \lambda; \\ \text{there exist } \widetilde{r}\in\mathcal{R}(T) \text{ and } \widetilde{c}\in\mathcal{C}(T) \\ \text{such that } S=\widetilde{c}\widetilde{r}T}} \mathbf{E}_S\mathbf{E}_T &= h^\lambda \underbrace{\mathbf{E}_T^2}_{=h^\lambda\mathbf{E}_T} = h^\lambda h^\lambda \mathbf{E}_T \\
&= \left(h^\lambda\right)^2 \mathbf{E}_T. \qquad (246)
\end{aligned}$$

Now, we can finish the calculation that we started at the beginning of this proof:

$$\begin{aligned}
\mathbf{E}_\lambda\mathbf{E}_T &= \underbrace{\sum_{\substack{S \text{ is an } n\text{-tableau of shape } \lambda; \\ \text{there exist } \widetilde{r}\in\mathcal{R}(T) \text{ and } \widetilde{c}\in\mathcal{C}(T) \\ \text{such that } S=\widetilde{c}\widetilde{r}T}} \mathbf{E}_S\mathbf{E}_T}_{\substack{=\left(h^\lambda\right)^2\mathbf{E}_T \\ \text{(by (246))}}} + \underbrace{\sum_{\substack{S \text{ is an } n\text{-tableau of shape } \lambda; \\ \text{there exist no } \widetilde{r}\in\mathcal{R}(T) \text{ and } \widetilde{c}\in\mathcal{C}(T) \\ \text{such that } S=\widetilde{c}\widetilde{r}T}} \mathbf{E}_S\mathbf{E}_T}_{\substack{=0 \\ \text{(by (245))}}} \\
&= \left(h^\lambda\right)^2 \mathbf{E}_T + 0 = \left(h^\lambda\right)^2 \mathbf{E}_T.
\end{aligned}$$

This proves Proposition 5.12.15. □

Note that the product $\mathbf{E}_\lambda \mathbf{E}_T$ in Proposition 5.12.15 can also be rewritten as $\mathbf{E}_T \mathbf{E}_\lambda$ (since Proposition 5.12.3 shows that $\mathbf{E}_\lambda \in Z(\mathbf{k}[S_n])$ and thus $\mathbf{E}_\lambda \mathbf{E}_T = \mathbf{E}_T \mathbf{E}_\lambda$).

*Proof of Theorem 5.12.14.* We have

$$\begin{aligned}
\mathbf{E}_\lambda^2 &= \mathbf{E}_\lambda \mathbf{E}_\lambda = \mathbf{E}_\lambda \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T \qquad \text{(by (241))} \\
&= \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \underbrace{\mathbf{E}_\lambda \mathbf{E}_T}_{\substack{= \left(h^\lambda\right)^2 \mathbf{E}_T \\ \text{(by Proposition 5.12.15)}}} \\
&= \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \left(h^\lambda\right)^2 \mathbf{E}_T = \left(h^\lambda\right)^2 \underbrace{\sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T}_{\substack{= \mathbf{E}_\lambda \\ \text{(by (241))}}} \\
&= \left(h^\lambda\right)^2 \mathbf{E}_\lambda.
\end{aligned}$$

This proves Theorem 5.12.14. □

Theorem 5.12.14 has an easy but surprising consequence. As we recall, $\mathbf{k}$-algebras are also known as *unital* $\mathbf{k}$*-algebras*, since they have unities. In contrast, a *nonunital* $\mathbf{k}$*-algebra* is defined just like a $\mathbf{k}$-algebra, except that it is not required to have a unity. Obviously, each $\mathbf{k}$-algebra becomes a nonunital $\mathbf{k}$-algebra if we forget its unity. A *nonunital* $\mathbf{k}$*-subalgebra* of a nonunital $\mathbf{k}$-algebra $A$ is defined as a $\mathbf{k}$-submodule $B$ of $A$ that is closed under multiplication. Somewhat surprisingly, it can happen that $A$ and $B$ are two unital $\mathbf{k}$-algebras, and $B$ is a nonunital $\mathbf{k}$-subalgebra of $A$ but not a unital $\mathbf{k}$-subalgebra of $A$ (despite having a unity!). This is the case whenever the unities of $A$ and $B$ are distinct, and this does indeed happen quite often[96]. The following corollary gives another example of this paradoxical phenomenon:

**Corollary 5.12.16.** Recall that $\mathcal{A} = \mathbf{k}[S_n]$. Let $\lambda$ be a partition of $n$.

**(a)** The subset $\mathcal{A}\mathbf{E}_\lambda$ of $\mathcal{A}$ is a $\mathbf{k}$-submodule and is closed under multiplication. That is, it is a nonunital subalgebra of $\mathcal{A}$.

**(b)** Now assume that $h^\lambda$ is invertible in $\mathbf{k}$. Then, the nonunital algebra $\mathcal{A}\mathbf{E}_\lambda$ from Corollary 5.12.16 **(a)** has a unity, namely $\dfrac{1}{\left(h^\lambda\right)^2}\mathbf{E}_\lambda$. (But this does not make $\mathcal{A}\mathbf{E}_\lambda$ a unital subalgebra of $\mathcal{A}$.)

[96] For example: If $R$ and $S$ are two nontrivial $\mathbf{k}$-algebras, then $R \times \{0_S\}$ is a nonunital $\mathbf{k}$-subalgebra of $R \times S$, but is itself a unital $\mathbf{k}$-algebra (with unity $(1_R, 0_S)$), despite not being a unital $\mathbf{k}$-subalgebra of $R \times S$ (since the unity of $R \times S$ is $(1_R, 1_S)$).

*Proof.* **(a)** This is obvious. In fact, if $R$ is any $\mathbf{k}$-algebra, and if $r \in R$ is any element, then the subset $Rr$ of $R$ is a $\mathbf{k}$-submodule and is closed under multiplication (since the product of two elements of $Rr$ has the form $xr \cdot yr = \underbrace{(xry)}_{\in R} r$ and thus belongs to $Rr$ again), and thus is a nonunital subalgebra of $R$. Applying this fact to $R = \mathcal{A}$ and $r = \mathbf{E}_\lambda$, we obtain the claim of Corollary 5.12.16 **(a)**.

**(b)** Proposition 5.12.3 yields $\mathbf{E}_\lambda \in Z(\mathbf{k}[S_n])$. In other words, $\mathbf{E}_\lambda \in Z(\mathcal{A})$ (since $\mathcal{A} = \mathbf{k}[S_n]$).

Let $\mathbf{r} = \dfrac{1}{(h^\lambda)^2}\mathbf{E}_\lambda$. Thus, $\mathbf{r} = \dfrac{1}{(h^\lambda)^2}\mathbf{E}_\lambda = \underbrace{\dfrac{1}{(h^\lambda)^2} \cdot 1}_{\in \mathcal{A}}\mathbf{E}_\lambda \in \mathcal{A}\mathbf{E}_\lambda$ and

$$\mathbf{E}_\lambda \mathbf{r} = \mathbf{E}_\lambda \cdot \frac{1}{(h^\lambda)^2}\mathbf{E}_\lambda = \frac{1}{(h^\lambda)^2} \underbrace{\mathbf{E}_\lambda^2}_{\substack{= \left(h^\lambda\right)^2 \mathbf{E}_\lambda \\ \text{(by Theorem 5.12.14)}}} = \frac{1}{(h^\lambda)^2}\left(h^\lambda\right)^2 \mathbf{E}_\lambda = \mathbf{E}_\lambda.$$

Now, we claim that each $\mathbf{a} \in \mathcal{A}\mathbf{E}_\lambda$ satisfies $\mathbf{ar} = \mathbf{ra} = \mathbf{a}$.

[*Proof:* Let $\mathbf{a} \in \mathcal{A}\mathbf{E}_\lambda$. Thus, $\mathbf{a} = \mathbf{b}\mathbf{E}_\lambda$ for some $\mathbf{b} \in \mathcal{A}$. Consider this $\mathbf{b}$. Now, from $\mathbf{a} = \mathbf{b}\mathbf{E}_\lambda$, we obtain

$$\mathbf{ar} = \mathbf{b} \underbrace{\mathbf{E}_\lambda \mathbf{r}}_{= \mathbf{E}_\lambda} = \mathbf{b}\mathbf{E}_\lambda = \mathbf{a}.$$

Furthermore, from $\mathbf{r} = \dfrac{1}{(h^\lambda)^2}\mathbf{E}_\lambda$, we obtain

$$\mathbf{ra} = \frac{1}{(h^\lambda)^2} \underbrace{\mathbf{E}_\lambda \mathbf{a}}_{\substack{= \mathbf{a}\mathbf{E}_\lambda \\ \text{(since } \mathbf{E}_\lambda \in Z(\mathcal{A})\text{)}}} = \frac{1}{(h^\lambda)^2}\mathbf{a}\mathbf{E}_\lambda = \mathbf{a} \cdot \underbrace{\frac{1}{(h^\lambda)^2}\mathbf{E}_\lambda}_{= \mathbf{r}} = \mathbf{ar},$$

so that $\mathbf{ar} = \mathbf{ra} = \mathbf{a}$ (since $\mathbf{ra} = \mathbf{ar} = \mathbf{a}$). Qed.]

We thus have shown that $\mathbf{r} \in \mathcal{A}\mathbf{E}_\lambda$ and that each $\mathbf{a} \in \mathcal{A}\mathbf{E}_\lambda$ satisfies $\mathbf{ar} = \mathbf{ra} = \mathbf{a}$. In other words, $\mathbf{r}$ is a unity of the nonunital algebra $\mathcal{A}\mathbf{E}_\lambda$. In other words, $\dfrac{1}{(h^\lambda)^2}\mathbf{E}_\lambda$ is a unity of the nonunital algebra $\mathcal{A}\mathbf{E}_\lambda$ (since $\mathbf{r} = \dfrac{1}{(h^\lambda)^2}\mathbf{E}_\lambda$). This proves Corollary 5.12.16 **(b)**. □

### 5.12.4. A partition of unity into centralized Young symmetrizers

One of our next goals is to make Theorem 5.12.4 **(a)** more concrete by providing explicit expansions for elements of $Z(\mathbf{k}[S_n])$ as $\mathbf{k}$-linear combinations of the $\mathbf{E}_\lambda$. We begin with the element 1:

**Theorem 5.12.17.** Let $\mathbf{k}$ be a commutative $\mathbb{Q}$-algebra. Then, in $\mathbf{k}[S_n]$, we have

$$1 = \sum_{\lambda \text{ is a partition of } n} \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}.$$

*Proof.* As in the proof of Proposition 5.12.10, we can WLOG assume that $\mathbf{k} = \mathbb{Q}$ (since base change allows us to extend the result from $\mathbf{k} = \mathbb{Q}$ to any commutative $\mathbb{Q}$-algebra $\mathbf{k}$). Assume this. Thus, $\mathbf{k}$ is a field.

As in the proof of Proposition 5.12.10, we can write 1 in the form

$$1 = \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\mu \qquad \text{for some scalars } \alpha_\mu \in \mathbf{k}.$$

Consider these scalars $\alpha_\mu$.

Now, let $\lambda$ be a partition of $n$. As in the proof of Proposition 5.12.10, we can show that $\mathbf{E}_\lambda \neq 0$ and

$$\alpha_\lambda \mathbf{E}_\lambda^2 = \mathbf{E}_\lambda. \tag{247}$$

But Theorem 5.12.14 shows that $\mathbf{E}_\lambda^2 = \left(h^\lambda\right)^2 \mathbf{E}_\lambda$. Thus, we can rewrite the equality (247) as

$$\alpha_\lambda \left(h^\lambda\right)^2 \mathbf{E}_\lambda = \mathbf{E}_\lambda.$$

Thus, $\left(\alpha_\lambda \left(h^\lambda\right)^2 - 1\right) \mathbf{E}_\lambda = \underbrace{\alpha_\lambda \left(h^\lambda\right)^2 \mathbf{E}_\lambda}_{=\mathbf{E}_\lambda} - \mathbf{E}_\lambda = \mathbf{E}_\lambda - \mathbf{E}_\lambda = 0.$

But $\mathbf{k}$ is a field. Thus, if the scalar $\alpha_\lambda \left(h^\lambda\right)^2 - 1$ was nonzero, then this scalar would be invertible, and thus we could divide the equality $\left(\alpha_\lambda \left(h^\lambda\right)^2 - 1\right) \mathbf{E}_\lambda = 0$ by $\alpha_\lambda \left(h^\lambda\right)^2 - 1$, and obtain $\mathbf{E}_\lambda = 0$, which would contradict $\mathbf{E}_\lambda \neq 0$. Hence, the scalar $\alpha_\lambda \left(h^\lambda\right)^2 - 1$ cannot be nonzero. In other words, $\alpha_\lambda \left(h^\lambda\right)^2 - 1 = 0$. Equivalently, $\alpha_\lambda \left(h^\lambda\right)^2 = 1$. Thus, $\alpha_\lambda = \dfrac{1}{\left(h^\lambda\right)^2}$.

Forget that we fixed $\lambda$. We thus have shown that each partition $\lambda$ of $n$ satisfies

$$\alpha_\lambda = \frac{1}{\left(h^\lambda\right)^2}. \tag{248}$$

Now, recall that

$$1 = \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\mu = \sum_{\lambda \text{ is a partition of } n} \underbrace{\alpha_\lambda}_{\substack{= \dfrac{1}{\left(h^\lambda\right)^2} \\ \text{(by (248))}}} \mathbf{E}_\lambda$$

$$= \sum_{\lambda \text{ is a partition of } n} \frac{1}{\left(h^\lambda\right)^2} \mathbf{E}_\lambda = \sum_{\lambda \text{ is a partition of } n} \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}.$$

This proves Theorem 5.12.17. □

### 5.12.5. A combinatorial consequence: the sum of all $\left(f^\lambda\right)^2$

Theorem 5.12.17 leads to an important purely combinatorial identity:

**Corollary 5.12.18.** We have

$$\sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 = n!.$$

**Example 5.12.19.** Let $n = 4$. Then, the partitions of $n$ are $(4)$, $(3,1)$, $(2,2)$, $(2,1,1)$ and $(1,1,1,1)$, and the corresponding numbers $f^\lambda$ are

$$f^{(4)} = 1, \qquad f^{(3,1)} = 3, \qquad f^{(2,2)} = 2, \qquad f^{(2,1,1)} = 3, \qquad f^{(1,1,1,1)} = 1.$$

Hence, Corollary 5.12.18 says that $1^2 + 3^2 + 2^2 + 3^2 + 1^2 = 4!$. And this is indeed true.

*Proof of Corollary 5.12.18.* Let $\mathbf{k} = \mathbb{Q}$. Then, Theorem 5.12.17 yields

$$1 = \sum_{\lambda \text{ is a partition of } n} \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}. \tag{249}$$

Now, $[1]\, 1 = 1$, so that

$$\begin{aligned} 1 = [1]\, 1 &= [1] \sum_{\lambda \text{ is a partition of } n} \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2} \qquad \text{(by (249))} \\ &= \sum_{\lambda \text{ is a partition of } n} \frac{[1]\, \mathbf{E}_\lambda}{\left(h^\lambda\right)^2}. \end{aligned} \tag{250}$$

However, each partition $\lambda$ of $n$ satisfies $[1]\, \mathbf{E}_\lambda = n!$ (by Lemma 5.12.7) and thus

$$\begin{aligned} \frac{[1]\, \mathbf{E}_\lambda}{\left(h^\lambda\right)^2} &= \frac{n!}{\left(h^\lambda\right)^2} = \frac{n!}{\left(n!/f^\lambda\right)^2} \qquad \left(\text{since } h^\lambda = \frac{n!}{f^\lambda} = n!/f^\lambda\right) \\ &= \frac{\left(f^\lambda\right)^2}{n!}. \end{aligned}$$

Thus, we can rewrite (250) as

$$1 = \sum_{\lambda \text{ is a partition of } n} \frac{\left(f^\lambda\right)^2}{n!}.$$

Multiplying both sides of this equality by $n!$, we find

$$n! = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2.$$

This proves Corollary 5.12.18. □

Corollary 5.12.18 is a crucial result in the combinatorics of permutations, and can be proved without any algebra. Such proofs can be found in [Leeuwe11, Proposition 1.3.3], [Ruther48, page 26, first formula], [Stanle18, Corollary 8.5], [MenRem15, Exercise 5.7], [Fulton97, §4.3, (4)], [Aigner07, Theorem 8.26 a.], [Sagan01, Theorem 2.6.5 part 3.], [Prasad15, (3.12)], [Knuth98, §5.1.4, Theorem A] and many other sources. (See also [Stanle23, Corollary 7.12.6] for a proof using symmetric functions.)

### 5.12.6. Explicit expansions in centralized Young symmetrizers

In Theorem 5.12.4 **(a)**, we required that $\mathbf{k}$ be a field of characteristic 0. However, this requirement is stronger than necessary; it suffices to assume that $n!$ be invertible in $\mathbf{k}$. In order to show this, we need some more lemmas.

We try to expand an arbitrary element $\mathbf{a}$ of $Z(\mathbf{k}[S_n])$ as a $\mathbf{k}$-linear combination of the $\mathbf{E}_\lambda$ (when $n!$ is invertible in $\mathbf{k}$). We first do this in the case when $\mathbf{k}$ is a field of characteristic 0:

**Lemma 5.12.20.** Let $\mathbf{k}$ be a field of characteristic 0. Let $\mathbf{a} \in Z(\mathbf{k}[S_n])$. Then,

$$\mathbf{a} = \sum_{\lambda \text{ is a partition of } n} \frac{[1](\mathbf{E}_\lambda \mathbf{a})}{\left(h^\lambda\right)^2 \cdot n!} \mathbf{E}_\lambda.$$

*Proof.* Theorem 5.12.9 tells us that the family $\left(\mathbf{E}_\mu\right)_{\mu \text{ is a partition of } n}$ is a basis of the $\mathbf{k}$-vector space $Z(\mathbf{k}[S_n])$. Thus, in particular, this family spans $Z(\mathbf{k}[S_n])$. Hence, the vector $\mathbf{a}$ (which belongs to $Z(\mathbf{k}[S_n])$) is a linear combination of this family. In other words, we can write $\mathbf{a}$ in the form

$$\mathbf{a} = \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\mu \qquad \text{for some scalars } \alpha_\mu \in \mathbf{k}.$$

Consider these scalars $\alpha_\mu$.

Let $\lambda$ be a partition of $n$. Multiplying the equality $\mathbf{a} = \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\mu$ by $\mathbf{E}_\lambda$ from the left, we get

$$\mathbf{E}_\lambda \mathbf{a} = \mathbf{E}_\lambda \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\mu = \alpha_\lambda \mathbf{E}_\lambda^2$$

(by the same computation that we used to prove $\mathbf{E}_\lambda \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\mu = \alpha_\lambda \mathbf{E}_\lambda^2$ during our above proof of Proposition 5.12.10). However, Theorem 5.12.14 yields $\mathbf{E}_\lambda^2 = \left(h^\lambda\right)^2 \mathbf{E}_\lambda$. Altogether, we thus find

$$\mathbf{E}_\lambda \mathbf{a} = \alpha_\lambda \underbrace{\mathbf{E}_\lambda^2}_{=\left(h^\lambda\right)^2 \mathbf{E}_\lambda} = \alpha_\lambda \left(h^\lambda\right)^2 \mathbf{E}_\lambda.$$

Hence,

$$[1]\left(\mathbf{E}_\lambda \mathbf{a}\right) = [1]\left(\alpha_\lambda \left(h^\lambda\right)^2 \mathbf{E}_\lambda\right) = \alpha_\lambda \left(h^\lambda\right)^2 \cdot \underbrace{[1]\,\mathbf{E}_\lambda}_{=n! \text{ (by Lemma 5.12.7)}} = \alpha_\lambda \left(h^\lambda\right)^2 \cdot n!.$$

Solving this equality for $\alpha_\lambda$ (since $\left(h^\lambda\right)^2 \cdot n!$ is a nonzero element of the field $\mathbf{k}$), we obtain

$$\alpha_\lambda = \frac{[1]\left(\mathbf{E}_\lambda \mathbf{a}\right)}{\left(h^\lambda\right)^2 \cdot n!}. \tag{251}$$

Now, forget that we fixed $\lambda$. We thus have proved (251) for each partition $\lambda$ of $n$.

Now, recall that

$$\begin{aligned}\mathbf{a} &= \sum_{\mu \text{ is a partition of } n} \alpha_\mu \mathbf{E}_\mu = \sum_{\lambda \text{ is a partition of } n} \underbrace{\alpha_\lambda}_{=\frac{[1]\left(\mathbf{E}_\lambda \mathbf{a}\right)}{\left(h^\lambda\right)^2 \cdot n!} \text{ (by (251))}} \mathbf{E}_\lambda \\ &= \sum_{\lambda \text{ is a partition of } n} \frac{[1]\left(\mathbf{E}_\lambda \mathbf{a}\right)}{\left(h^\lambda\right)^2 \cdot n!} \mathbf{E}_\lambda.\end{aligned}$$

This proves Lemma 5.12.20. □

It stands to reason that Lemma 5.12.20 should hold more generally for any commutative ring $\mathbf{k}$ in which $n!$ is invertible (since all denominators are easily seen to divide $n!^3$). This is indeed the case. To show this, we first consider conjugacy class sums over an arbitrary commutative ring $\mathbf{k}$:

**Lemma 5.12.21.** Let $C$ be a conjugacy class of $S_n$. Let $\mathbf{z}_C := \sum_{c \in C} c \in \mathbf{k}\left[S_n\right]$. Then,

$$n!^3 \cdot \mathbf{z}_C = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1]\left(\mathbf{E}_\lambda \mathbf{z}_C\right) \cdot \mathbf{E}_\lambda.$$

*Proof. Step 1:* In order to prove Lemma 5.12.21 for all commutative rings $\mathbf{k}$, it suffices to prove it for $\mathbf{k} = \mathbb{Z}$.

Indeed, the same base change argument that we used in the proof of Theorem 5.11.3 can be used here to reduce the general case to the case of $\mathbf{k} = \mathbb{Z}$. Namely, we consider the canonical ring morphism $f : \mathbb{Z} \to \mathbf{k}$ and the base change morphism $f_* : \mathbb{Z}[S_n] \to \mathbf{k}[S_n]$ corresponding to $f$. Then, $f_*$ sends $\mathbf{z}_C \in \mathbb{Z}[S_n]$ to $\mathbf{z}_C \in \mathbf{k}[S_n]$ and sends $\mathbf{E}_\lambda \in \mathbb{Z}[S_n]$ to $\mathbf{E}_\lambda \in \mathbf{k}[S_n]$. Hence, $f_*$ sends the product $\mathbf{E}_\lambda \mathbf{z}_C \in \mathbb{Z}[S_n]$ to the corresponding product $\mathbf{E}_\lambda \mathbf{z}_C \in \mathbf{k}[S_n]$. Thus, the coefficient $[1](\mathbf{E}_\lambda \mathbf{z}_C)$ computed using $\mathbb{Z}[S_n]$ is transformed into the coefficient $[1](\mathbf{E}_\lambda \mathbf{z}_C)$ computed using $\mathbf{k}[S_n]$ by the morphism $f$. Thus, if the equality

$$n!^3 \cdot \mathbf{z}_C = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1](\mathbf{E}_\lambda \mathbf{z}_C) \cdot \mathbf{E}_\lambda \qquad \text{in } \mathbb{Z}[S_n]$$

has been proved, then we can apply the base change morphism $f_*$ to this equality, and obtain the analogous equality in $\mathbf{k}[S_n]$. In other words, if Lemma 5.12.21 has been proved for $\mathbf{k} = \mathbb{Z}$, then it holds for all $\mathbf{k}$.

*Step 2:* In order to prove Lemma 5.12.21 for $\mathbf{k} = \mathbb{Z}$, it suffices to prove it for $\mathbf{k} = \mathbb{Q}$.

Indeed, this too follows from a base change argument, except that we apply the base change morphism backwards now. Namely, the canonical ring morphism $f : \mathbb{Z} \to \mathbb{Q}$ is injective (indeed, it is just the inclusion map), and thus the base change morphism $f_* : \mathbb{Z}[S_n] \to \mathbb{Q}[S_n]$ corresponding to $f$ is injective as well. However, as we already saw in Step 1, the equality

$$n!^3 \cdot \mathbf{z}_C = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1](\mathbf{E}_\lambda \mathbf{z}_C) \cdot \mathbf{E}_\lambda \qquad \text{in } \mathbb{Q}[S_n]$$

can be obtained from the equality

$$n!^3 \cdot \mathbf{z}_C = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1](\mathbf{E}_\lambda \mathbf{z}_C) \cdot \mathbf{E}_\lambda \qquad \text{in } \mathbb{Z}[S_n]$$

by applying the base change morphism $f_* : \mathbb{Z}[S_n] \to \mathbb{Q}[S_n]$. Thus, conversely, the latter equality can be obtained from the former equality by "un-applying" this morphism $f_*$ (which is allowed, since $f_*$ is injective). Hence, if the former equality holds, then so does the latter equality. In other words, if Lemma 5.12.21 has been proved for $\mathbf{k} = \mathbb{Q}$, then it holds for $\mathbf{k} = \mathbb{Z}$ and therefore for all $\mathbf{k}$ (as we saw in Step 1).

*Step 3:* It thus remains to prove Lemma 5.12.21 for $\mathbf{k} = \mathbb{Q}$.

So let us do this. Let $\mathbf{k} = \mathbb{Q}$. Then, $\mathbf{k}$ is a field of characteristic 0.

Theorem 3.9.5 (applied to $G = S_n$) shows that the center $Z(\mathbf{k}[S_n])$ of the group algebra $\mathbf{k}[S_n]$ is the $\mathbf{k}$-linear span of the conjugacy class sums. Hence, in particular, each conjugacy class sum of $S_n$ belongs to $Z(\mathbf{k}[S_n])$.

The element $\mathbf{z}_C = \sum_{c \in C} c$ is a conjugacy class sum (since it is the sum of all permutations in the conjugacy class $C$). Thus, it belongs to $Z(\mathbf{k}[S_n])$ (since each conjugacy class sum of $S_n$ belongs to $Z(\mathbf{k}[S_n])$). Hence, Lemma 5.12.20 (applied to $\mathbf{a} = \mathbf{z}_C$) yields

$$\mathbf{z}_C = \sum_{\lambda \text{ is a partition of } n} \frac{[1](\mathbf{E}_\lambda \mathbf{z}_C)}{(h^\lambda)^2 \cdot n!} \mathbf{E}_\lambda.$$

Multiplying both sides of this equality by $n!^3$, we obtain

$$\begin{aligned} n!^3 \cdot \mathbf{z}_C &= n!^3 \cdot \sum_{\lambda \text{ is a partition of } n} \frac{[1](\mathbf{E}_\lambda \mathbf{z}_C)}{(h^\lambda)^2 \cdot n!} \mathbf{E}_\lambda \\ &= \sum_{\lambda \text{ is a partition of } n} \underbrace{n!^3 \cdot \frac{[1](\mathbf{E}_\lambda \mathbf{z}_C)}{(h^\lambda)^2 \cdot n!}}_{\substack{= \left(\frac{n!}{h^\lambda}\right)^2 \cdot [1](\mathbf{E}_\lambda \mathbf{z}_C) \\ = \left(f^\lambda\right)^2 \cdot [1](\mathbf{E}_\lambda \mathbf{z}_C) \\ \left(\text{since } \frac{n!}{h^\lambda} = f^\lambda \text{ (because } h^\lambda = \frac{n!}{f^\lambda}\right)\right)}} \mathbf{E}_\lambda \\ &= \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1](\mathbf{E}_\lambda \mathbf{z}_C) \cdot \mathbf{E}_\lambda. \end{aligned}$$

Thus, Lemma 5.12.21 is proved for $\mathbf{k} = \mathbb{Q}$. As we explained above, this completes the proof of Lemma 5.12.21 for all $\mathbf{k}$. □

We can now extend Lemma 5.12.20 to all commutative rings $\mathbf{k}$ (upon multiplication by $n!^3$ in order to clear all denominators):

**Proposition 5.12.22.** Let $\mathbf{a} \in Z(\mathbf{k}[S_n])$. Then,

$$n!^3 \cdot \mathbf{a} = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1](\mathbf{E}_\lambda \mathbf{a}) \cdot \mathbf{E}_\lambda.$$

*Proof.* Theorem 3.9.5 (applied to $G = S_n$) shows that the center $Z(\mathbf{k}[S_n])$ of the group algebra $\mathbf{k}[S_n]$ is the $\mathbf{k}$-linear span of the conjugacy class sums. In other words, the conjugacy class sums span $Z(\mathbf{k}[S_n])$.

We need to prove the equality

$$n!^3 \cdot \mathbf{a} = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1](\mathbf{E}_\lambda \mathbf{a}) \cdot \mathbf{E}_\lambda.$$

Both sides of this equality are $\mathbf{k}$-linear in $\mathbf{a}$. Hence, in proving this equality, we can WLOG assume that $\mathbf{a}$ is a conjugacy class sum (since the conjugacy class

sums span $Z(\mathbf{k}[S_n])$). Assume this. Thus, $\mathbf{a} = \mathbf{z}_C$ for some conjugacy class $C$ of $S_n$ (where $\mathbf{z}_C$ is as defined in Definition 3.9.4 **(c)**). Consider this $C$. Thus, $\mathbf{z}_C = \sum_{c \in C} c$ (by the definition of $\mathbf{z}_C$). Hence, Lemma 5.12.21 yields

$$n!^3 \cdot \mathbf{z}_C = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1]\left(\mathbf{E}_\lambda \mathbf{z}_C\right) \cdot \mathbf{E}_\lambda.$$

Since $\mathbf{a} = \mathbf{z}_C$, we can rewrite this as

$$n!^3 \cdot \mathbf{a} = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1]\left(\mathbf{E}_\lambda \mathbf{a}\right) \cdot \mathbf{E}_\lambda.$$

Thus, Proposition 5.12.22 is proved. □

As a consequence, we can now generalize Theorem 5.12.4 **(a)** to all commutative rings $\mathbf{k}$ in which $n!$ is invertible:

**Theorem 5.12.23.** Assume that $n!$ is invertible in $\mathbf{k}$. Then, the family

$$\left(\mathbf{E}_\lambda\right)_{\lambda \text{ is a partition of } n}$$

is a basis of the $\mathbf{k}$-module $Z(\mathbf{k}[S_n])$.

*Proof.* Let $P$ be the set of all partitions of $n$. Thus, we must show that the family $\left(\mathbf{E}_\lambda\right)_{\lambda \in P}$ is a basis of the $\mathbf{k}$-module $Z(\mathbf{k}[S_n])$.

Clearly, this family really consists of vectors in $Z(\mathbf{k}[S_n])$, since Proposition 5.12.3 shows that each $\lambda \in P$ satisfies $\mathbf{E}_\lambda \in Z(\mathbf{k}[S_n])$. Thus, it remains to prove the following two claims:

*Claim 1:* The family $\left(\mathbf{E}_\lambda\right)_{\lambda \in P}$ is $\mathbf{k}$-linearly independent.

*Claim 2:* The family $\left(\mathbf{E}_\lambda\right)_{\lambda \in P}$ spans the $\mathbf{k}$-module $Z(\mathbf{k}[S_n])$.

*Proof of Claim 1.* This proof is mostly analogous to the proof of Claim 1 during the above proof of Theorem 5.12.9. The only part that needs to be changed is the argument that we use to derive $\alpha_\lambda = 0$ from $0 = \alpha_\lambda \mathbf{E}_\lambda^2$ (since $\mathbf{k}$ is no longer assumed to be a field). This part can now be done as follows: We have $f^\lambda h^\lambda = n!$ (since $h^\lambda = \dfrac{n!}{f^\lambda}$). Thus, the number $h^\lambda$ is invertible in $\mathbf{k}$ (since $f^\lambda h^\lambda = n!$ is invertible in $\mathbf{k}$). Hence, its square $\left(h^\lambda\right)^2$ is invertible in $\mathbf{k}$ as well. But we also have

$$0 = \alpha_\lambda \underbrace{\mathbf{E}_\lambda^2}_{= \left(h^\lambda\right)^2 \mathbf{E}_\lambda \text{ (by Theorem 5.12.14)}} = \alpha_\lambda \left(h^\lambda\right)^2 \mathbf{E}_\lambda.$$

We can divide this equality by $\left(h^\lambda\right)^2$ (since $\left(h^\lambda\right)^2$ is invertible in $\mathbf{k}$), and thus obtain $0 = \alpha_\lambda \mathbf{E}_\lambda$. Hence,

$$[1]\,0 = [1]\,(\alpha_\lambda \mathbf{E}_\lambda) = \alpha_\lambda \cdot \underbrace{[1]\,\mathbf{E}_\lambda}_{\substack{=n! \\ \text{(by Lemma 5.12.7)}}} = \alpha_\lambda \cdot n!.$$

In other words, $\alpha_\lambda \cdot n! = [1]\,0 = 0$. We can divide this equality by $n!$ (since $n!$ is invertible in $\mathbf{k}$), and obtain $\alpha_\lambda = 0$. Thus, $\alpha_\lambda = 0$ has been derived from $0 = \alpha_\lambda \mathbf{E}_\lambda^2$. As we said above, the rest of the proof of Claim 1 proceeds precisely as in the proof of Theorem 5.12.9. □

*Proof of Claim 2.* Let $\mathbf{a} \in Z\left(\mathbf{k}\left[S_n\right]\right)$. Then, Proposition 5.12.22 yields

$$n!^3 \cdot \mathbf{a} = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1]\,(\mathbf{E}_\lambda \mathbf{a}) \cdot \mathbf{E}_\lambda.$$

We can divide this equality by $n!^3$ (since $n!$ and therefore also $n!^3$ is invertible in $\mathbf{k}$), and obtain

$$\begin{aligned}
\mathbf{a} &= \frac{1}{n!^3} \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot [1]\,(\mathbf{E}_\lambda \mathbf{a}) \cdot \mathbf{E}_\lambda \\
&= \sum_{\lambda \text{ is a partition of } n} \frac{\left(f^\lambda\right)^2 \cdot [1]\,(\mathbf{E}_\lambda \mathbf{a})}{n!^3} \cdot \mathbf{E}_\lambda \\
&= \sum_{\lambda \in P} \frac{\left(f^\lambda\right)^2 \cdot [1]\,(\mathbf{E}_\lambda \mathbf{a})}{n!^3} \cdot \mathbf{E}_\lambda
\end{aligned}$$

(since the set of all partitions of $n$ is $P$). This shows that $\mathbf{a}$ is a $\mathbf{k}$-linear combination of the family $\left(\mathbf{E}_\lambda\right)_{\lambda \in P}$.

Forget that we fixed $\mathbf{a}$. We thus have shown that each $\mathbf{a} \in Z\left(\mathbf{k}\left[S_n\right]\right)$ is a $\mathbf{k}$-linear combination of the family $\left(\mathbf{E}_\lambda\right)_{\lambda \in P}$. In other words, the family $\left(\mathbf{E}_\lambda\right)_{\lambda \in P}$ spans the $\mathbf{k}$-module $Z\left(\mathbf{k}\left[S_n\right]\right)$. This proves Claim 2. □

Combining Claim 1 with Claim 2, we conclude that the family $\left(\mathbf{E}_\lambda\right)_{\lambda \in P}$ is a basis of $Z\left(\mathbf{k}\left[S_n\right]\right)$. As we said, this proves Theorem 5.12.23. □

Proposition 5.12.22 also allows us to extend Theorem 5.12.17 to a larger class of rings $\mathbf{k}$:

**Corollary 5.12.24.** Assume that $n!$ is invertible in $\mathbf{k}$. Then:

**(a)** For any partition $\lambda$ of $n$, the number $\left(h^\lambda\right)^2$ is invertible in $\mathbf{k}$.

**(b)** In $\mathbf{k}\left[S_n\right]$, we have

$$1 = \sum_{\lambda \text{ is a partition of } n} \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}.$$

*Proof.* **(a)** Let $\lambda$ be a partition of $n$. Then, the definition of $h^\lambda$ yields $h^\lambda = \dfrac{n!}{f^\lambda} \mid n!$, so that $h^\lambda$ is invertible in $\mathbf{k}$ (since $n!$ is invertible in $\mathbf{k}$). Therefore, $\left(h^\lambda\right)^2$ is also invertible in $\mathbf{k}$. This proves Corollary 5.12.24 **(a)**.

**(b)** Corollary 5.12.24 **(a)** shows that $\left(h^\lambda\right)^2$ is invertible in $\mathbf{k}$ for any partition $\lambda$ of $n$. Thus, the sum $\sum\limits_{\lambda \text{ is a partition of } n} \dfrac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}$ is well-defined.

We have $1 \in Z\left(\mathbf{k}\left[S_n\right]\right)$. Thus, Proposition 5.12.22 (applied to $\mathbf{a} = 1$) yields

$$\begin{aligned} n!^3 \cdot 1 &= \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot \underbrace{\left[1\right]\left(\mathbf{E}_\lambda 1\right)}_{\substack{=[1]\mathbf{E}_\lambda = n! \\ \text{(by Lemma 5.12.7)}}} \cdot \mathbf{E}_\lambda \\ &= \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot n! \cdot \mathbf{E}_\lambda. \end{aligned}$$

Dividing this equality by $n!^3$, we find

$$\begin{aligned} 1 &= \frac{1}{n!^3} \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \cdot n! \cdot \mathbf{E}_\lambda \\ &= \sum_{\lambda \text{ is a partition of } n} \underbrace{\frac{1}{n!^3}\left(f^\lambda\right)^2 \cdot n!}_{\substack{=\left(f^\lambda/n!\right)^2 = 1/\left(h^\lambda\right)^2 \\ \text{(since } h^\lambda = n!/f^\lambda \text{ and} \\ \text{thus } 1/\left(h^\lambda\right)^2 = 1/\left(n!/f^\lambda\right)^2 = \left(f^\lambda/n!\right)^2\text{)}}} \cdot \mathbf{E}_\lambda \\ &= \sum_{\lambda \text{ is a partition of } n} \left(1/\left(h^\lambda\right)^2\right) \cdot \mathbf{E}_\lambda = \sum_{\lambda \text{ is a partition of } n} \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}. \end{aligned}$$

This proves Corollary 5.12.24 **(b)**. □

**Exercise 5.12.1.** [2] Prove that $n!^2 = \sum\limits_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 \mathbf{E}_\lambda$ in $\mathbf{k}\left[S_n\right]$ (for arbitrary $\mathbf{k}$).

[Be careful not to divide by $n!$ without justification, since $n!$ might not be invertible.]

Combining Theorem 5.12.14, Corollary 5.12.24 and Proposition 5.12.6, we obtain a decomposition of $\mathcal{A} = \mathbf{k}\left[S_n\right]$ as a left $\mathcal{A}$-module:

**Corollary 5.12.25.** Assume that $n!$ is invertible in $\mathbf{k}$. Then,

$$\mathcal{A} = \bigoplus_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda \qquad \text{(an internal direct sum)}$$

as left $\mathcal{A}$-modules.

*Proof.* Let $P$ be the set of all partitions of $n$. We must then show that $\mathcal{A} = \bigoplus_{\lambda \in P} \mathcal{A}\mathbf{E}_\lambda$ (an internal direct sum).

First, we shall show that the sum $\sum_{\lambda \in P} \mathcal{A}\mathbf{E}_\lambda$ is direct.

Indeed, let $(\mathbf{a}_\lambda)_{\lambda \in P} \in \prod_{\lambda \in P} \mathcal{A}\mathbf{E}_\lambda$ be a family of scalars such that $\sum_{\lambda \in P} \mathbf{a}_\lambda = 0$. We shall show that $\mathbf{a}_\lambda = 0$ for all $\lambda \in P$.

Indeed, we have $(\mathbf{a}_\lambda)_{\lambda \in P} \in \prod_{\lambda \in P} \mathcal{A}\mathbf{E}_\lambda$. In other words,

$$\mathbf{a}_\lambda \in \mathcal{A}\mathbf{E}_\lambda \qquad \text{for each } \lambda \in P. \tag{252}$$

Now, let $\mu \in P$ be arbitrary. Then, $\mu$ is a partition of $n$ (since $P$ is the set of all partitions of $n$). Thus, Corollary 5.12.24 **(a)** (applied to $\lambda = \mu$) shows that the number $(h^\mu)^2$ is invertible in $\mathbf{k}$. We have

$$\sum_{\lambda \in P} \mathbf{a}_\lambda \mathbf{E}_\mu = \mathbf{a}_\mu \mathbf{E}_\mu + \sum_{\substack{\lambda \in P; \\ \lambda \neq \mu}} \mathbf{a}_\lambda \mathbf{E}_\mu \tag{253}$$

(here, we have split off the addend for $\lambda = \mu$ from the sum).

We have $\mathbf{a}_\mu \in \mathcal{A}\mathbf{E}_\mu$ (by (252), applied to $\lambda = \mu$). In other words, $\mathbf{a}_\mu = \mathbf{c}\mathbf{E}_\mu$ for some $\mathbf{c} \in \mathcal{A}$. Consider this $\mathbf{c}$. Then,

$$\begin{aligned} \underbrace{\mathbf{a}_\mu}_{=\mathbf{c}\mathbf{E}_\mu} \mathbf{E}_\mu = \mathbf{c} \underbrace{\mathbf{E}_\mu \mathbf{E}_\mu}_{\substack{=\mathbf{E}_\mu^2 = (h^\mu)^2 \mathbf{E}_\mu \\ \text{(by Theorem 5.12.14,} \\ \text{applied to } \lambda = \mu\text{)}}} &= \mathbf{c}\,(h^\mu)^2\,\mathbf{E}_\mu = (h^\mu)^2 \underbrace{\mathbf{c}\mathbf{E}_\mu}_{=\mathbf{a}_\mu} \\ &= (h^\mu)^2\,\mathbf{a}_\mu. \end{aligned} \tag{254}$$

For any $\lambda \in P$ satisfying $\lambda \neq \mu$, we have $\mathbf{a}_\lambda \mathbf{E}_\mu = 0$ (this follows easily from Proposition 5.12.6)[97]. Hence, $\sum_{\substack{\lambda \in P; \\ \lambda \neq \mu}} \underbrace{\mathbf{a}_\lambda \mathbf{E}_\mu}_{=0} = \sum_{\substack{\lambda \in P; \\ \lambda \neq \mu}} 0 = 0$. Thus, (253) becomes

$$\sum_{\lambda \in P} \mathbf{a}_\lambda \mathbf{E}_\mu = \underbrace{\mathbf{a}_\mu \mathbf{E}_\mu}_{\substack{=(h^\mu)^2 \mathbf{a}_\mu \\ \text{(by (254))}}} + \underbrace{\sum_{\substack{\lambda \in P; \\ \lambda \neq \mu}} \mathbf{a}_\lambda \mathbf{E}_\mu}_{=0} = (h^\mu)^2\,\mathbf{a}_\mu.$$

Hence,

$$(h^\mu)^2\,\mathbf{a}_\mu = \sum_{\lambda \in P} \mathbf{a}_\lambda \mathbf{E}_\mu = \underbrace{\left( \sum_{\lambda \in P} \mathbf{a}_\lambda \right)}_{=0} \mathbf{E}_\mu = 0.$$

[97] *Proof.* Let $\lambda \in P$ be such that $\lambda \neq \mu$. Then, $\lambda, \mu \in P$. In other words, $\lambda$ and $\mu$ are two partitions of $n$ (since $P$ is the set of all partitions of $n$). Moreover, $\lambda$ and $\mu$ are distinct (since $\lambda \neq \mu$). Thus, Proposition 5.12.6 yields $\mathbf{E}_\lambda \mathbf{E}_\mu = 0$. However, (252) yields $\mathbf{a}_\lambda \in \mathcal{A}\mathbf{E}_\lambda$. In other words, $\mathbf{a}_\lambda = \mathbf{b}\mathbf{E}_\lambda$ for some $\mathbf{b} \in \mathcal{A}$. Consider this $\mathbf{b}$. Then, $\underbrace{\mathbf{a}_\lambda}_{=\mathbf{b}\mathbf{E}_\lambda} \mathbf{E}_\mu = \mathbf{b} \underbrace{\mathbf{E}_\lambda \mathbf{E}_\mu}_{=0} = 0$, qed.

We can divide this equality by $(h^\mu)^2$ (since $(h^\mu)^2$ is invertible in $\mathbf{k}$), and thus obtain $\mathbf{a}_\mu = 0$.

Forget that we fixed $\mu$. We thus have shown that $\mathbf{a}_\mu = 0$ for each $\mu \in P$. In other words, $\mathbf{a}_\lambda = 0$ for each $\lambda \in P$.

Forget that we fixed $(\mathbf{a}_\lambda)_{\lambda\in P}$. We thus have shown that if $(\mathbf{a}_\lambda)_{\lambda\in P} \in \prod_{\lambda\in P} \mathcal{A}\mathbf{E}_\lambda$ is a family of elements such that $\sum_{\lambda\in P} \mathbf{a}_\lambda = 0$, then we have $\mathbf{a}_\lambda = 0$ for all $\lambda \in P$. In other words, the sum $\sum_{\lambda\in P} \mathcal{A}\mathbf{E}_\lambda$ is direct. Thus, we can write it as $\bigoplus_{\lambda\in P} \mathcal{A}\mathbf{E}_\lambda$ (an internal direct sum).

We shall now show that this sum is $\mathcal{A}$. Indeed, Corollary 5.12.24 **(b)** yields

$$1 = \underbrace{\sum_{\lambda \text{ is a partition of } n}}_{\substack{=\sum_{\lambda\in P} \\ \text{(by the definition of } P\text{)}}} \frac{\mathbf{E}_\lambda}{(h^\lambda)^2} = \sum_{\lambda\in P} \frac{\mathbf{E}_\lambda}{(h^\lambda)^2}. \tag{255}$$

Hence, for each $\mathbf{a} \in \mathcal{A}$, we have

$$\begin{aligned}
\mathbf{a} = \mathbf{a}\cdot 1 &= \mathbf{a}\cdot \sum_{\lambda\in P} \frac{\mathbf{E}_\lambda}{(h^\lambda)^2} \qquad \text{(by (255))} \\
&= \sum_{\lambda\in P} \underbrace{\frac{\mathbf{a}\cdot \mathbf{E}_\lambda}{(h^\lambda)^2}}_{=\frac{\mathbf{a}}{(h^\lambda)^2}\mathbf{E}_\lambda} = \sum_{\lambda\in P} \underbrace{\frac{\mathbf{a}}{(h^\lambda)^2}}_{\in\mathcal{A}} \mathbf{E}_\lambda \in \sum_{\lambda\in P} \mathcal{A}\mathbf{E}_\lambda.
\end{aligned}$$

In other words, $\mathcal{A} \subseteq \sum_{\lambda\in P} \mathcal{A}\mathbf{E}_\lambda$. Since $\sum_{\lambda\in P} \mathcal{A}\mathbf{E}_\lambda$ is clearly a subset of $\mathcal{A}$, we thus obtain

$$\begin{aligned}
\mathcal{A} &= \sum_{\lambda\in P} \mathcal{A}\mathbf{E}_\lambda = \bigoplus_{\lambda\in P} \mathcal{A}\mathbf{E}_\lambda \qquad \text{(as we saw above)} \\
&= \bigoplus_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda \qquad \text{(since } P \text{ is the set of all partitions of } n\text{)}.
\end{aligned}$$

This proves Corollary 5.12.25. □

### 5.12.7. The structure of $Z(\mathbf{k}[S_n])$

Now we are ready to describe $Z(\mathbf{k}[S_n])$ as a $\mathbf{k}$-algebra when $n!$ is invertible in $\mathbf{k}$ (and thus prove Theorem 3.15.1 **(b)**):

**Theorem 5.12.26.** Assume that $n!$ is invertible in $\mathbf{k}$. Then, there exists a $\mathbf{k}$-algebra isomorphism

$$\mathbf{k}^{\{\text{partitions of } n\}} \to Z\left(\mathbf{k}\left[S_n\right]\right),$$
$$\left(\beta_\lambda\right)_{\lambda \text{ is a partition of } n} \mapsto \sum_{\lambda \text{ is a partition of } n} \beta_\lambda \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}.$$

(Here, $\mathbf{k}^{\{\text{partitions of } n\}}$ is a direct product of copies of $\mathbf{k}$, one copy for each partition of $n$. The addition and the multiplication on this direct product are entrywise, as usual.)

*Proof.* We begin with some simple observations. Theorem 5.12.23 shows that the family

$$\left(\mathbf{E}_\lambda\right)_{\lambda \text{ is a partition of } n}$$

is a basis of the $\mathbf{k}$-module $Z\left(\mathbf{k}\left[S_n\right]\right)$. Hence, this family is $\mathbf{k}$-linearly independent and spans $Z\left(\mathbf{k}\left[S_n\right]\right)$.

Next, we claim the following:

*Claim 1:* For any family $\left(\beta_\lambda\right)_{\lambda \text{ is a partition of } n} \in \mathbf{k}^{\{\text{partitions of } n\}}$, the sum $\sum_{\lambda \text{ is a partition of } n} \beta_\lambda \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}$ is a well-defined element of $Z\left(\mathbf{k}\left[S_n\right]\right)$.

*Proof of Claim 1.* It clearly suffices to prove that $\frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}$ is a well-defined element of $Z\left(\mathbf{k}\left[S_n\right]\right)$ whenever $\lambda$ is a partition of $n$.

So let us do this. Let $\lambda$ be a partition of $n$. Then, Proposition 5.12.3 shows that $\mathbf{E}_\lambda$ belongs to $Z\left(\mathbf{k}\left[S_n\right]\right)$. Thus, $\frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}$ is a well-defined element of $Z\left(\mathbf{k}\left[S_n\right]\right)$ (since Corollary 5.12.24 **(a)** shows that $\left(h^\lambda\right)^2$ is invertible in $\mathbf{k}$). This completes our proof of Claim 1. □

Claim 1 allows us to define a map

$$\Phi : \mathbf{k}^{\{\text{partitions of } n\}} \to Z\left(\mathbf{k}\left[S_n\right]\right),$$
$$\left(\beta_\lambda\right)_{\lambda \text{ is a partition of } n} \mapsto \sum_{\lambda \text{ is a partition of } n} \beta_\lambda \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}.$$

Consider this map $\Phi$. It is clearly $\mathbf{k}$-linear. Our goal is now to show that $\Phi$ is a $\mathbf{k}$-algebra isomorphism (since this is precisely what Theorem 5.12.26 claims). We shall achieve this by proving the following four claims:

*Claim 2:* The map $\Phi$ is injective.

*Claim 3:* The map $\Phi$ is surjective.

*Claim 4:* The map $\Phi$ respects multiplication (i.e., we have $\Phi(\beta\gamma) = \Phi(\beta)\Phi(\gamma)$ for all $\beta, \gamma \in \mathbf{k}^{\{\text{partitions of } n\}}$).

*Claim 5:* The map $\Phi$ respects the unity (i.e., we have $\Phi\left(1_{\mathbf{k}^{\{\text{partitions of } n\}}}\right) = 1$).

*Proof of Claim 2.* Let $\beta \in \operatorname{Ker}\Phi$. We shall show that $\beta = 0$.

We have $\beta \in \operatorname{Ker}\Phi \subseteq \mathbf{k}^{\{\text{partitions of } n\}}$. Therefore, write $\beta$ as $\beta = (\beta_\lambda)_{\lambda \text{ is a partition of } n}$ for some scalars $\beta_\lambda \in \mathbf{k}$. Then, the definition of $\Phi$ yields

$$\Phi(\beta) = \sum_{\lambda \text{ is a partition of } n} \underbrace{\beta_\lambda \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}}_{= \frac{\beta_\lambda}{\left(h^\lambda\right)^2}\mathbf{E}_\lambda} = \sum_{\lambda \text{ is a partition of } n} \frac{\beta_\lambda}{\left(h^\lambda\right)^2}\mathbf{E}_\lambda.$$

Hence,

$$\sum_{\lambda \text{ is a partition of } n} \frac{\beta_\lambda}{\left(h^\lambda\right)^2}\mathbf{E}_\lambda = \Phi(\beta) = 0 \qquad (\text{since } \beta \in \operatorname{Ker}\Phi).$$

From this equality, we conclude that

$$\frac{\beta_\lambda}{\left(h^\lambda\right)^2} = 0 \qquad \text{for all partitions } \lambda \text{ of } n$$

(since the family $(\mathbf{E}_\lambda)_{\lambda \text{ is a partition of } n}$ is $\mathbf{k}$-linearly independent). Multiplying this equality by $\left(h^\lambda\right)^2$, we find that

$$\beta_\lambda = 0 \qquad \text{for all partitions } \lambda \text{ of } n.$$

In other words, $(\beta_\lambda)_{\lambda \text{ is a partition of } n} = (0)_{\lambda \text{ is a partition of } n} = 0$. Hence, $\beta = (\beta_\lambda)_{\lambda \text{ is a partition of } n} = 0$.

Forget that we fixed $\beta$. We thus have shown that $\beta = 0$ for each $\beta \in \operatorname{Ker}\Phi$. In other words, $\operatorname{Ker}\Phi = 0$. Since $\Phi$ is a $\mathbf{k}$-linear map, this entails that $\Phi$ is injective. Thus, Claim 2 is proved. □

*Proof of Claim 3.* Let $\mathbf{a} \in Z(\mathbf{k}[S_n])$. Then, $\mathbf{a}$ is a $\mathbf{k}$-linear combination of the family $(\mathbf{E}_\lambda)_{\lambda \text{ is a partition of } n}$ (since this family spans $Z(\mathbf{k}[S_n])$). In other words, we can write $\mathbf{a}$ as $\mathbf{a} = \sum_{\lambda \text{ is a partition of } n} \gamma_\lambda \mathbf{E}_\lambda$ for some scalars $\gamma_\lambda \in \mathbf{k}$. Consider these scalars $\gamma_\lambda$. Now, the definition of $\Phi$ yields

$$\begin{aligned}\Phi\left(\left(\left(h^\lambda\right)^2\gamma_\lambda\right)_{\lambda \text{ is a partition of } n}\right) &= \sum_{\lambda \text{ is a partition of } n} \underbrace{\left(h^\lambda\right)^2\gamma_\lambda\frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}}_{=\gamma_\lambda\mathbf{E}_\lambda} \\ &= \sum_{\lambda \text{ is a partition of } n} \gamma_\lambda\mathbf{E}_\lambda = \mathbf{a},\end{aligned}$$

so that $\mathbf{a} = \Phi\left(\left(\left(h^\lambda\right)^2 \gamma_\lambda\right)_{\lambda \text{ is a partition of } n}\right) \in \operatorname{Im}\Phi$.

Forget that we fixed $\mathbf{a}$. We thus have shown that each $\mathbf{a} \in Z\left(\mathbf{k}\left[S_n\right]\right)$ satisfies $\mathbf{a} \in \operatorname{Im}\Phi$. In other words, the map $\Phi$ is surjective. This proves Claim 3. $\square$

*Proof of Claim 4.* Let $\beta, \gamma \in \mathbf{k}^{\{\text{partitions of } n\}}$. We must show that $\Phi(\beta\gamma) = \Phi(\beta)\Phi(\gamma)$.

Write $\beta$ as $\beta = (\beta_\lambda)_{\lambda \text{ is a partition of } n}$ for some scalars $\beta_\lambda \in \mathbf{k}$. Likewise, write $\gamma$ as $\gamma = (\gamma_\lambda)_{\lambda \text{ is a partition of } n}$ for some scalars $\gamma_\lambda \in \mathbf{k}$. Thus,

$$\beta\gamma = (\beta_\lambda)_{\lambda \text{ is a partition of } n} (\gamma_\lambda)_{\lambda \text{ is a partition of } n} = (\beta_\lambda \gamma_\lambda)_{\lambda \text{ is a partition of } n}.$$

Hence, the definition of $\Phi$ yields

$$\Phi(\beta\gamma) = \sum_{\lambda \text{ is a partition of } n} \beta_\lambda \gamma_\lambda \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}. \tag{256}$$

However, the definition of $\Phi$ also yields

$$\Phi(\beta) = \sum_{\lambda \text{ is a partition of } n} \beta_\lambda \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2} \qquad \left(\text{since } \beta = (\beta_\lambda)_{\lambda \text{ is a partition of } n}\right)$$

and

$$\begin{aligned}\Phi(\gamma) &= \sum_{\lambda \text{ is a partition of } n} \gamma_\lambda \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2} \qquad \left(\text{since } \gamma = (\gamma_\lambda)_{\lambda \text{ is a partition of } n}\right) \\ &= \sum_{\mu \text{ is a partition of } n} \gamma_\mu \frac{\mathbf{E}_\mu}{\left(h^\mu\right)^2}.\end{aligned}$$

Multiplying these two equalities, we obtain

$$\begin{aligned}
&\Phi(\beta)\,\Phi(\gamma) \\
&= \left( \sum_{\lambda \text{ is a partition of } n} \beta_\lambda \frac{\mathbf{E}_\lambda}{(h^\lambda)^2} \right) \left( \sum_{\mu \text{ is a partition of } n} \gamma_\mu \frac{\mathbf{E}_\mu}{(h^\mu)^2} \right) \\
&= \underbrace{\sum_{\lambda \text{ is a partition of } n} \sum_{\mu \text{ is a partition of } n}}_{= \sum\limits_{\substack{\lambda \text{ and } \mu \text{ are two} \\ \text{partitions of } n}}} \underbrace{\beta_\lambda \frac{\mathbf{E}_\lambda}{(h^\lambda)^2} \gamma_\mu \frac{\mathbf{E}_\mu}{(h^\mu)^2}}_{= \frac{\beta_\lambda \gamma_\mu}{(h^\lambda)^2 (h^\mu)^2} \mathbf{E}_\lambda \mathbf{E}_\mu} \\
&= \sum_{\substack{\lambda \text{ and } \mu \text{ are two} \\ \text{partitions of } n}} \frac{\beta_\lambda \gamma_\mu}{(h^\lambda)^2 (h^\mu)^2} \mathbf{E}_\lambda \mathbf{E}_\mu \\
&= \sum_{\substack{\lambda \text{ and } \mu \text{ are two} \\ \text{partitions of } n; \\ \lambda = \mu}} \frac{\beta_\lambda \gamma_\mu}{(h^\lambda)^2 (h^\mu)^2} \mathbf{E}_\lambda \mathbf{E}_\mu + \sum_{\substack{\lambda \text{ and } \mu \text{ are two} \\ \text{partitions of } n; \\ \lambda \neq \mu}} \frac{\beta_\lambda \gamma_\mu}{(h^\lambda)^2 (h^\mu)^2} \underbrace{\mathbf{E}_\lambda \mathbf{E}_\mu}_{=0} \\
&\qquad \text{(by Proposition 5.12.6)} \\
&= \sum_{\substack{\lambda \text{ and } \mu \text{ are two} \\ \text{partitions of } n; \\ \lambda = \mu}} \frac{\beta_\lambda \gamma_\mu}{(h^\lambda)^2 (h^\mu)^2} \mathbf{E}_\lambda \mathbf{E}_\mu + \underbrace{\sum_{\substack{\lambda \text{ and } \mu \text{ are two} \\ \text{partitions of } n; \\ \lambda \neq \mu}} \frac{\beta_\lambda \gamma_\mu}{(h^\lambda)^2 (h^\mu)^2} 0}_{=0} \\
&= \sum_{\substack{\lambda \text{ and } \mu \text{ are two} \\ \text{partitions of } n; \\ \lambda = \mu}} \frac{\beta_\lambda \gamma_\mu}{(h^\lambda)^2 (h^\mu)^2} \mathbf{E}_\lambda \mathbf{E}_\mu \\
&= \sum_{\lambda \text{ is a partition of } n} \frac{\beta_\lambda \gamma_\lambda}{(h^\lambda)^2 (h^\lambda)^2} \underbrace{\mathbf{E}_\lambda \mathbf{E}_\lambda}_{= \mathbf{E}_\lambda^2 = \left(h^\lambda\right)^2 \mathbf{E}_\lambda} \\
&\qquad \text{(by Theorem 5.12.14)} \\
&\qquad \left( \begin{array}{c} \text{here, we substituted } \lambda \text{ for } \mu \text{ in the sum,} \\ \text{because of the requirement } \lambda = \mu \end{array} \right) \\
&= \sum_{\lambda \text{ is a partition of } n} \underbrace{\frac{\beta_\lambda \gamma_\lambda}{(h^\lambda)^2 (h^\lambda)^2} \left(h^\lambda\right)^2 \mathbf{E}_\lambda}_{= \beta_\lambda \gamma_\lambda \frac{\mathbf{E}_\lambda}{(h^\lambda)^2}} = \sum_{\lambda \text{ is a partition of } n} \beta_\lambda \gamma_\lambda \frac{\mathbf{E}_\lambda}{(h^\lambda)^2}.
\end{aligned}$$

Comparing this with (256), we find $\Phi(\beta\gamma) = \Phi(\beta)\,\Phi(\gamma)$. Thus, Claim 4 is proved. $\square$

*Proof of Claim 5.* We have $1_{\mathbf{k}^{\{\text{partitions of } n\}}} = (1)_{\lambda \text{ is a partition of } n}$. Thus, the defini-

tion of $\Phi$ yields

$$\Phi\left(1_{\mathbf{k}^{\{\text{partitions of } n\}}}\right) = \sum_{\lambda \text{ is a partition of } n} 1\frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2} = \sum_{\lambda \text{ is a partition of } n} \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2} = 1$$

(by Corollary 5.12.24 **(b)**). In other words, $\Phi$ respects the unity. This proves Claim 5. □

Now, it is time to combine what we have shown. The map $\Phi$ is $\mathbf{k}$-linear and respects multiplication (by Claim 4) and respects the unity (by Claim 5). Hence, it is a $\mathbf{k}$-algebra morphism. Moreover, the map $\Phi$ is injective (by Claim 2) and surjective (by Claim 3), hence bijective, thus invertible. Hence, it is a $\mathbf{k}$-algebra isomorphism (since it is a $\mathbf{k}$-algebra morphism). This completes the proof of Theorem 5.12.26. □

Thus, Theorem 3.15.1 **(b)** is proved as well.

We have now reached a full understanding of the $\mathbf{k}$-algebra $Z\left(\mathbf{k}\left[S_n\right]\right)$ when $n!$ is invertible in $\mathbf{k}$. This is more or less the most general case in which $Z\left(\mathbf{k}\left[S_n\right]\right)$ is a direct product of $\mathbf{k}$'s. The structure of $Z\left(\mathbf{k}\left[S_n\right]\right)$ when $\mathbf{k}$ has "small characteristic" (i.e., $n!$ is not invertible in $\mathbf{k}$) is much less clear and might not be easily described.[98]

**Question 5.12.27.** What can we say about the structure of $Z\left(\mathbf{k}\left[S_n\right]\right)$ when $n$ is prime and $\mathbf{k}$ is a field of characteristic $n$ ? (This case appears to be the most tractable, as it comes the closest to the well-behaved "$n!$ invertible in $\mathbf{k}$" case.)

## 5.13. Irreducibility, non-isomorphism and completeness

Our next goal is the classification of irreps (i.e., irreducible representations) of the symmetric group $S_n$ over a field of characteristic 0. The goal here is to prove the following:

**Theorem 5.13.1.** Assume that $\mathbf{k}$ is a field of characteristic 0. Then:

**(a)** Each Specht module $\mathcal{S}^\lambda$ (where $\lambda$ is a partition of $n$) is irreducible as a left $\mathbf{k}\left[S_n\right]$-module (i.e., as a representation of $S_n$).

**(b)** The Specht modules $\mathcal{S}^\lambda$ for different partitions $\lambda$ of $n$ are mutually non-isomorphic.

**(c)** Each irreducible representation of $S_n$ is isomorphic to a Specht module $\mathcal{S}^\lambda$ for some partition $\lambda$ of $n$.

[98] For example, if $\mathbf{k}$ is a field with $0 < \operatorname{char}\mathbf{k} \leq n$, then the center $Z\left(\mathbf{k}\left[S_n\right]\right)$ contains the nonzero nilpotent element $\nabla$ (it is nilpotent since $\nabla^2 = n! \cdot \nabla = 0$, because $n! = 0$ in $\mathbf{k}$), whereas a direct product of $\mathbf{k}$'s cannot contain any nonzero nilpotent elements (since $\mathbf{k}$ is a field). Thus, $Z\left(\mathbf{k}\left[S_n\right]\right)$ cannot be isomorphic to a direct product of $\mathbf{k}$'s in this case.

This is a famous result (see, e.g., [CurRei62, (28.15)], [EGHetc11, Theorem 5.12.2], [Prasad15, Theorem 3.3.2], [Wildon18, Corollary 3.5, Corollary 4.4, and the paragraph below Example 4.5], [Waerde91, §14.7] for proofs at various levels of generality[99]), and is commonly proved using some representation theory of finite groups[100]. However, we shall give a proof that requires no new representation theory apart from a few lemmas that we will prove. It relies instead on the properties of Young symmetrizers and centralized Young symmetrizers we have derived above.

### 5.13.1. Irreducibility

We begin by preparing for the proof of Theorem 5.13.1 **(a)**. To prove this, we need a lemma:

**Lemma 5.13.2.** Let $\mathbf{k}$ be a field. Let $R$ be a $\mathbf{k}$-algebra. Let $e \in R$ be an element. Assume that $e$ has the following two properties:

1. Quasi-idempotency: There is a nonzero scalar $\kappa \in \mathbf{k}$ such that $e^2 = \kappa e$.
2. "Sandwich property": For each $a \in R$, the product $eae$ is a scalar multiple of $e$ (that is, we have $eae = \tau e$ for some $\tau \in \mathbf{k}$).

Then:

**(a)** Any left $R$-module morphism from $Re$ to $Re$ is a scalar multiple of the identity map id.

**(b)** In particular, any left $R$-module morphism from $Re$ to $Re$ is either zero or an isomorphism.

**(c)** The left $R$-module $Re$ cannot be written as a direct sum of two nonzero $R$-submodules.

**(d)** Assume that $e \neq 0$. Furthermore, assume that $R = \mathbf{k}[G]$ for a finite group $G$, and that $\operatorname{char} \mathbf{k} = 0$. Then, the left $R$-module $Re$ is irreducible.

*Proof.* By the quasi-idempotency property, there is a nonzero scalar $\kappa \in \mathbf{k}$ such

[99] In [EGHetc11, Theorem 5.12.2], the result is only stated for $\mathbf{k} = \mathbb{C}$; in [Waerde91, §14.7], only for algebraically closed fields $\mathbf{k}$; in [CurRei62, (28.15)], only for $\mathbf{k} = \mathbb{Q}$.

[100] In particular, there is a known fact (see, e.g., [CurRei62, (27.25)]) that if $G$ is an arbitrary finite group and $\mathbf{k}$ is a field of characteristic 0, then

$$\begin{aligned}&(\#\text{ of irreducible representations of } G \text{ (counted up to isomorphism)})\\ &\leq (\#\text{ of conjugacy classes of } G)\,. \end{aligned} \tag{257}$$

Using this fact, we can easily derive part **(c)** of Theorem 5.13.1 from parts **(a)** and **(b)** (since the # of conjugacy classes of $S_n$ is the # of partitions of $n$). It is worth noting (and this is much better known) that the inequality (257) becomes an equality whenever $\mathbf{k}$ is algebraically closed (see, e.g., [EGHetc11, Corollary 4.2.2] or [Serre77, §2.5, Theorem 7]).

that $e^2 = \kappa e$. Consider this $\kappa$. Note that $\kappa$ is nonzero and thus invertible (since $\mathbf{k}$ is a field).

**(a)** Let $f : Re \to Re$ be a left $R$-module morphism. We must show that $f$ is a scalar multiple of id.

We have $e \in Re$ (since $e = 1e$); thus, $f(e)$ is well-defined. Moreover, $f(e) \in Re$, so that $f(e) = ae$ for some $a \in R$. Consider this $a$. The "sandwich property" shows that $eae$ is a scalar multiple of $e$. In other words, we have $eae = \tau e$ for some $\tau \in \mathbf{k}$. Consider this $\tau$.

We have $e^2 = \kappa e$, so that $e = \frac{1}{\kappa} e^2 = \frac{1}{\kappa} ee$. Hence,

$$\begin{aligned} f(e) &= f\left(\frac{1}{\kappa} ee\right) = \frac{1}{\kappa} e \underbrace{f(e)}_{=ae} \qquad \text{(since } f \text{ is left } R\text{-linear)} \\ &= \frac{1}{\kappa} \underbrace{eae}_{=\tau e} = \frac{1}{\kappa} \tau e = \frac{\tau}{\kappa} e. \end{aligned}$$

Setting $\lambda := \frac{\tau}{\kappa}$, we can rewrite this as $f(e) = \lambda e$. Thus, for every $r \in R$, we have

$$\begin{aligned} f(re) &= r \underbrace{f(e)}_{=\lambda e} \qquad \text{(since } f \text{ is left } R\text{-linear)} \\ &= r \cdot \lambda e = \lambda re. \end{aligned}$$

In other words, $f(b) = \lambda b$ for each $b \in Re$ (since each $b \in Re$ has the form $re$ for some $r \in R$). In other words, $f = \lambda \operatorname{id}$. Thus, $f$ is a scalar multiple of id. This proves Lemma 5.13.2 **(a)**.

**(b)** Let $f : Re \to Re$ be a left $R$-module morphism. We must show that $f$ is either zero or an isomorphism.

Lemma 5.13.2 **(a)** shows that $f$ is a scalar multiple of id. In other words, $f = \lambda \operatorname{id}$ for some $\lambda \in \mathbf{k}$. If $\lambda = 0$, then this entails that $f$ is zero. If $\lambda \neq 0$, then this shows that $f$ is an isomorphism (with inverse $\lambda^{-1} \operatorname{id}$, which is well-defined because $\mathbf{k}$ is a field). Thus, we conclude that $f$ is either zero or an isomorphism. This proves Lemma 5.13.2 **(b)**.

**(c)** Assume the contrary. Thus, $Re = W \oplus U$ for some nonzero left $R$-submodules $W$ and $U$. Consider these $W$ and $U$. Thus, each element of $Re$ can be written uniquely as a sum $w + u$ for some $w \in W$ and some $u \in U$. This allows us to define a map

$$\begin{aligned} &f : Re \to Re, \\ &w + u \mapsto w \qquad \text{for each } w \in W \text{ and } u \in U. \end{aligned}$$

This map $f$ is easily seen to be a left $R$-module morphism (since $W$ and $U$ are $R$-submodules). Hence, by Lemma 5.13.2 **(b)**, it is either zero or an isomorphism.

But it cannot be zero (since its image is $f(Re) = W \neq 0$), and it cannot be an isomorphism either (since its kernel is $\operatorname{Ker} f = U \neq 0$). This is a contradiction. Thus, our assumption was false, and Lemma 5.13.2 **(c)** is proved.

**(d)** We recall that $G$-representations are the same as left $\mathbf{k}[G]$-modules, i.e., as left $R$-modules (since $\mathbf{k}[G] = R$).

From $e \neq 0$, we obtain $Re \neq 0$ (since $e = 1e \in Re$).

Next, we shall show that the $G$-representation $Re$ has no subrepresentation $W$ that is distinct from $Re$ and $\{0\}$.

Indeed, assume the contrary. Thus, $Re$ has such a subrepresentation $W$. Consider this $W$.

Now, $W$ is a $G$-subrepresentation, i.e., a left $\mathbf{k}[G]$-submodule of $Re$. Hence, it is a $\mathbf{k}$-vector subspace of $Re$, and thus a direct addend of $Re$ as a $\mathbf{k}$-vector space (by Proposition 4.4.9, since $\mathbf{k}$ is a field). Moreover, $|G|$ is invertible in $\mathbf{k}$ (since $\operatorname{char}\mathbf{k} = 0$). Thus, Maschke's theorem (Theorem 4.4.14, applied to $V = Re$) shows that the left $\mathbf{k}[G]$-submodule $W$ is also a direct addend of $Re$ as a left $\mathbf{k}[G]$-module. In other words, $Re = W \oplus U$ for some further left $\mathbf{k}[G]$-submodule $U$. Now, this latter submodule $U$ must be nonzero (since otherwise, $Re = W \oplus U$ would become $Re = W$, contradicting the fact that $W$ is distinct from $Re$). Hence, $Re = W \oplus U$ shows that $R$ is a direct sum of two nonzero $R$-submodules. But this contradicts Lemma 5.13.2 **(c)**.

This contradiction shows that our assumption was wrong. Thus we have proved that the $G$-representation $Re$ has no subrepresentation $W$ that is distinct from $Re$ and $\{0\}$. Since $Re \neq 0$, this shows that $Re$ is irreducible. This proves Lemma 5.13.2 **(d)**. □

**Corollary 5.13.3.** Assume that $\mathbf{k}$ is a field of characteristic 0. Let $\lambda$ be a partition of $n$. Then, the Specht module $\mathcal{S}^\lambda$ is irreducible as a left $\mathbf{k}[S_n]$-module (i.e., as a representation of $S_n$).

*Proof.* Recall that $\mathcal{A} = \mathbf{k}[S_n]$. Pick any $n$-tableau $T$ of shape $\lambda$ (clearly, such an $n$-tableau exists, since $|Y(\lambda)| = |\lambda| = n$). Then, (228) (applied to $D = Y(\lambda)$) yields $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$ (as left $\mathcal{A}$-modules). In other words, $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T$ (since $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)}$). We observe the following:

- The element $\mathbf{E}_T$ is quasi-idempotent: i.e., there is a nonzero scalar $\kappa \in \mathbf{k}$ such that $\mathbf{E}_T^2 = \kappa \mathbf{E}_T$. (Indeed, $\kappa = \dfrac{n!}{f^\lambda}$ works, because of Theorem 5.11.3.)
- The element $\mathbf{E}_T$ has the "sandwich property": For each $\mathbf{a} \in \mathcal{A}$, the product $\mathbf{E}_T \mathbf{a} \mathbf{E}_T$ is a scalar multiple of $\mathbf{E}_T$. (This is just the claim of Proposition 5.11.5.)
- We have $\mathbf{E}_T \neq 0$ (since Lemma 5.11.13 yields $[1]\,\mathbf{E}_T = 1 \neq 0$).
- We have $\mathcal{A} = \mathbf{k}[G]$ for some finite group $G$ (namely, for $G = S_n$).

- We have $\operatorname{char} \mathbf{k} = 0$.

Thus, Lemma 5.13.2 **(d)** (applied to $R = \mathcal{A}$ and $e = \mathbf{E}_T$) shows that the left $\mathcal{A}$-module $\mathcal{A}\mathbf{E}_T$ is irreducible. In other words, the left $\mathbf{k}[S_n]$-module $\mathcal{A}\mathbf{E}_T$ is irreducible (since $\mathcal{A} = \mathbf{k}[S_n]$). Hence, the left $\mathbf{k}[S_n]$-module $\mathcal{S}^\lambda$ is irreducible (since $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T$). This proves Corollary 5.13.3. □

This proves Theorem 5.13.1 **(a)**.

It is easy to generalize Corollary 5.13.3 to the case when $\mathbf{k}$ is a field of characteristic larger than $n$. However, we cannot allow $\mathbf{k}$ to be just a commutative $\mathbb{Q}$-algebra, since irreducibility becomes a far less well-behaved notion when $\mathbf{k}$ is not a field (for example, if $\mathbf{k}$ is the polynomial ring $\mathbb{Q}[x]$, then any Specht module $\mathcal{S}^\lambda$ has subrepresentations $x\mathcal{S}^\lambda$, $x^2\mathcal{S}^\lambda$, $x^3\mathcal{S}^\lambda$, ..., which are all distinct and nonzero).

A much more interesting question is what happens when $\mathbf{k}$ is a field of positive characteristic $\leq n$. The Specht modules $\mathcal{S}^\lambda$ are no longer always irreducible in this case, and indeed there is a blatant counterexample in characteristic 2 ([Wildon18, Example 4.5]):

**Proposition 5.13.4.** If $2 = 0$ in $\mathbf{k}$, then

$$\mathcal{S}^{(5,1,1)} \cong \mathcal{S}^{(5,2)} \oplus \mathcal{S}^{(7)} \qquad \text{as left } \mathbf{k}[S_7]\text{-modules.}$$

Thus, $\mathcal{S}^{(5,1,1)}$ is not only not irreducible over a characteristic-2 field, but even decomposes as a direct sum of two other Specht modules! Non-irreducible Specht modules $\mathcal{S}^\lambda$ also exist in other positive characteristics, although non-trivial direct sum decompositions only exist in characteristic 2 (see [James78, Corollary 13.18] for a proof).

Some Specht modules do remain irreducible in positive characteristic. For example, here is the situation for $n = 4$:

| | $\operatorname{char} \mathbf{k} = 2$ | $\operatorname{char} \mathbf{k} = 3$ | $\operatorname{char} \mathbf{k} \notin \{2, 3\}$ |
|---|---|---|---|
| $\mathcal{S}^{(4)}$ | irreducible | irreducible | irreducible |
| $\mathcal{S}^{(3,1)}$ | reducible | irreducible | irreducible |
| $\mathcal{S}^{(2,2)}$ | irreducible | reducible | irreducible |
| $\mathcal{S}^{(2,1,1)}$ | reducible | irreducible | irreducible |
| $\mathcal{S}^{(1,1,1,1)}$ | irreducible | irreducible | irreducible |

.

**Exercise 5.13.1.** 3+3+1 Verify all the claims in the above table. (The claims in the first and the last rows are trivial, since 1-dimensional $\mathbf{k}$-modules are always irreducible. Each remaining claim in any of the first two columns is worth 1 point; the whole third column is worth 1 point.)

See [JamMat98] for a classification of the Specht modules $\mathcal{S}^\lambda$ that are irreducible over a characteristic-2 field $\mathbf{k}$.

Arguably, for fields of positive characteristic, the proper candidates for the irreducible $S_n$-representations are not the Specht modules $\mathcal{S}^\lambda$ themselves, but rather their quotients. See [Wildon18, §5] and [James78, §11] for the details of their construction and [JamKer81, Chapter 6] for more about the modular (i.e., positive-characteristic) representation theory of $S_n$. In a nutshell, the irreducible $S_n$-representations can still be nicely characterized, but their fine structure (bases, dimensions, etc.) are mostly unknown (e.g., a "standard basis theorem" is missing).

### 5.13.2. Non-isomorphism

Next we aim at proving Theorem 5.13.1 **(b)**. Here, we shall use the following lemma to distinguish $\mathcal{S}^\lambda$ from $\mathcal{S}^\mu$:

**Lemma 5.13.5.** Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $\lambda$. As usual in algebra, we shall write $\mathbf{a}M$ for the set $\{\mathbf{a}m \mid m \in M\}$ when $\mathbf{a}$ is an element of $\mathbf{k}[S_n]$ and when $M$ is a left $\mathbf{k}[S_n]$-module. Then:

**(a)** If $h^\lambda \neq 0$ in $\mathbf{k}$, then $\mathbf{E}_T\mathcal{S}^\lambda \neq 0$.

**(b)** If $\mu$ is any partition of $n$ that is distinct from $\lambda$, then $\mathbf{E}_T\mathcal{S}^\mu = 0$.

*Proof.* **(a)** Assume that $h^\lambda \neq 0$ in $\mathbf{k}$. Then,

$$\begin{aligned}\mathbf{E}_T 1 \mathbf{E}_T = \mathbf{E}_T^2 &= \frac{n!}{f^\lambda}\mathbf{E}_T \qquad \text{(by Theorem 5.11.3)} \\ &= h^\lambda \mathbf{E}_T \qquad \left(\text{since } \frac{n!}{f^\lambda} = h^\lambda\right),\end{aligned}$$

so that $[1]\left(\mathbf{E}_T 1 \mathbf{E}_T\right) = [1]\left(h^\lambda \mathbf{E}_T\right) = h^\lambda \cdot \underbrace{[1]\,\mathbf{E}_T}_{\substack{=1 \\ \text{(by Lemma 5.11.13)}}} = h^\lambda \neq 0$. Therefore, $\mathbf{E}_T 1 \mathbf{E}_T \neq 0$. Hence, $\mathbf{E}_T \mathcal{A} \mathbf{E}_T \neq 0$ (since $\mathbf{E}_T 1 \mathbf{E}_T$ is clearly an element of $\mathbf{E}_T \mathcal{A} \mathbf{E}_T$).

But (228) (applied to $D = Y(\lambda)$) yields $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$ (as left $\mathcal{A}$-modules). In other words, $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T$ (since $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)}$). Hence, $\mathbf{E}_T\mathcal{S}^\lambda \cong \mathbf{E}_T\mathcal{A}\mathbf{E}_T$ (since the action of $\mathbf{E}_T$ on isomorphic left $\mathbf{k}[S_n]$-modules is isomorphic). Therefore, $\mathbf{E}_T\mathcal{S}^\lambda \neq 0$ (since $\mathbf{E}_T\mathcal{A}\mathbf{E}_T \neq 0$). This proves Lemma 5.13.5 **(a)**.

**(b)** Let $\mu$ be any partition of $n$ that is distinct from $\lambda$. Pick any $n$-tableau $S$ of shape $\mu$ (such an $n$-tableau clearly exists). Then, $S$ is an $n$-tableau of shape $Y(\mu)$, whereas $T$ is an $n$-tableau of shape $Y(\lambda)$. Therefore, Proposition 5.11.15 (applied to $T$ and $S$ instead of $S$ and $T$) yields that

$$\mathbf{E}_T \mathbf{a} \mathbf{E}_S = 0 \qquad \text{for each } \mathbf{a} \in \mathbf{k}[S_n].$$

In other words, $\mathbf{E}_T \mathcal{A} \mathbf{E}_S = 0$.

However, in our above proof of Lemma 5.13.5 **(a)**, we have shown that $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules. Similarly, we can see that $\mathcal{S}^\mu \cong \mathcal{A}\mathbf{E}_S$ as left $\mathcal{A}$-modules (since $S$ is an $n$-tableau $S$ of shape $\mu$). Therefore, $\mathbf{E}_T \mathcal{S}^\mu \cong \mathbf{E}_T \mathcal{A} \mathbf{E}_S = 0$, so that $\mathbf{E}_T \mathcal{S}^\mu = 0$. This proves Lemma 5.13.5 **(b).** □

**Corollary 5.13.6.** Assume that $\mathbf{k}$ is a field of characteristic 0. Let $\lambda$ and $\mu$ be two distinct partitions of $n$. Then, $\mathcal{S}^\lambda \ncong \mathcal{S}^\mu$ as $S_n$-representations.

*First proof of Corollary 5.13.6.* Assume the contrary. Thus, we have $\mathcal{S}^\lambda \cong \mathcal{S}^\mu$ as $S_n$-representations.

Note that $h^\lambda \neq 0$ in $\mathbf{k}$ (since $\mathbf{k}$ is a field of characteristic 0).

Pick any $n$-tableau $T$ of shape $\lambda$ (such a $T$ clearly exists). Then, Lemma 5.13.5 **(a)** yields $\mathbf{E}_T \mathcal{S}^\lambda \neq 0$. But Lemma 5.13.5 **(b)** yields $\mathbf{E}_T \mathcal{S}^\mu = 0$. However, from $\mathcal{S}^\lambda \cong \mathcal{S}^\mu$, we obtain $\mathbf{E}_T \mathcal{S}^\lambda \cong \mathbf{E}_T \mathcal{S}^\mu = 0$, contradicting $\mathbf{E}_T \mathcal{S}^\lambda \neq 0$. This contradiction shows that our assumption was wrong. Thus, Corollary 5.13.6 is proved. □

*Second proof of Corollary 5.13.6 (sketched).* Let me sketch a different proof of Corollary 5.13.6 (coming from [James78, Corollary 13.17]), which establishes a more general result: Namely, it shows that Corollary 5.13.6 still holds whenever $\mathbf{k}$ is a field of characteristic $\neq 2$ (rather than characteristic 0). Even better, it works whenever $\mathbf{k}$ is any nontrivial commutative ring in which 2 is invertible.

So let $\mathbf{k}$ be a nontrivial commutative ring in which 2 is invertible. Let $\lambda$ and $\mu$ be two distinct partitions of $n$.

Consider the **smallest** $i \geq 1$ satisfying $\lambda_i \neq \mu_i$. Assume WLOG that $\lambda_i > \mu_i$ (otherwise, we can swap $\lambda$ with $\mu$). Then,

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i > \mu_1 + \mu_2 + \cdots + \mu_i \tag{258}$$

(as we saw in the proof of Proposition 5.11.15).

We must show that $\mathcal{S}^\lambda \ncong \mathcal{S}^\mu$ as $S_n$-representations.

Indeed, assume the contrary. Thus, there is a left $\mathbf{k}[S_n]$-module isomorphism $f : \mathcal{S}^\lambda \to \mathcal{S}^\mu$. Consider this $f$.

Pick any $n$-tableau $T$ of shape $\lambda$. Write the image $f(\mathbf{e}_T) \in \mathcal{S}^\mu = \mathcal{S}^{Y(\mu)} \subseteq \mathcal{M}^{Y(\mu)}$ as

$$f(\mathbf{e}_T) = \sum_{\overline{S}} a_{\overline{S}} \overline{S}, \tag{259}$$

where the sum ranges over all $n$-tabloids $\overline{S}$ of shape $Y(\mu)$, and where the coef-

ficients $a_{\overline{S}}$ are scalars in $\mathbf{k}$. For any $w \in S_n$, we have

$$\begin{aligned} f(\mathbf{e}_{wT}) &= f(w\mathbf{e}_T) && \text{(since Lemma 5.4.6 \textbf{(a)} yields } \mathbf{e}_{wT} = w\mathbf{e}_T\text{)} \\ &= wf(\mathbf{e}_T) && \text{(since } f \text{ is a left } \mathbf{k}[S_n]\text{-module morphism)} \\ &= w\sum_{\overline{S}} a_{\overline{S}}\overline{S} && \text{(by (259))} \\ &= \sum_{\overline{S}} a_{\overline{S}} w\overline{S} = \sum_{\overline{S}} a_{w^{-1}\overline{S}}\overline{S} \end{aligned} \tag{260}$$

(here, we have substituted $w^{-1}\overline{S}$ for $\overline{S}$ in the sum).

Now, let $w \in S_n$ be a transposition $t_{u,v}$ swapping two distinct numbers $u, v \in [n]$ that lie in the same column of $T$. Then, $w \in \mathcal{C}(T)$, so that Lemma 5.4.6 **(c)** yields $\mathbf{e}_{wT} = (-1)^w \mathbf{e}_T = -\mathbf{e}_T$ (since the transposition $w$ has sign $(-1)^w = -1$), and thus we have

$$f(\mathbf{e}_{wT}) = f(-\mathbf{e}_T) = -f(\mathbf{e}_T) = -\sum_{\overline{S}} a_{\overline{S}}\overline{S} \qquad \text{(by (259))}.$$

Comparing this with (260), we obtain

$$\sum_{\overline{S}} a_{w^{-1}\overline{S}}\overline{S} = -\sum_{\overline{S}} a_{\overline{S}}\overline{S} = \sum_{\overline{S}} \left(-a_{\overline{S}}\right)\overline{S}.$$

Since the $n$-tabloids $\overline{S}$ form a basis of $\mathcal{M}^{Y(\mu)}$, we thus obtain

$$a_{w^{-1}\overline{S}} = -a_{\overline{S}} \qquad \text{for each } n\text{-tabloid } \overline{S} \text{ of shape } Y(\mu). \tag{261}$$

Forget that we fixed $w$. We thus have shown that if $w \in S_n$ is a transposition $t_{u,v}$ swapping two distinct numbers $u, v \in [n]$ that lie in the same column of $T$, then (261) holds.

Now, let $\overline{S}$ be any $n$-tabloid of shape $Y(\mu)$. Then, by Theorem 5.10.1 **(b)** (or, rather, by its contrapositive), we can see that there are two distinct integers $u$ and $v$ that lie in the same row of $S$ and simultaneously lie in the same column of $T$ (since otherwise, (258) would be violated). Consider these $u$ and $v$. Let $w$ be the transposition $t_{u,v} \in S_n$. Then, (261) shows that $a_{w^{-1}\overline{S}} = -a_{\overline{S}}$. But $w = t_{u,v} \in \mathcal{R}(S)$ (since $u$ and $v$ lie in the same row of $S$), so that $wS$ is row-equivalent to $S$. Therefore, $\overline{wS} = \overline{S}$. Thus, $w\overline{S} = \overline{wS} = \overline{S}$, so that $w^{-1}\overline{S} = \overline{S}$ as well. Hence, the equality $a_{w^{-1}\overline{S}} = -a_{\overline{S}}$ can be rewritten as $a_{\overline{S}} = -a_{\overline{S}}$. In other words, $2a_{\overline{S}} = 0$. Thus, $a_{\overline{S}} = 0$ (since 2 is invertible in $\mathbf{k}$).

Forget that we fixed $\overline{S}$. We thus have shown that $a_{\overline{S}} = 0$ for each $n$-tabloid $\overline{S}$ of shape $Y(\mu)$. Hence, (259) rewrites as $f(\mathbf{e}_T) = \sum_{\overline{S}} 0\overline{S} = 0$.

Forget that we fixed $T$. We have now shown that $f(\mathbf{e}_T) = 0$ for each $n$-tableau $T$ of shape $\lambda$. In other words, the $\mathbf{k}$-linear map $f$ sends each polytabloid $\mathbf{e}_T$ to 0. Since the Specht module $\mathcal{S}^\lambda$ is spanned by the polytabloids $\mathbf{e}_T$, this entails that $f$ sends every vector to 0. In other words, $f = 0$. Hence, $f\left(\mathcal{S}^\lambda\right) = 0$.

But $f$ is an isomorphism, and thus $f\left(\mathcal{S}^{\lambda}\right) = \mathcal{S}^{\mu}$. Thus, $\mathcal{S}^{\mu} = f\left(\mathcal{S}^{\lambda}\right) = 0$. But Corollary 5.7.10 shows that $\mathcal{S}^{\mu} \neq 0$. This clearly contradicts $\mathcal{S}^{\mu} = 0$. This contradiction shows that our assumption was false, and thus Corollary 5.13.6 is proved again. □

Note that Corollary 5.13.6 fails in characteristic 2, since we have

$$\mathcal{S}^{(2)} \cong \mathcal{S}^{(1,1)} \qquad \text{when } \mathbf{k} \text{ is a field of characteristic 2}$$

(for fairly obvious reasons).

Of course, Theorem 5.13.1 **(b)** follows from Corollary 5.13.6 (in fact, it is just a restatement of the latter).

**Exercise 5.13.2.** [2] Tweak the above second proof of Corollary 5.13.6 a bit further to work under the even weaker assumption "$\mathbf{k}$ is a commutative ring such that $2 \neq 0$ in $\mathbf{k}$" instead of "$\mathbf{k}$ is a field of characteristic 0".

### 5.13.3. Completeness

To prove Theorem 5.13.1 **(c)**, we need the following general fact from representation theory:

**Proposition 5.13.7.** Let $\mathbf{k}$ be a field. Let $R$ be a $\mathbf{k}$-algebra that is finite-dimensional as a $\mathbf{k}$-vector space. Let $J_1, J_2, \ldots, J_m$ be $m$ irreducible left $R$-submodules of the left regular $R$-module $R$ satisfying $R = J_1 + J_2 + \cdots + J_m$. Let $K$ be a further irreducible left $R$-module. Then, $K \cong J_i$ for some $i \in [m]$.

*Proof sketch.* If some $i \in [m]$ satisfies $J_1 + J_2 + \cdots + J_i = J_1 + J_2 + \cdots + J_{i-1}$, then we have

$$\begin{aligned} R &= J_1 + J_2 + \cdots + J_m \\ &= \underbrace{\left(J_1 + J_2 + \cdots + J_i\right)}_{= J_1 + J_2 + \cdots + J_{i-1}} + \left(J_{i+1} + J_{i+2} + \cdots + J_m\right) \\ &= \left(J_1 + J_2 + \cdots + J_{i-1}\right) + \left(J_{i+1} + J_{i+2} + \cdots + J_m\right), \end{aligned}$$

and therefore the equality $R = J_1 + J_2 + \cdots + J_m$ remains valid if we remove the addend $J_i$ from its right hand side. Thus, in this case, we can reduce the claim of Proposition 5.13.7 to an analogous claim with a smaller value of $m$ (by removing $J_i$ from the list $(J_1, J_2, \ldots, J_m)$ and replacing $m$ by $m-1$).

We can perform such reductions until there is no longer any $i \in [m]$ remaining that satisfies $J_1 + J_2 + \cdots + J_i = J_1 + J_2 + \cdots + J_{i-1}$. Hence, we can WLOG assume that no $i \in [m]$ satisfies $J_1 + J_2 + \cdots + J_i = J_1 + J_2 + \cdots + J_{i-1}$. Assume this. Thus,

$$J_1 + J_2 + \cdots + J_i \neq J_1 + J_2 + \cdots + J_{i-1} \tag{262}$$

for each $i \in [m]$.

Now, for each $i \in \{0, 1, \ldots, m\}$, let us define the left $R$-submodule

$$F_i := J_1 + J_2 + \cdots + J_i \text{ of } R.$$

Thus, $F_0 = \{0\}$ (since an empty sum of submodules is $\{0\}$ by definition) and $F_m = J_1 + J_2 + \cdots + J_m = R$. Moreover, each $i \in [m]$ satisfies

$$\begin{aligned} F_i &= J_1 + J_2 + \cdots + J_i = \underbrace{\left(J_1 + J_2 + \cdots + J_{i-1}\right)}_{\substack{= F_{i-1} \\ \text{(by the definition of } F_{i-1}\text{)}}} + J_i \\ &= F_{i-1} + J_i \end{aligned} \tag{263}$$

and therefore $F_{i-1} \subseteq F_i$. Hence, $\{0\} = F_0 \subseteq F_1 \subseteq F_2 \subseteq \cdots \subseteq F_m = R$. This shows that $(F_0, F_1, \ldots, F_m)$ is a filtration of the left $R$-module $R$.

Next, let us show that the subquotients $F_i/F_{i-1}$ of this filtration are isomorphic to the respective $J_i$:

*Claim 1:* Let $i \in [m]$. Then, $F_i/F_{i-1} \cong J_i$ as left $R$-modules.

*Proof of Claim 1.* First of all, we have $F_{i-1} = J_1 + J_2 + \cdots + J_{i-1}$ (by the definition of $F_{i-1}$). But (262) says that $J_1 + J_2 + \cdots + J_i \neq J_1 + J_2 + \cdots + J_{i-1}$. In view of $F_i = J_1 + J_2 + \cdots + J_i$ and $F_{i-1} = J_1 + J_2 + \cdots + J_{i-1}$, we can rewrite this as $F_i \neq F_{i-1}$. Thus, $F_i/F_{i-1} \neq \{0\}$.

From (263), we have $F_i = F_{i-1} + J_i$. Hence, $J_i \subseteq F_i$. Thus, there is a left $R$-linear map

$$\begin{aligned} g : J_i &\to F_i/F_{i-1}, \\ x &\mapsto \overline{x} \end{aligned}$$

(where $\overline{x}$ denotes the residue class of $x \in F_i$ modulo $F_{i-1}$). Consider this map $g$.

Let us show that $g$ is surjective. Indeed, let $v \in F_i/F_{i-1}$ be arbitrary. Then, we can write $v$ as $v = \overline{y}$ for some $y \in F_i$ (obviously). Consider this $y$. Then, $y \in F_i = F_{i-1} + J_i$, so that we can write $y$ as $y = z + w$ for some $z \in F_{i-1}$ and $w \in J_i$. Consider these $z$ and $w$. From $y = z + w$, we obtain $y - w = z \in F_{i-1}$ and therefore $\overline{y} = \overline{w}$. Hence, $v = \overline{y} = \overline{w} = g(w)$ (since the definition of $g$ yields $g(w) = \overline{w}$). Therefore, $v = g(w) \in \operatorname{Im} g$.

Forget that we fixed $v$. We thus have shown that $v \in \operatorname{Im} g$ for each $v \in F_i/F_{i-1}$. In other words, the map $g$ is surjective. That is, $g(J_i) = F_i/F_{i-1} \neq \{0\}$.

The kernel $\operatorname{Ker} g$ of the map $g : J_i \to F_i/F_{i-1}$ is a left $R$-submodule of $J_i$ (since $g$ is left $R$-linear), and thus must be either $J_i$ or $\{0\}$ (since the $R$-module $J_i$ is irreducible, and thus its only $R$-submodules are $J_i$ and $\{0\}$). If $\operatorname{Ker} g$ was $J_i$, then $g(J_i)$ would be $\{0\}$, which would contradict $g(J_i) \neq \{0\}$. Hence, $\operatorname{Ker} g$ must be $\{0\}$ (since $\operatorname{Ker} g$ is either $J_i$ or $\{0\}$). In other words, $g$ is injective (since $g$ is a left $R$-linear map). Since $g$ is also surjective, we thus conclude that $g$ is bijective, i.e., invertible. Hence, $g$ is a left $R$-module isomorphism (since $g$ is left $R$-linear). Thus, $F_i/F_{i-1} \cong J_i$ as left $R$-modules. This proves Claim 1. $\square$

Claim 1 shows that each subquotient $F_i/F_{i-1}$ of our filtration $(F_0, F_1, \ldots, F_m)$ is isomorphic to the corresponding $J_i$, and thus is irreducible (since $J_i$ is irreducible). Hence, the filtration $(F_0, F_1, \ldots, F_m)$ is a composition series of $R$. Thus, Theorem 4.4.24 (applied to $m$ and $(F_0, F_1, \ldots, F_m)$ instead of $k$ and $(W_0, W_1, \ldots, W_k)$) shows that each irreducible left $R$-module is isomorphic to $F_i/F_{i-1}$ for some $i \in [m]$. In particular, $K$ is isomorphic to $F_i/F_{i-1}$ for some $i \in [m]$ (since $K$ is an irreducible left $R$-module). For this $i$, we then have $K \cong F_i/F_{i-1} \cong J_i$ (by Claim 1). This proves Proposition 5.13.7. □

**Corollary 5.13.8.** Assume that $\mathbf{k}$ is a field of characteristic 0. Let $K$ be any irreducible representation of $S_n$ over $\mathbf{k}$. Then, $K \cong \mathcal{S}^\lambda$ for some partition $\lambda$ of $n$.

*Proof.* Recall that we denote the ring $\mathbf{k}[S_n]$ by $\mathcal{A}$. Thus, $K$ is an irreducible left $\mathcal{A}$-module (since $K$ is an irreducible representation of $S_n$ over $\mathbf{k}$).

Let $(T_1, T_2, \ldots, T_m)$ be a list of **all** the $n$-tableaux of shape $\lambda$ for **all** the partitions $\lambda$ of $n$. (Thus, the length $m$ of this list is $n! \cdot p(n)$, where $p(n)$ is the # of all partitions of $n$.)

Theorem 5.12.17 yields

$$\begin{aligned} 1 &= \sum_{\lambda \text{ is a partition of } n} \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2} \\ &= \sum_{\lambda \text{ is a partition of } n} \frac{\sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T}{\left(h^\lambda\right)^2} \qquad \text{(by (241))} \\ &= \sum_{\lambda \text{ is a partition of } n} \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \frac{\mathbf{E}_T}{\left(h^\lambda\right)^2}. \end{aligned} \tag{264}$$

Thus, each $\mathbf{a} \in \mathcal{A}$ satisfies

$$\begin{aligned} \mathbf{a} = \mathbf{a} \cdot 1 &= \mathbf{a} \cdot \sum_{\lambda \text{ is a partition of } n} \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \frac{\mathbf{E}_T}{\left(h^\lambda\right)^2} \qquad \text{(by (264))} \\ &= \sum_{\lambda \text{ is a partition of } n} \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \underbrace{\frac{\mathbf{a}\mathbf{E}_T}{\left(h^\lambda\right)^2}}_{\in \mathcal{A}\mathbf{E}_T} \\ &\in \sum_{\lambda \text{ is a partition of } n} \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathcal{A}\mathbf{E}_T \\ &= \mathcal{A}\mathbf{E}_{T_1} + \mathcal{A}\mathbf{E}_{T_2} + \cdots + \mathcal{A}\mathbf{E}_{T_m} \end{aligned}$$

(since $T_1, T_2, \ldots, T_m$ are all the $n$-tableaux of shape $\lambda$ for all the partitions $\lambda$ of $n$). Hence,

$$\mathcal{A} \subseteq \mathcal{A}\mathbf{E}_{T_1} + \mathcal{A}\mathbf{E}_{T_2} + \cdots + \mathcal{A}\mathbf{E}_{T_m}.$$

This entails

$$\mathcal{A} = \mathcal{A}\mathbf{E}_{T_1} + \mathcal{A}\mathbf{E}_{T_2} + \cdots + \mathcal{A}\mathbf{E}_{T_m}$$

(since we obviously have $\mathcal{A}\mathbf{E}_{T_1} + \mathcal{A}\mathbf{E}_{T_2} + \cdots + \mathcal{A}\mathbf{E}_{T_m} \subseteq \mathcal{A}$ as well). Moreover, the left $\mathcal{A}$-submodules $\mathcal{A}\mathbf{E}_{T_i}$ of $\mathcal{A}$ are isomorphic to Specht modules of the form $\mathcal{S}^\lambda$ (by (228), since $T_i$ is an $n$-tableau of shape $\lambda$ for some partition $\lambda$ of $n$), and thus irreducible (since Corollary 5.13.3 says that any Specht module of the form $\mathcal{S}^\lambda$ is irreducible). Hence, Proposition 5.13.7 (applied to $R = \mathcal{A}$ and $J_i = \mathcal{A}\mathbf{E}_{T_i}$) yields that $K \cong \mathcal{A}\mathbf{E}_{T_i}$ for some $i \in [m]$. Therefore, $K \cong \mathcal{S}^\lambda$ for some partition $\lambda$ of $n$ (since, as we just said, each $\mathcal{A}\mathbf{E}_{T_i}$ is isomorphic to a Specht module of the form $\mathcal{S}^\lambda$). This proves Corollary 5.13.8. □

Thus, Theorem 5.13.1 **(c)** follows (being a restatement of Corollary 5.13.8).

Furthermore, we obtain the following corollary, which says that each finite-dimensional representation of $S_n$ over a field of characteristic 0 can be decomposed into a direct sum of Specht modules $\mathcal{S}^\lambda$:

**Corollary 5.13.9.** Assume that $\mathbf{k}$ is a field of characteristic 0. Let $P$ be the set of all partitions of $n$. Let $V$ be any finite-dimensional representation of $S_n$ over $\mathbf{k}$. Then,

$$V \cong \bigoplus_{\lambda \in P} \left(\mathcal{S}^\lambda\right)^{k_\lambda} \qquad \text{for some integers } k_\lambda \in \mathbb{N}.$$

Here, $M^k$ means the external direct sum $\underbrace{M \oplus M \oplus \cdots \oplus M}_{k \text{ times}}$ whenever $M$ is an $S_n$-representation and $k \in \mathbb{N}$.

*Proof.* Theorem 4.4.23 **(a)** (applied to $R = \mathbf{k}[S_n]$) yields that there exists a composition series $(W_0, W_1, \ldots, W_k)$ of $V$. Consider this composition series. For each $i \in [k]$, the quotient $W_i/W_{i-1}$ is an irreducible left $\mathbf{k}[S_n]$-module (by the definition of a composition series), i.e., an irreducible representation of $S_n$, and thus satisfies $W_i/W_{i-1} \cong \mathcal{S}^\lambda$ for some partition $\lambda$ of $n$ (by Corollary 5.13.8, applied to $K = W_i/W_{i-1}$). Let us denote this partition $\lambda$ by $\lambda_i$. Thus,

$$W_i/W_{i-1} \cong \mathcal{S}^{\lambda_i} \qquad \text{for each } i \in [k]. \tag{265}$$

Clearly, $\lambda_i \in P$ for each $i \in [k]$ (since $\lambda_i$ is a partition of $n$).

But $(W_0, W_1, \ldots, W_k)$ is a composition series of $V$, thus a filtration of $V$. Hence, Corollary 4.4.22 (applied to $G = S_n$) yields

$$\begin{aligned} V &\cong (W_1/W_0) \oplus (W_2/W_1) \oplus \cdots \oplus (W_k/W_{k-1}) \\ &= \bigoplus_{i \in [k]} \underbrace{(W_i/W_{i-1})}_{\substack{\cong \mathcal{S}^{\lambda_i} \\ \text{(by (265))}}} \cong \bigoplus_{i \in [k]} \mathcal{S}^{\lambda_i} \cong \bigoplus_{\lambda \in P} \left(\mathcal{S}^\lambda\right)^{(\#\text{ of } i \in [k] \text{ satisfying } \lambda_i = \lambda)} \end{aligned}$$

(here, we have rewritten our direct sum by collecting equal addends together – for instance, $\mathcal{S}^{(4)} \oplus \mathcal{S}^{(3,1)} \oplus \mathcal{S}^{(4)} \oplus \mathcal{S}^{(4)} = \left(\mathcal{S}^{(4)}\right)^3 \oplus \left(\mathcal{S}^{(3,1)}\right)^1$). Hence,

$$V \cong \bigoplus_{\lambda \in P} \left(\mathcal{S}^{\lambda}\right)^{k_\lambda} \qquad \text{for some integers } k_\lambda \in \mathbb{N}$$

(namely, for $k_\lambda = (\#\text{ of } i \in [k] \text{ satisfying } \lambda_i = \lambda)$). This proves Corollary 5.13.9. □

## 5.14. The Artin–Wedderburn theorem

### 5.14.1. The theorem

Another crucial fact that we can now prove is the *Artin–Wedderburn theorem for symmetric groups*, stated here in its coordinate-free version (i.e., using endomorphism rings instead of matrix rings):

**Theorem 5.14.1.** Assume that $n!$ is invertible in $\mathbf{k}$. For each partition $\lambda$ of $n$, we let $\rho_\lambda : \mathbf{k}[S_n] \to \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ be the curried form of the left $\mathbf{k}[S_n]$-action on the Specht module $\mathcal{S}^\lambda$. (See Definition 4.3.28 for the meaning of "curried form". Explicitly, $\rho_\lambda$ is given by $(\rho_\lambda(\mathbf{a}))(\mathbf{v}) = \mathbf{a}\mathbf{v}$ for each $\mathbf{a} \in \mathbf{k}[S_n]$ and each $\mathbf{v} \in \mathcal{S}^\lambda$.)

Then, the map

$$\mathbf{k}[S_n] \to \underbrace{\prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)}_{\text{a direct product of } \mathbf{k}\text{-algebras}},$$
$$\mathbf{a} \mapsto \left(\rho_\lambda(\mathbf{a})\right)_{\lambda \text{ is a partition of } n}$$

is a $\mathbf{k}$-algebra isomorphism.

### 5.14.2. Linear algebra lemmas

To prove this theorem, we recall a few basic results from linear algebra over commutative rings. The first one allows us to (sometimes) simplify our life when proving that two linear maps are mutually inverse:

**Lemma 5.14.2.** Let $r \in \mathbb{N}$. Let $V$ and $W$ be two free $\mathbf{k}$-modules of rank $r$. Let $f : V \to W$ and $g : W \to V$ be two $\mathbf{k}$-linear maps such that $f \circ g = \mathrm{id}_W$. Then, $g \circ f = \mathrm{id}_V$.

This fact is well-known in the case when $\mathbf{k}$ is a field (see, e.g., [Axler24, 3.68] for this case), but actually holds for an arbitrary commutative ring $\mathbf{k}$. The general case can be reduced to matrix algebra in a standard way:

*Proof of Lemma 5.14.2.* Choose a basis $(v_1, v_2, \ldots, v_r)$ of $V$ and a basis $(w_1, w_2, \ldots, w_r)$ of $W$. Let $A \in \mathbf{k}^{r \times r}$ and $B \in \mathbf{k}^{r \times r}$ be the matrices that represent the $\mathbf{k}$-linear maps $f$ and $g$ with respect to these two bases. Then, the matrices $AB$ and $BA$ represent the compositions $f \circ g$ and $g \circ f$ with respect to these two bases (since matrix multiplication corresponds to composition of linear maps). Hence, from $f \circ g = \mathrm{id}_W$, we obtain $AB = I_r$ (where $I_r$ denotes the $r \times r$ identity matrix). But this entails $BA = I_r$ by a known result about square matrices (see, e.g., [Grinbe15, Corollary 6.112 **(a)**]). This, in turn, entails $g \circ f = \mathrm{id}_V$ (since the matrix $BA$ represents the linear map $g \circ f$). This proves Lemma 5.14.2. □

**Remark 5.14.3.** Lemma 5.14.2 can be generalized: Instead of requiring $V$ and $W$ to be free $\mathbf{k}$-modules of the same rank $r \in \mathbb{N}$, we can get by just requiring $V$ and $W$ to be two isomorphic finitely generated $\mathbf{k}$-modules. For a proof of this version, see [Conrad, "Universal identities", Corollary 5.6] (why does this generalize Lemma 5.14.2?).

Next, we recall two simple properties of ranks of free modules. The first one is the straightforward generalization (to arbitrary rings) of the classical fact that the dimension of a direct product is the sum of the dimensions of the factors:

**Lemma 5.14.4.** Let $M_1, M_2, \ldots, M_k$ be $k$ free $\mathbf{k}$-modules of ranks $r_1, r_2, \ldots, r_k$, respectively. Then, their direct product $M_1 \times M_2 \times \cdots \times M_k$ is a free $\mathbf{k}$-module of rank $r_1 + r_2 + \cdots + r_k$.

*Proof idea.* For each $i \in [k]$, choose a basis $\left(m_{i,1}, m_{i,2}, \ldots, m_{i,r_i}\right)$ of the free $\mathbf{k}$-module $M_i$. Let $m'_{i,j}$ be the image of each basis vector $m_{i,j} \in M_i$ under the canonical embedding $M_i \to M_1 \times M_2 \times \cdots \times M_k$. Then, the list

$$\begin{aligned}&(m'_{1,1}, m'_{1,2}, \ldots, m'_{1,r_1},\\ &\ m'_{2,1}, m'_{2,2}, \ldots, m'_{2,r_2},\\ &\ \ldots,\\ &\ m'_{k,1}, m'_{k,2}, \ldots, m'_{k,r_k})\end{aligned}$$

is a basis of $M_1 \times M_2 \times \cdots \times M_k$. Since this basis consists of $r_1 + r_2 + \cdots + r_k$ vectors, it follows that $M_1 \times M_2 \times \cdots \times M_k$ is a free $\mathbf{k}$-module of rank $r_1 + r_2 + \cdots + r_k$. □

The next property we need is a formula for the rank (i.e., dimension) of the endomorphism ring of a free $\mathbf{k}$-module:

**Lemma 5.14.5.** Let $M$ be a free $\mathbf{k}$-module of rank $r$. Then, $\mathrm{End}_{\mathbf{k}} M$ is a free $\mathbf{k}$-module of rank $r^2$.

*Proof idea.* Choose a basis $(m_1, m_2, \ldots, m_r)$ of $M$. Then, each endomorphism $f \in \mathrm{End}_{\mathbf{k}} M$ is represented by a unique $r \times r$-matrix $A_f \in \mathbf{k}^{r \times r}$ with respect

to this basis. This gives a one-to-one correspondence between endomorphisms $f \in \operatorname{End}_{\mathbf{k}} M$ and $r \times r$-matrices $A \in \mathbf{k}^{r \times r}$. That is, we obtain a bijection

$$\begin{aligned} \operatorname{End}_{\mathbf{k}} M &\to \mathbf{k}^{r \times r}, \\ f &\mapsto A_f. \end{aligned}$$

Moreover, this bijection is $\mathbf{k}$-linear (since addition of matrices corresponds to addition of linear maps, and the same holds for scaling), and thus is a $\mathbf{k}$-module isomorphism. Hence, the $\mathbf{k}$-module $\operatorname{End}_{\mathbf{k}} M$ is isomorphic to $\mathbf{k}^{r \times r}$.

But the $\mathbf{k}$-module $\mathbf{k}^{r \times r}$ has a basis consisting of the $r^2$ elementary matrices (each of which has a 1 in some cell and 0's in all remaining cells). Hence, the $\mathbf{k}$-module $\mathbf{k}^{r \times r}$ is free of rank $r^2$. Thus, the $\mathbf{k}$-module $\operatorname{End}_{\mathbf{k}} M$ is free of rank $r^2$ as well (since the $\mathbf{k}$-module $\operatorname{End}_{\mathbf{k}} M$ is isomorphic to $\mathbf{k}^{r \times r}$). □

We also recall that the endomorphism ring of a free $\mathbf{k}$-module is isomorphic to a matrix ring:

**Lemma 5.14.6.** Let $M$ be a free $\mathbf{k}$-module of rank $r$. Then, $\operatorname{End}_{\mathbf{k}} M \cong \mathbf{k}^{r \times r}$ as $\mathbf{k}$-algebras.

*Proof idea.* Argue as in the proof of Lemma 5.14.5, but observe additionally that the bijection $\operatorname{End}_{\mathbf{k}} M \to \mathbf{k}^{r \times r}$ is a $\mathbf{k}$-algebra morphism (since multiplication of matrices corresponds to composition of linear maps) and thus is a $\mathbf{k}$-algebra isomorphism. □

### 5.14.3. The proof

We can now prove Theorem 5.14.1:

*Proof of Theorem 5.14.1. Step 1:* In the following, for any given partition $\lambda$ of $n$, we shall abbreviate the summation sign "$\sum_{T \text{ is an } n\text{-tableau of shape } \lambda}$" as "$\sum_{T \text{ of shape } \lambda}$". Using this abbreviation, we can rewrite the equality (241) as

$$\mathbf{E}_\lambda = \sum_{T \text{ of shape } \lambda} \mathbf{E}_T \tag{266}$$

(for any partition $\lambda$ of $n$).

*Step 2:* Define a map

$$\begin{aligned} \Phi : \mathbf{k}\left[S_n\right] &\to \prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right), \\ \mathbf{a} &\mapsto \left(\rho_\lambda\left(\mathbf{a}\right)\right)_{\lambda \text{ is a partition of } n}. \end{aligned}$$

This map $\Phi$ is precisely the map

$$\mathbf{k}\left[S_n\right] \to \prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right),$$
$$\mathbf{a} \mapsto \left(\rho_\lambda\left(\mathbf{a}\right)\right)_{\lambda \text{ is a partition of } n}$$

in Theorem 5.14.1. Thus, our goal is to prove that this map $\Phi$ is a $\mathbf{k}$-algebra isomorphism.

We recall that if $R$ is any $\mathbf{k}$-algebra, and if $V$ is any left $R$-module, then the curried form $\rho : R \to \operatorname{End}_{\mathbf{k}} V$ of the left $R$-action on $V$ is a $\mathbf{k}$-algebra morphism (by Theorem 4.3.27). Hence, for each partition $\lambda$ of $n$, the curried form $\rho_\lambda : \mathbf{k}\left[S_n\right] \to \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ of the left $\mathbf{k}\left[S_n\right]$-action on the Specht module $\mathcal{S}^\lambda$ is a $\mathbf{k}$-algebra morphism. Thus, the map $\Phi$ (which is defined by applying all these morphisms $\rho_\lambda$ to its input and collecting the resulting outputs in a family) is a $\mathbf{k}$-algebra morphism as well[101].

We shall now focus on proving that $\Phi$ is invertible.

*Step 3:* Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $\lambda$. Then, $T$ is an $n$-tableau of shape $Y\left(\lambda\right)$. Hence, (228) (applied to $D = Y\left(\lambda\right)$) yields $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$ (as left $\mathbf{k}\left[S_n\right]$-modules, i.e., as left $\mathcal{A}$-modules). Hence, $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$. In other words, there exists an $\mathcal{A}$-module isomorphism $\alpha_{\lambda,T} : \mathcal{S}^\lambda \to \mathcal{A}\mathbf{E}_T$. Fix such an isomorphism $\alpha_{\lambda,T}$ once and for all.[102]

Forget that we fixed $\lambda$ and $T$. Thus, for any partition $\lambda$ of $n$ and any $n$-tableau $T$ of shape $\lambda$, we have defined an $\mathcal{A}$-module isomorphism $\alpha_{\lambda,T} : \mathcal{S}^\lambda \to \mathcal{A}\mathbf{E}_T$. In particular, $\alpha_{\lambda,T}$ is thus a $\mathbf{k}$-module isomorphism. Therefore, if $\zeta \in \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ is a $\mathbf{k}$-module endomorphism of $\mathcal{S}^\lambda$, then $\alpha_{\lambda,T} \circ \zeta \circ \alpha_{\lambda,T}^{-1} \in \operatorname{End}_{\mathbf{k}}\left(\mathcal{A}\mathbf{E}_T\right)$ is a $\mathbf{k}$-module endomorphism of $\mathcal{A}\mathbf{E}_T$.

*Step 4:* Now, define a map

$$\Psi : \prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right) \to \mathbf{k}\left[S_n\right],$$

which sends each family $\left(\zeta_\lambda\right)_{\lambda \text{ is a partition of } n} \in \prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ to

$$\sum_{\lambda \text{ is a partition of } n} \frac{1}{\left(h^\lambda\right)^2} \sum_{T \text{ of shape } \lambda} \underbrace{\left(\alpha_{\lambda,T} \circ \zeta_\lambda \circ \alpha_{\lambda,T}^{-1}\right)\left(\mathbf{E}_T\right)}_{\substack{\text{(this is well-defined,} \\ \text{since } \mathbf{E}_T = 1\mathbf{E}_T \in \mathcal{A}\mathbf{E}_T \\ \text{and } \alpha_{\lambda,T} \circ \zeta_\lambda \circ \alpha_{\lambda,T}^{-1} \in \operatorname{End}_{\mathbf{k}}\left(\mathcal{A}\mathbf{E}_T\right))}} \in \mathcal{A} = \mathbf{k}\left[S_n\right].$$

[101] since the $\mathbf{k}$-algebra structure on $\prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ is entrywise

[102] A canonical choice of $\alpha_{\lambda,T}$ can be obtained from Theorem 5.5.13 **(b)**, but we don't need it to be canonical here.

This map $\Psi$ is clearly $\mathbf{k}$-linear (since each $\left(\alpha_{\lambda,T} \circ \zeta_\lambda \circ \alpha_{\lambda,T}^{-1}\right)(\mathbf{E}_T)$ depends $\mathbf{k}$-linearly on $\zeta_\lambda$ and thus on the input of $\Psi$).

*Step 5:* We shall next show that $\Psi \circ \Phi = \mathrm{id}$.

Indeed, let $\mathbf{a} \in \mathbf{k}[S_n]$. Then, the definition of $\Phi$ yields

$$\Phi(\mathbf{a}) = (\rho_\lambda(\mathbf{a}))_{\lambda \text{ is a partition of } n}.$$

Hence,

$$\begin{aligned}\Psi(\Phi(\mathbf{a})) &= \Psi\left((\rho_\lambda(\mathbf{a}))_{\lambda \text{ is a partition of } n}\right) \\ &= \sum_{\lambda \text{ is a partition of } n} \frac{1}{(h^\lambda)^2} \sum_{T \text{ of shape } \lambda} \left(\alpha_{\lambda,T} \circ (\rho_\lambda(\mathbf{a})) \circ \alpha_{\lambda,T}^{-1}\right)(\mathbf{E}_T)\end{aligned}$$

(by the definition of $\Psi$).

However, we can simplify the addends in this sum significantly, due to the following claim:

*Claim 1:* Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $\lambda$. Then,

$$\left(\alpha_{\lambda,T} \circ (\rho_\lambda(\mathbf{a})) \circ \alpha_{\lambda,T}^{-1}\right)(\mathbf{E}_T) = \mathbf{a}\mathbf{E}_T.$$

*Proof of Claim 1.* Set $\mathbf{u} = \alpha_{\lambda,T}^{-1}(\mathbf{E}_T)$. Then, $\alpha_{\lambda,T}(\mathbf{u}) = \mathbf{E}_T$. Furthermore, we have $(\rho_\lambda(\mathbf{a}))(\mathbf{u}) = \mathbf{a}\mathbf{u}$ (since $\rho_\lambda$ is just the curried form of the left $\mathbf{k}[S_n]$-action on $\mathcal{S}^\lambda$). Moreover, $\alpha_{\lambda,T}(\mathbf{a}\mathbf{u}) = \mathbf{a}\alpha_{\lambda,T}(\mathbf{u})$ (since $\alpha_{\lambda,T}$ is an $\mathcal{A}$-module morphism). Now,

$$\begin{aligned}\left(\alpha_{\lambda,T} \circ (\rho_\lambda(\mathbf{a})) \circ \alpha_{\lambda,T}^{-1}\right)(\mathbf{E}_T) &= \alpha_{\lambda,T}\left((\rho_\lambda(\mathbf{a}))\left(\underbrace{\alpha_{\lambda,T}^{-1}(\mathbf{E}_T)}_{=\mathbf{u}}\right)\right) \\ &= \alpha_{\lambda,T}\left(\underbrace{(\rho_\lambda(\mathbf{a}))(\mathbf{u})}_{=\mathbf{a}\mathbf{u}}\right) = \alpha_{\lambda,T}(\mathbf{a}\mathbf{u}) = \mathbf{a}\underbrace{\alpha_{\lambda,T}(\mathbf{u})}_{=\mathbf{E}_T} \\ &= \mathbf{a}\mathbf{E}_T.\end{aligned}$$

This proves Claim 1. □

Now,

$$\begin{aligned}
(\Psi \circ \Phi)(\mathbf{a}) &= \Psi(\Phi(\mathbf{a})) \\
&= \sum_{\lambda \text{ is a partition of } n} \frac{1}{\left(h^\lambda\right)^2} \sum_{T \text{ of shape } \lambda} \underbrace{\left(\alpha_{\lambda,T} \circ (\rho_\lambda(\mathbf{a})) \circ \alpha_{\lambda,T}^{-1}\right)(\mathbf{E}_T)}_{=\mathbf{a}\mathbf{E}_T \text{ (by Claim 1)}} \\
&\qquad \text{(by our above computation of } \Psi(\Phi(\mathbf{a}))) \\
&= \sum_{\lambda \text{ is a partition of } n} \frac{1}{\left(h^\lambda\right)^2} \sum_{T \text{ of shape } \lambda} \mathbf{a}\mathbf{E}_T \\
&= \mathbf{a} \sum_{\lambda \text{ is a partition of } n} \frac{1}{\left(h^\lambda\right)^2} \underbrace{\sum_{T \text{ of shape } \lambda} \mathbf{E}_T}_{=\mathbf{E}_\lambda \text{ (by (266))}} \\
&= \mathbf{a} \sum_{\lambda \text{ is a partition of } n} \frac{1}{\left(h^\lambda\right)^2} \mathbf{E}_\lambda = \mathbf{a} \underbrace{\sum_{\lambda \text{ is a partition of } n} \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}}_{=1 \text{ (by Corollary 5.12.24 \textbf{(b)})}} = \mathbf{a}.
\end{aligned}$$

Forget that we fixed $\mathbf{a}$. We thus have shown that $(\Psi \circ \Phi)(\mathbf{a}) = \mathbf{a}$ for each $\mathbf{a} \in \mathbf{k}[S_n]$. In other words, $\Psi \circ \Phi = \mathrm{id}$.

*Step 6:* Next we want to show that $\Phi \circ \Psi = \mathrm{id}$ as well. We will derive this from $\Psi \circ \Phi = \mathrm{id}$ using Lemma 5.14.2. In order to do so, we need to check that $\mathbf{k}[S_n]$ and $\prod_{\lambda \text{ is a partition of } n} \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ are two free $\mathbf{k}$-modules of the same rank (namely, $n!$).

Let us now do this.

Let $\lambda$ be a partition of $n$. Then, Lemma 5.9.17 shows that $\mathcal{S}^\lambda$ is a free $\mathbf{k}$-module of rank $f^\lambda$. Hence, Lemma 5.14.5 (applied to $M = \mathcal{S}^\lambda$ and $r = f^\lambda$) shows that $\mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ is a free $\mathbf{k}$-module of rank $\left(f^\lambda\right)^2$.

Forget that we fixed $\lambda$. We thus have shown that $\mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ is a free $\mathbf{k}$-module of rank $\left(f^\lambda\right)^2$ whenever $\lambda$ is a partition of $n$. Hence, by Lemma 5.14.4, we conclude that the direct product $\prod_{\lambda \text{ is a partition of } n} \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ of these free $\mathbf{k}$-modules $\mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ is a free $\mathbf{k}$-module of rank $\sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2$. In other words, the direct product $\prod_{\lambda \text{ is a partition of } n} \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ is a free $\mathbf{k}$-module of rank $n!$ (since Corollary 5.12.18 yields $\sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 = n!$).

But the $\mathbf{k}$-module $\mathbf{k}[S_n]$ is also a free $\mathbf{k}$-module of rank $n!$ (since it has a basis $(w)_{w \in S_n}$, which consists of $|S_n| = n!$ many vectors).

Thus, Lemma 5.14.2 (applied to $V = \prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right)$ and $W = \mathbf{k}\left[S_n\right]$ and $r = n!$ and $f = \Psi$ and $g = \Phi$) shows that $\Phi \circ \Psi = \mathrm{id}$ (since $\Psi \circ \Phi = \mathrm{id}$).

*Step 7:* Combining $\Phi \circ \Psi = \mathrm{id}$ with $\Psi \circ \Phi = \mathrm{id}$, we see that the maps $\Phi$ and $\Psi$ are mutually inverse. Hence, the map $\Phi$ is invertible, and thus is a $\mathbf{k}$-algebra isomorphism (since it is a $\mathbf{k}$-algebra morphism). As explained above, this proves Theorem 5.14.1. □

**Exercise 5.14.1.** Let $\mathbf{a} \in \mathbf{k}\left[S_n\right]$. Assume that $\mathbf{a}\mathcal{S}^{\lambda} = 0$ for each partition $\lambda$ of $n$. (As usual, $\mathbf{a}\mathcal{S}^{\lambda}$ means $\left\{\mathbf{av} \mid \mathbf{v} \in \mathcal{S}^{\lambda}\right\}$.)

**(a)** [1] Prove that $\mathbf{a} = 0$ if $n!$ is invertible in $\mathbf{k}$.

**(b)** [3] Prove that $n!^2 \cdot \mathbf{a} = 0$ always.

### 5.14.4. Proof of Theorem 3.15.1 (a)

We can now easily verify Theorem 3.15.1 **(a)** as well (which is just a less specific version of Theorem 5.14.1):

*Proof of Theorem 3.15.1 (a).* Let $\lambda$ be a partition of $n$. Then, $\mathcal{S}^{\lambda}$ is a free $\mathbf{k}$-module of rank $f^{\lambda}$ (by Lemma 5.9.17). Hence,

$$\operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right) \cong \mathbf{k}^{f^{\lambda} \times f^{\lambda}} \qquad \text{as } \mathbf{k}\text{-algebras} \tag{267}$$

(by Lemma 5.14.6, applied to $M = \mathcal{S}^{\lambda}$ and $r = f^{\lambda}$).

Forget that we fixed $\lambda$. We thus have proved (267) for each partition $\lambda$ of $n$. Now, Theorem 5.14.1 constructs a $\mathbf{k}$-algebra isomorphism

$$\mathbf{k}\left[S_n\right] \to \prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right).$$

Thus, we see that

$$\mathbf{k}\left[S_n\right] \cong \prod_{\lambda \text{ is a partition of } n} \underbrace{\operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right)}_{\substack{\cong \mathbf{k}^{f^{\lambda} \times f^{\lambda}} \\ \text{(by (267))}}} \cong \prod_{\lambda \text{ is a partition of } n} \mathbf{k}^{f^{\lambda} \times f^{\lambda}}$$

as $\mathbf{k}$-algebras. In other words,

$$\mathbf{k}\left[S_n\right] \cong \prod_{\lambda \in \Lambda} \mathbf{k}^{f_{\lambda} \times f_{\lambda}},$$

where $\Lambda$ is the set of all partitions of $n$, and where each $f_{\lambda}$ is what we call $f^{\lambda}$. This proves Theorem 3.15.1 **(a)**. □

We have now proved all of Theorem 3.15.1 except for part **(c)**. The latter part will require significantly more preparation: We will have to introduce the *seminormal basis*, which connects Young tableaux with the jucys–murphies.

### 5.14.5. The kernel and the surjectivity of $\rho_\lambda : \mathbf{k}[S_n] \to \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$

Theorem 5.14.1 claims something like a "joint bijectivity" of the maps $\rho_\lambda : \mathcal{A} \to \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ – meaning that even if each of these maps is not bijective per se, their combination (sending $\mathbf{a} \mapsto (\rho_\lambda(\mathbf{a}))_{\lambda \text{ is a partition of } n}$) is (as long as $n!$ is invertible in $\mathbf{k}$). One might wonder what can be said about a single map $\rho_\lambda$ rather than all of them in combination. And indeed, there are things that can be said about it, and even under a slightly weaker assumption (requiring the invertibility of $h^\lambda$ rather than the stronger invertibility of $n!$). We begin with an easy description of its kernel:

**Proposition 5.14.7.** Let $\lambda$ be a partition of $n$. Assume that $h^\lambda$ is invertible in $\mathbf{k}$. Let $\rho_\lambda : \mathbf{k}[S_n] \to \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ be the curried form of the left $\mathbf{k}[S_n]$-action on the Specht module $\mathcal{S}^\lambda$ (as in Theorem 5.14.1). Let $\mathbf{G}_\lambda := 1 - \dfrac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2} \in \mathcal{A}$. Then,

$$\mathrm{Ker}(\rho_\lambda) = \mathcal{A}\mathbf{G}_\lambda.$$

This will follow easily from the following fact, which is interesting by itself:

**Proposition 5.14.8.** Let $\lambda$ be a partition of $n$. Let $\mathbf{v} \in \mathcal{S}^\lambda$. Then, $\mathbf{E}_\lambda \mathbf{v} = \left(h^\lambda\right)^2 \mathbf{v}$.

*Proof.* The diagram $Y(\lambda)$ has size $|Y(\lambda)| = |\lambda| = n$ (since $\lambda$ is a partition of $n$). Pick any $n$-tableau $T$ of shape $Y(\lambda)$. (Such a $T$ clearly exists, since $|Y(\lambda)| = n$.)

Thus, (228) (applied to $D = Y(\lambda)$) yields $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules. In other words, $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules (since $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)}$). In other words, there exists a left $\mathcal{A}$-module isomorphism $f : \mathcal{S}^\lambda \to \mathcal{A}\mathbf{E}_T$. Consider this $f$. Thus, $f(\mathbf{v}) \in \mathcal{A}\mathbf{E}_T$. In other words, $f(\mathbf{v}) = \mathbf{a}\mathbf{E}_T$ for some $\mathbf{a} \in \mathcal{A}$. Consider this $\mathbf{a}$.

Proposition 5.12.3 yields that $\mathbf{E}_\lambda$ belongs to the center $Z(\mathbf{k}[S_n])$ of $\mathbf{k}[S_n]$. Hence, $\mathbf{E}_\lambda \mathbf{a} = \mathbf{a}\mathbf{E}_\lambda$ (since $\mathbf{a} \in \mathcal{A} = \mathbf{k}[S_n]$).

Now, recall that $f$ is a left $\mathcal{A}$-module morphism. Hence,

$$f(\mathbf{E}_\lambda \mathbf{v}) = \mathbf{E}_\lambda \underbrace{f(\mathbf{v})}_{=\mathbf{a}\mathbf{E}_T} = \underbrace{\mathbf{E}_\lambda \mathbf{a}}_{=\mathbf{a}\mathbf{E}_\lambda} \mathbf{E}_T = \mathbf{a} \underbrace{\mathbf{E}_\lambda \mathbf{E}_T}_{=\left(h^\lambda\right)^2 \mathbf{E}_T \text{ (by Proposition 5.12.15)}}$$

$$= \left(h^\lambda\right)^2 \underbrace{\mathbf{a}\mathbf{E}_T}_{=f(\mathbf{v})} = \left(h^\lambda\right)^2 f(\mathbf{v}) = f\left(\left(h^\lambda\right)^2 \mathbf{v}\right)$$

(since $f$ is $\mathbf{k}$-linear). But the map $f$ is injective (since $f$ is an isomorphism). Thus, from $f(\mathbf{E}_\lambda \mathbf{v}) = f\left(\left(h^\lambda\right)^2 \mathbf{v}\right)$, we obtain $\mathbf{E}_\lambda \mathbf{v} = \left(h^\lambda\right)^2 \mathbf{v}$. This proves Proposition 5.14.8. □

*Proof of Proposition 5.14.7. Step 1:* We shall first show that $\mathcal{A}\mathbf{G}_\lambda \subseteq \operatorname{Ker}(\rho_\lambda)$.

*Proof.* Let $\mathbf{a} \in \mathcal{A}\mathbf{G}_\lambda$. Then, $\mathbf{a} = \mathbf{b}\mathbf{G}_\lambda$ for some $\mathbf{b} \in \mathcal{A}$. Consider this $\mathbf{b}$.

Let $\mathbf{v} \in \mathcal{S}^\lambda$. Then, from $\mathbf{G}_\lambda = 1 - \dfrac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}$, we obtain

$$\mathbf{G}_\lambda \mathbf{v} = \left(1 - \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}\right)\mathbf{v} = \mathbf{v} - \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}\mathbf{v} = \mathbf{v} - \frac{1}{\left(h^\lambda\right)^2} \underbrace{\mathbf{E}_\lambda \mathbf{v}}_{\substack{=\left(h^\lambda\right)^2 \mathbf{v} \\ \text{(by Proposition 5.14.8)}}}$$

$$= \mathbf{v} - \underbrace{\frac{1}{\left(h^\lambda\right)^2}\left(h^\lambda\right)^2}_{=1}\mathbf{v} = \mathbf{v} - \mathbf{v} = 0.$$

Now,

$$\underbrace{\mathbf{a}}_{=\mathbf{b}\mathbf{G}_\lambda}\mathbf{v} = \mathbf{b}\underbrace{\mathbf{G}_\lambda \mathbf{v}}_{=0} = 0.$$

The definition of $\rho_\lambda$ yields $(\rho_\lambda(\mathbf{a}))(\mathbf{v}) = \mathbf{a}\mathbf{v} = 0$.

Forget that we fixed $\mathbf{v}$. We thus have shown that $(\rho_\lambda(\mathbf{a}))(\mathbf{v}) = 0$ for each $\mathbf{v} \in \mathcal{S}^\lambda$. In other words, $\rho_\lambda(\mathbf{a}) = 0$. In other words, $\mathbf{a} \in \operatorname{Ker}(\rho_\lambda)$.

Forget that we fixed $\mathbf{a}$. We thus have proved that $\mathbf{a} \in \operatorname{Ker}(\rho_\lambda)$ for each $\mathbf{a} \in \mathcal{A}\mathbf{G}_\lambda$. In other words, $\mathcal{A}\mathbf{G}_\lambda \subseteq \operatorname{Ker}(\rho_\lambda)$.

*Step 2:* We shall now prove that $\operatorname{Ker}(\rho_\lambda) \subseteq \mathcal{A}\mathbf{G}_\lambda$.

*Proof.* Let $\mathbf{d} \in \operatorname{Ker}(\rho_\lambda)$. Thus, $\mathbf{d} \in \mathbf{k}[S_n] = \mathcal{A}$ and $\rho_\lambda(\mathbf{d}) = 0$.

Now we shall show that $\mathbf{d}\mathbf{E}_T = 0$ for any $n$-tableau $T$ of shape $\lambda$.

Indeed, let $T$ be an $n$-tableau of shape $\lambda$. Thus, $T$ is an $n$-tableau of shape $Y(\lambda)$. Hence, (228) (applied to $D = Y(\lambda)$) yields $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules. In other words, $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules (since $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)}$). In other words, there exists a left $\mathcal{A}$-module isomorphism $f : \mathcal{S}^\lambda \to \mathcal{A}\mathbf{E}_T$. Consider this $f$. We have $\mathbf{E}_T = \underbrace{1}_{\in\mathcal{A}}\mathbf{E}_T \in \mathcal{A}\mathbf{E}_T$. Hence, there exists some $\mathbf{v} \in \mathcal{S}^\lambda$ such that $\mathbf{E}_T = f(\mathbf{v})$ (since $f$ is an isomorphism and thus surjective). Consider this $\mathbf{v}$. (Actually, $\mathbf{v} = \mathbf{e}_T$, but we don't need this.)

The definition of $\rho_\lambda$ yields $(\rho_\lambda(\mathbf{d}))(\mathbf{v}) = \mathbf{d}\mathbf{v}$, so that $\mathbf{d}\mathbf{v} = \underbrace{(\rho_\lambda(\mathbf{d}))}_{=0}(\mathbf{v}) = 0$.

Hence, $f(\mathbf{d}\mathbf{v}) = f(0) = 0$ (since $f$ is $\mathbf{k}$-linear). But $f$ is a left $\mathcal{A}$-module morphism; thus, $f(\mathbf{d}\mathbf{v}) = \mathbf{d}\underbrace{f(\mathbf{v})}_{=\mathbf{E}_T} = \mathbf{d}\mathbf{E}_T$. Hence, $\mathbf{d}\mathbf{E}_T = f(\mathbf{d}\mathbf{v}) = 0$.

Forget that we fixed $T$. We thus have shown that

$$\mathbf{d}\mathbf{E}_T = 0 \qquad \text{for any } n\text{-tableau } T \text{ of shape } \lambda. \tag{268}$$

Now,

$$\begin{aligned}\mathbf{d}\mathbf{E}_\lambda &= \mathbf{d} \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T \qquad \text{(by (241))} \\ &= \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \underbrace{\mathbf{d}\mathbf{E}_T}_{\substack{=0 \\ \text{(by (268))}}} = 0.\end{aligned}$$

Hence,

$$\begin{aligned}\mathbf{d}\mathbf{G}_\lambda &= \mathbf{d}\left(1 - \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}\right) \qquad \left(\text{since } \mathbf{G}_\lambda = 1 - \frac{\mathbf{E}_\lambda}{\left(h^\lambda\right)^2}\right) \\ &= \mathbf{d} - \frac{\mathbf{d}\mathbf{E}_\lambda}{\left(h^\lambda\right)^2} = \mathbf{d} - \frac{0}{\left(h^\lambda\right)^2} \qquad (\text{since } \mathbf{d}\mathbf{E}_\lambda = 0) \\ &= \mathbf{d},\end{aligned}$$

so that $\mathbf{d} = \underbrace{\mathbf{d}}_{\in \mathcal{A}} \mathbf{G}_\lambda \in \mathcal{A}\mathbf{G}_\lambda$.

Forget that we fixed $\mathbf{d}$. We thus have shown that $\mathbf{d} \in \mathcal{A}\mathbf{G}_\lambda$ for each $\mathbf{d} \in \operatorname{Ker}(\rho_\lambda)$. In other words, $\operatorname{Ker}(\rho_\lambda) \subseteq \mathcal{A}\mathbf{G}_\lambda$.

*Step 3:* Combining $\operatorname{Ker}(\rho_\lambda) \subseteq \mathcal{A}\mathbf{G}_\lambda$ with $\mathcal{A}\mathbf{G}_\lambda \subseteq \operatorname{Ker}(\rho_\lambda)$, we obtain $\operatorname{Ker}(\rho_\lambda) = \mathcal{A}\mathbf{G}_\lambda$. This proves Proposition 5.14.7. □

More subtle is the surjectivity of $\rho_\lambda$:

**Theorem 5.14.9.** Let $\lambda$ be a partition of $n$. Assume that $h^\lambda$ is invertible in $\mathbf{k}$. Let $\rho_\lambda : \mathbf{k}[S_n] \to \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ be the curried form of the left $\mathbf{k}[S_n]$-action on the Specht module $\mathcal{S}^\lambda$ (as in Theorem 5.14.1). Then:

**(a)** The map $\rho_\lambda$ is surjective.

**(b)** The restriction of $\rho_\lambda$ to $\mathcal{A}\mathbf{E}_\lambda$ is a $\mathbf{k}$-algebra isomorphism from $\mathcal{A}\mathbf{E}_\lambda$ to $\operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$. (See Corollary 5.12.16 **(b)** for why $\mathcal{A}\mathbf{E}_\lambda$ is a $\mathbf{k}$-algebra.)

We will prove Theorem 5.14.9 in full in the next section (Subsection 5.15.6), but it is already not hard to prove it under the slightly more restrictive requirement that $n!$ be invertible:

**Exercise 5.14.2.** [3] Prove Theorem 5.14.9 in the case when $n!$ is invertible in $\mathbf{k}$.

[**Hint:** Show first that if $e$ is an idempotent element of a $\mathbf{k}$-algebra $R$, then $Re \cong R/R(1-e)$ as $\mathbf{k}$-algebras.]

## 5.15. The Young natural basis of $\mathbf{k}[S_n]$

Corollary 5.12.18 shows that $n!$ equals the number of **pairs** of standard tableaux of shape $Y(\lambda)$, summed over all partitions $\lambda$ of $n$. (Indeed, for any given partition $\lambda$ of $n$, the number of such pairs is $\left(f^\lambda\right)^2$.) The Artin–Wedderburn theorem (Theorem 3.15.1 **(a)**) provides an algebraic interpretation of this fact. Another linear-algebraic consequence of this fact is that $\mathcal{A}$ (as a $\mathbf{k}$-module) has a basis indexed by such pairs.

Of course, the mere existence of such a basis is obvious (since the number of such pairs is $n!$, and clearly $\mathcal{A}$ has a basis of size $n!$). But it is reasonable to wonder whether there is a "natural" basis of this kind, i.e., a basis of combinatorially defined elements that are meaningfully related to pairs of standard tableaux.

The answer is "yes", at least when $n!$ is invertible in $\mathbf{k}$. In this section, we will construct such a basis, which we call the *Young natural basis* (Corollary 5.15.28). It consists of Young symmetrizers $\mathbf{E}_P$, tweaked slightly to depend on two $n$-tableaux $P$ and $Q$ rather than a single $n$-tableau $P$. These tweaked Young symmetrizers are called $\mathbf{E}_{P,Q}$, and we will define them shortly.

Once the elements $\mathbf{E}_{P,Q}$ are defined, the basis will be easy to find: It is the family $\left(\mathbf{E}_{U,V}\right)_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \mathrm{SYT}(\lambda)}$, where $\mathrm{SYT}(\lambda)$ denotes the set of all standard $n$-tableaux of shape $Y(\lambda)$. Proving that this family really is a basis of $\mathcal{A}$ (when $n!$ is invertible) is not all that easy. The journey is long, although none of the steps involved is particularly difficult.

The standard basis theorem for Specht modules will play a role in the proof, but we will also need to understand the behavior of the elements $\mathbf{E}_{P,Q}$ under transposing the tableaux $P$ and $Q$ (that is, reflecting them across the northwest-to-southeast diagonal). This behavior will be described in (5.15.16) below, which is just one of a number of auxiliary results that we will use as stepping stones towards the proof. Another remarkable result we will find along the way (Corollary 5.15.25) is an explicit decomposition of $\mathcal{A}$ into a direct sum of submodules of the form $\mathcal{A}\mathbf{E}_T$, each of which is isomorphic to a Specht module $\mathcal{S}^\lambda$. Some further comments about the Young natural basis will be made in Section 6.4.

One of our auxiliary results (Corollary 5.15.24) will then be used to prove Theorem 5.14.9 (as promised in the previous section).

We shall use Definition 5.12.11 throughout this section.

### 5.15.1. The $w_{P,Q}$ and the $\mathbf{E}_{P,Q}$

We begin our journey with the definitions of the elements $w_{P,Q}$ and $\mathbf{E}_{P,Q}$.

The following lemma is nearly obvious, and has been implicitly used a few times already (see, e.g., Proposition 5.3.13 **(b)**):

**Lemma 5.15.1.** Let $D$ be any diagram. Let $P$ and $Q$ be two $n$-tableaux of shape $D$. Then, there is a unique permutation $w \in S_n$ that satisfies $wQ = P$.

*Proof.* We know that $P$ is an $n$-tableau of shape $D$. In other words, $P$ is an injective map from $D$ to $\{1, 2, 3, \ldots\}$ whose image is $[n]$. Thus, we can view $P$ as a bijection from $D$ to $[n]$. Likewise, we can view $Q$ as a bijection from $D$ to $[n]$. Hence, the inverse $Q^{-1}$ of $Q$ is well-defined and is a bijection from $[n]$ to $D$. Therefore, $P \circ Q^{-1}$ is a well-defined bijection from $[n]$ to $[n]$. In other words, $P \circ Q^{-1}$ is a well-defined permutation in $S_n$.

But any permutation $w \in S_n$ satisfies $wQ = w \rightharpoonup Q = w \circ Q$ (by Definition 5.3.9). Thus, for any permutation $w \in S_n$, we have the following chain of equivalences:

$$\begin{aligned}
(wQ = P) &\iff (w \circ Q = P) && (\text{since } wQ = w \circ Q) \\
&\iff \left(w \circ Q = P \circ Q^{-1} \circ Q\right) && \left(\text{since } P = P \circ Q^{-1} \circ Q\right) \\
&\iff \left(w = P \circ Q^{-1}\right)
\end{aligned}$$

(here, we have cancelled the $Q$ on both sides, since $Q$ is a bijection). Hence, a permutation $w \in S_n$ satisfies $wQ = P$ if and only if it satisfies $w = P \circ Q^{-1}$. In other words, there is exactly one permutation $w \in S_n$ that satisfies $wQ = P$, namely the permutation $P \circ Q^{-1}$ (since we already know that $P \circ Q^{-1}$ is a well-defined permutation in $S_n$). This proves Lemma 5.15.1. □

Let us give the unique permutation $w$ in Lemma 5.15.1 a name:

**Definition 5.15.2.** Let $D$ be any diagram. Let $P$ and $Q$ be two $n$-tableaux of shape $D$. Lemma 5.15.1 shows that there is a unique permutation $w \in S_n$ that satisfies $wQ = P$. We shall denote this permutation $w$ by $w_{P,Q}$.

Thus, $w_{P,Q} \in S_n$ is the permutation satisfying

$$w_{P,Q} Q = P. \tag{269}$$

**Example 5.15.3.** Let $n = 4$ and $D = Y(2,2)$ and $P = \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 3 \\ \hline \end{array}$ and $Q = \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 4 & 3 \\ \hline \end{array}$. Then, $w_{P,Q} = \operatorname{oln}(4132)$, since $\operatorname{oln}(4132)\, Q = P$.

We now define a slight generalization of the Young symmetrizers $\mathbf{E}_T$:

**Definition 5.15.4.** Let $D$ be any diagram. Let $P$ and $Q$ be two $n$-tableaux of shape $D$. Then, we define an element $\mathbf{E}_{P,Q} \in \mathbf{k}[S_n]$ by

$$\mathbf{E}_{P,Q} := \nabla^{-}_{\operatorname{Col} P} w_{P,Q} \nabla_{\operatorname{Row} Q}.$$

(See Definition 5.5.1 and Definition 5.15.2 for the meaning of the terms used here.)

**Example 5.15.5.** Let $n = 3$ and $D = Y(2,1)$. Then, for the two $n$-tableaux $31\backslash\backslash 2$ and $12\backslash\backslash 3$ of shape $D$, we have

$$\begin{aligned}\mathbf{E}_{31\backslash\backslash 2,\ 12\backslash\backslash 3} &= \underbrace{\nabla^{-}_{\mathrm{Col}(31\backslash\backslash 2)}}_{\substack{=1-t_{3,2}\\=1-s_2}} \underbrace{w_{31\backslash\backslash 2,\ 12\backslash\backslash 3}}_{\substack{=\mathrm{oln}(312)\\=\mathrm{cyc}_{1,3,2}\\=s_2s_1}} \underbrace{\nabla_{\mathrm{Row}(12\backslash\backslash 3)}}_{\substack{=1+t_{1,2}\\=1+s_1}} \\ &= \underbrace{(1-s_2)\,s_2}_{=s_2-1} \underbrace{s_1\,(1+s_1)}_{=s_1+1} = (s_2-1)(s_1+1) = s_2s_1 - s_1 + s_2 - 1.\end{aligned}$$

The two standard tableaux of shape $D$ are $12\backslash\backslash 3$ and $13\backslash\backslash 2$. The corresponding $\mathbf{E}_{P,Q}$ are

$$\begin{aligned}\mathbf{E}_{12\backslash\backslash 3,\ 12\backslash\backslash 3} &= 1 - t_{1,3} + t_{1,2} - \mathrm{cyc}_{1,2,3};\\ \mathbf{E}_{12\backslash\backslash 3,\ 13\backslash\backslash 2} &= t_{2,3} - \mathrm{cyc}_{1,3,2} + \mathrm{cyc}_{1,2,3} - t_{1,2};\\ \mathbf{E}_{13\backslash\backslash 2,\ 12\backslash\backslash 3} &= t_{2,3} - \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2} - t_{1,3};\\ \mathbf{E}_{13\backslash\backslash 2,\ 13\backslash\backslash 2} &= 1 - t_{1,2} + t_{1,3} - \mathrm{cyc}_{1,3,2}\,.\end{aligned}$$

**Remark 5.15.6.** The $w_{P,Q}$ and the $\mathbf{E}_{P,Q}$ go all the way back to Young's 1928 paper [Young77, *On Quantitative Substitutional Analysis (Third Paper)*, §4]. A more recent source is Rutherford's [Ruther48, §8 and §11], which studies them in the case when the shape $D$ is the Young diagram $Y(\alpha)$ of a partition $\alpha$. In this setting, Rutherford numbers all $n$-tableaux of shape $D$ by $S^{\alpha}_1, S^{\alpha}_2, \ldots, S^{\alpha}_{n!}$ (in an arbitrary but fixed order), and uses the notations $\sigma^{\alpha}_{rs}$ and $E^{\alpha}_{rs}$ for what we call $w_{P,Q}$ and $\mathbf{E}_{P,Q}$ when $P = S^{\alpha}_r$ and $Q = S^{\alpha}_s$. (He often omits the superscript $\alpha$.)

The elements $\mathbf{E}_{P,Q}$ we introduced in Definition 5.15.4 are very close to the Young symmetrizers $\mathbf{E}_P$ (which is why we had no need to introduce them until now). Indeed, we can describe the former through the latter in two easy ways:

**Proposition 5.15.7.** Let $D$ be any diagram. Let $P$ and $Q$ be two $n$-tableaux of shape $D$. Then,

$$\mathbf{E}_{P,Q} = w_{P,Q}\mathbf{E}_Q \tag{270}$$
$$= \mathbf{E}_P w_{P,Q}. \tag{271}$$

*Proof.* Let us denote the permutation $w_{P,Q} \in S_n$ by $w$. Then, $w = w_{P,Q}$, so that $wQ = w_{P,Q}Q = P$ (by (269)). In other words, $w \rightharpoonup Q = P$ (since $wQ$ is just a shorthand for $w \rightharpoonup Q$). But Proposition 5.5.11 (applied to $T = Q$) yields $\nabla_{\mathrm{Row}(w\rightharpoonup Q)} = w\nabla_{\mathrm{Row}\,Q}w^{-1}$ and $\nabla^{-}_{\mathrm{Col}(w\rightharpoonup Q)} = w\nabla^{-}_{\mathrm{Col}\,Q}w^{-1}$. In view of

$w \rightharpoonup Q = P$, we can rewrite these two equalities as

$$\nabla_{\operatorname{Row} P} = w \nabla_{\operatorname{Row} Q} w^{-1} \qquad \text{and} \qquad \nabla^{-}_{\operatorname{Col} P} = w \nabla^{-}_{\operatorname{Col} Q} w^{-1}.$$

Now, the definition of $\mathbf{E}_Q$ yields $\mathbf{E}_Q = \nabla^{-}_{\operatorname{Col} Q} \nabla_{\operatorname{Row} Q}$. Likewise, we obtain $\mathbf{E}_P = \nabla^{-}_{\operatorname{Col} P} \nabla_{\operatorname{Row} P}$. But the definition of $\mathbf{E}_{P,Q}$ yields

$$\begin{aligned}
\mathbf{E}_{P,Q} &= \underbrace{\nabla^{-}_{\operatorname{Col} P}}_{= w \nabla^{-}_{\operatorname{Col} Q} w^{-1}} \underbrace{w_{P,Q}}_{= w} \nabla_{\operatorname{Row} Q} \\
&= w \nabla^{-}_{\operatorname{Col} Q} \underbrace{w^{-1} w}_{=1} \nabla_{\operatorname{Row} Q} = w \underbrace{\nabla^{-}_{\operatorname{Col} Q} \nabla_{\operatorname{Row} Q}}_{= \mathbf{E}_Q} = \underbrace{w}_{= w_{P,Q}} \mathbf{E}_Q = w_{P,Q} \mathbf{E}_Q
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{E}_{P,Q} &= \nabla^{-}_{\operatorname{Col} P} \underbrace{w_{P,Q}}_{= w} \underbrace{\nabla_{\operatorname{Row} Q}}_{\substack{= w^{-1} \nabla_{\operatorname{Row} P} w \\ (\text{since } \nabla_{\operatorname{Row} P} = w \nabla_{\operatorname{Row} Q} w^{-1})}} \\
&= \nabla^{-}_{\operatorname{Col} P} \underbrace{w w^{-1}}_{=1} \nabla_{\operatorname{Row} P} w = \underbrace{\nabla^{-}_{\operatorname{Col} P} \nabla_{\operatorname{Row} P}}_{= \mathbf{E}_P} w = \mathbf{E}_P \underbrace{w}_{= w_{P,Q}} = \mathbf{E}_P w_{P,Q}.
\end{aligned}$$

Together, these two equalities prove Proposition 5.15.7. □

**Exercise 5.15.1.** [1] Let $D$ be any diagram. Let $P$ and $Q$ be two $n$-tableaux of shape $D$. Let $u, v \in S_n$. Prove that

$$u \mathbf{E}_{P,Q} v = \mathbf{E}_{uP, v^{-1}Q}.$$

**Exercise 5.15.2.** [1] Let $\lambda$ be a partition of $n$. Let $P$, $Q$ and $R$ be three $n$-tableaux of shape $\lambda$. Show that $\mathbf{E}_{P,Q} \mathbf{E}_{Q,R} = h^{\lambda} \mathbf{E}_{P,R}$.

The Young symmetrizers $\mathbf{E}_P$ are particular cases of the newly-defined elements $\mathbf{E}_{P,P}$:

**Proposition 5.15.8.** Let $D$ be any diagram. Let $P$ be an $n$-tableau of shape $D$. Then, $w_{P,P} = \mathrm{id}$ and $\mathbf{E}_{P,P} = \mathbf{E}_P$.

*Proof.* Definition 5.15.2 tells us that $w_{P,P}$ is defined as the unique permutation $w \in S_n$ that satisfies $wP = P$. Hence, if some permutation $w \in S_n$ satisfies $wP = P$, then $w = w_{P,P}$ (by the uniqueness claim in the preceding sentence). Applying this to $w = \mathrm{id}$, we obtain $\mathrm{id} = w_{P,P}$ (since $\mathrm{id}\, P = P$). Thus, we conclude that $w_{P,P} = \mathrm{id}$.

Furthermore, Definition 5.15.4 yields $\mathbf{E}_{P,P} = \nabla^{-}_{\operatorname{Col} P} \underbrace{w_{P,P}}_{= \mathrm{id}} \nabla_{\operatorname{Row} P} = \nabla^{-}_{\operatorname{Col} P} \nabla_{\operatorname{Row} P} = \mathbf{E}_P$ (since $\mathbf{E}_P$ was defined to be $\nabla^{-}_{\operatorname{Col} P} \nabla_{\operatorname{Row} P}$). The proof of Proposition 5.15.8 is thus complete. □

One reason we introduced the $\mathbf{E}_{P,Q}$'s is that they form a basis of the left ideal avatar $\mathcal{A}\mathbf{E}_T$ of a skew Specht module $\mathcal{S}^{Y(\lambda/\mu)}$:

**Proposition 5.15.9.** Let $D$ be a skew Young diagram. Let $T$ be an $n$-tableau of shape $D$. Let $\operatorname{SYT}(D)$ be the set of all standard $n$-tableaux of shape $D$. Then, the family $(\mathbf{E}_{P,T})_{P\in\operatorname{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_T$.

*Proof.* This is essentially just the standard basis theorem (Theorem 5.9.1), after applying the isomorphism $\mathcal{S}^D \to \mathcal{A}\mathbf{E}_T$. In more detail:

We know that $T$ is an $n$-tableau of shape $D$, thus a bijection from $D$ to $[n]$. Hence, $|D| = |[n]| = n$.

Theorem 5.9.1 says that the standard polytabloids $\mathbf{e}_P$ (that is, the $\mathbf{e}_P$ where $P$ ranges over all standard tableaux of shape $D$) form a basis of the $\mathbf{k}$-module $\mathcal{S}^D$. In other words, the $\mathbf{e}_P$ for $P \in \operatorname{SYT}(D)$ form a basis of the $\mathbf{k}$-module $\mathcal{S}^D$ (since $\operatorname{SYT}(D)$ is the set of all standard $n$-tableaux of shape $D$). In other words, the family $(\mathbf{e}_P)_{P\in\operatorname{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{S}^D$.

Corollary 5.5.23 shows that there is a left $\mathbf{k}[S_n]$-module isomorphism

$$\overline{\alpha} : \mathbf{k}[S_n]\cdot \nabla^-_{\operatorname{Col} T}\nabla_{\operatorname{Row} T} \to \mathcal{S}^D$$

that sends $\nabla^-_{\operatorname{Col} T}\nabla_{\operatorname{Row} T}$ to $\mathbf{e}_T$. In other words, there is a left $\mathcal{A}$-module isomorphism

$$\overline{\alpha} : \mathcal{A}\mathbf{E}_T \to \mathcal{S}^D$$

that sends $\mathbf{E}_T$ to $\mathbf{e}_T$ (since $\mathbf{k}[S_n] = \mathcal{A}$ and $\nabla^-_{\operatorname{Col} T}\nabla_{\operatorname{Row} T} = \mathbf{E}_T$). Consider this isomorphism $\overline{\alpha}$. Thus, $\overline{\alpha}(\mathbf{E}_T) = \mathbf{e}_T$ (since $\overline{\alpha}$ sends $\mathbf{E}_T$ to $\mathbf{e}_T$).

Next, let $P \in \operatorname{SYT}(D)$. Then, Proposition (270) (applied to $Q = T$) yields $\mathbf{E}_{P,T} = w_{P,T}\mathbf{E}_T$. Hence,

$$\begin{aligned}\overline{\alpha}(\mathbf{E}_{P,T}) &= \overline{\alpha}(w_{P,T}\mathbf{E}_T) = w_{P,T}\underbrace{\overline{\alpha}(\mathbf{E}_T)}_{=\mathbf{e}_T} \qquad \text{(since } \overline{\alpha} \text{ is a left } \mathcal{A}\text{-module morphism)}\\ &= w_{P,T}\mathbf{e}_T = \mathbf{e}_{w_{P,T}T} \qquad (272)\end{aligned}$$

(since Lemma 5.4.6 **(a)** (applied to $u = w_{P,T}$) yields $\mathbf{e}_{w_{P,T}T} = w_{P,T}\mathbf{e}_T$). But (269) (applied to $Q = T$) yields $w_{P,T}T = P$. Hence, we can rewrite (272) as $\overline{\alpha}(\mathbf{E}_{P,T}) = \mathbf{e}_P$. Therefore, $\overline{\alpha}^{-1}(\mathbf{e}_P) = \mathbf{E}_{P,T}$ (since $\overline{\alpha}$ is an isomorphism and thus invertible).

Forget that we fixed $P$. We thus have shown that

$$\overline{\alpha}^{-1}(\mathbf{e}_P) = \mathbf{E}_{P,T} \qquad \text{for each } P \in \operatorname{SYT}(D). \qquad (273)$$

In other words,

$$\left(\overline{\alpha}^{-1}(\mathbf{e}_P)\right)_{P\in\operatorname{SYT}(D)} = (\mathbf{E}_{P,T})_{P\in\operatorname{SYT}(D)}. \qquad (274)$$

But the map $\overline{\alpha} : \mathcal{A}\mathbf{E}_T \to \mathcal{S}^D$ is a $\mathbf{k}$-module isomorphism (since it is a left $\mathcal{A}$-module isomorphism). Hence, it has an inverse $\overline{\alpha}^{-1} : \mathcal{S}^D \to \mathcal{A}\mathbf{E}_T$, which is

also a $\mathbf{k}$-module isomorphism. It is well-known that a $\mathbf{k}$-module isomorphism sends any basis of its domain to a basis of its target. Applying this to the $\mathbf{k}$-module isomorphism $\overline{\alpha}^{-1}$, we see that $\overline{\alpha}^{-1}$ sends any basis of $\mathcal{S}^D$ to a basis of $\mathcal{A}\mathbf{E}_T$. In other words, if $(v_i)_{i\in I}$ is a basis of $\mathcal{S}^D$, then $\left(\overline{\alpha}^{-1}(v_i)\right)_{i\in I}$ is a basis of $\mathcal{A}\mathbf{E}_T$.

Applying this to $(v_i)_{i\in I} = (\mathbf{e}_P)_{P\in \mathrm{SYT}(D)}$, we conclude that the family $\left(\overline{\alpha}^{-1}(\mathbf{e}_P)\right)_{P\in\mathrm{SYT}(D)}$ is a basis of $\mathcal{A}\mathbf{E}_T$ (since $(\mathbf{e}_P)_{P\in\mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{S}^D$). In view of (274), we can rewrite this as follows: The family $(\mathbf{E}_{P,T})_{P\in\mathrm{SYT}(D)}$ is a basis of $\mathcal{A}\mathbf{E}_T$. This proves Proposition 5.15.9. □

Thus, we found an explicit basis of $\mathcal{A}\mathbf{E}_T$. This is not very surprising, as $\mathcal{A}\mathbf{E}_T$ is isomorphic to the Specht module $\mathcal{S}^D$. What is more interesting is to find an explicit basis for $\mathcal{A}\mathbf{E}_\lambda$ when $\lambda$ is a partition of $n$. This will take us the next few pages (and require some invertibility conditions on $\mathbf{k}$). As an intermediate step, let us find a basis of $\mathbf{E}_T\mathcal{A}$. This is quite similar to $\mathcal{A}\mathbf{E}_T$, and indeed we will derive a basis of $\mathbf{E}_T\mathcal{A}$ from our above basis of $\mathcal{A}\mathbf{E}_T$. To do so, we shall now briefly study transposition of Young tableaux, which will help us relate the Young symmetrizers $\mathbf{E}_T$ to the antipode of $\mathcal{A}$.

### 5.15.2. Transposes of Young tableaux

In Theorem 5.1.10, we defined the map $\mathbf{r} : \mathbb{Z}^2 \to \mathbb{Z}^2$, which sends each cell $(i,j) \in \mathbb{Z}^2$ to $(j,i)$. Geometrically, this map is the reflection across the northwest-to-southeast diagonal. We observed that the image of a Young diagram $Y(\lambda)$ under this map $\mathbf{r}$ is a Young diagram again, namely the Young diagram of the partition $\lambda^t$ known as the conjugate of $\lambda$ (see Theorem 5.1.10 **(a)**). It is also clear that the map $\mathbf{r}$ is an involution (i.e., it is inverse to itself, i.e., it satisfies $\mathbf{r}\circ\mathbf{r} = \mathrm{id}$), and thus is a bijection.

We shall now use the map $\mathbf{r}$ to reflect a tableau across the northwest-to-southeast diagonal:

**Definition 5.15.10.** Let $D$ be any diagram. Consider the map $\mathbf{r} : \mathbb{Z}^2 \to \mathbb{Z}^2$ from Theorem 5.1.10. Then, $\mathbf{r}(\mathbf{r}(D)) = D$ (since $\mathbf{r}$ is an involution). Hence, we can restrict the bijection $\mathbf{r}$ to obtain a bijection $\mathbf{r}|_{\mathbf{r}(D)} : \mathbf{r}(D) \to D$.

Now, let $T$ be an $n$-tableau of shape $D$. Thus, $T$ is a bijection from $D$ to $[n]$. Composing this bijection $T$ with the bijection $\mathbf{r}|_{\mathbf{r}(D)} : \mathbf{r}(D) \to D$, we obtain a bijection $T \circ \left(\mathbf{r}|_{\mathbf{r}(D)}\right) : \mathbf{r}(D) \to [n]$, which we shall call $T\mathbf{r}$. Thus, $T\mathbf{r}$ is an $n$-tableau of shape $\mathbf{r}(D)$. Explicitly, the entries of this $n$-tableau $T\mathbf{r}$ are given

by

$$\begin{aligned}(T\mathbf{r})(i,j) &= T\left(\left(\mathbf{r}\mid_{\mathbf{r}(D)}\right)(i,j)\right) \qquad \left(\text{since } T\mathbf{r} = T\circ\left(\mathbf{r}\mid_{\mathbf{r}(D)}\right)\right)\\ &= T(j,i) \qquad \left(\begin{array}{c}\text{since } \left(\mathbf{r}\mid_{\mathbf{r}(D)}\right)(i,j) = \mathbf{r}(i,j) = (j,i)\\ \text{(by the definition of } \mathbf{r})\end{array}\right)\\ &\qquad \text{for each } (i,j)\in\mathbf{r}(D).\end{aligned}$$

Thus, informally speaking, the $n$-tableau $T\mathbf{r}$ is obtained by reflecting $T$ across the northwest-to-southeast diagonal.[103] We call $T\mathbf{r}$ the *transpose* of the $n$-tableau $T$.

**Example 5.15.11.** Let $D = Y((4,4,2)/(2))$ and $T = \begin{array}{|c|c|c|c|}\hline \multicolumn{2}{c}{} & 3 & 6\\ \hline 1 & 4 & 5 & 7\\ \hline 2 & 8\\ \cline{1-2}\end{array}$. Then, $T\mathbf{r} = \begin{array}{|c|c|c|}\hline \multicolumn{1}{c}{} & 1 & 2\\ \cline{2-3} \multicolumn{1}{c}{} & 4 & 8\\ \cline{2-3} 3 & 5\\ \cline{1-2} 6 & 7\\ \cline{1-2}\end{array}$.

**Proposition 5.15.12.** Let $D$ be any diagram. Then:

**(a)** If $T$ is a standard $n$-tableau of shape $D$, then $T\mathbf{r}$ is a standard $n$-tableau of shape $\mathbf{r}(D)$.

**(b)** If $T$ is any $n$-tableau of shape $D$, then $(T\mathbf{r})\,\mathbf{r} = T$.

**(c)** For any diagram $E$, let $\mathrm{SYT}(E)$ be the set of all standard $n$-tableaux of shape $E$. Then, the map

$$\begin{aligned}\mathrm{SYT}(D) &\to \mathrm{SYT}(\mathbf{r}(D)),\\ T &\mapsto T\mathbf{r}\end{aligned}$$

is a bijection.

**(d)** If $T$ is any $n$-tableau of shape $D$, and if $w\in S_n$ is arbitrary, then $(wT)\,\mathbf{r} = w(T\mathbf{r})$.

**(e)** If $T$ is any $n$-tableau of shape $D$, then $\mathcal{R}(T) = \mathcal{C}(T\mathbf{r})$ and $\mathcal{C}(T) = \mathcal{R}(T\mathbf{r})$.

The proof of Proposition 5.15.12 is essentially trivial, exploiting obvious symmetries of the definitions and the fact that $\mathbf{r}$ is an involution. Thus we are banishing it to the appendix (Section A.2).

[103] No, you don't replace 8's by $\infty$'s.

**Corollary 5.15.13.** Let $\lambda$ be a partition of $n$. Then, $f^\lambda = f^{\lambda^t}$. (See Theorem 5.1.10 **(a)** for the meaning of $\lambda^t$.)

*Proof.* We have $|Y(\lambda)| = |\lambda| = n$ (since $\lambda$ is a partition of $n$). Hence, Proposition 5.2.6 (applied to $D = Y(\lambda)$) shows that the standard tableaux of shape $Y(\lambda)$ are precisely the standard $n$-tableaux of shape $Y(\lambda)$.

We shall use the notation $\mathrm{SYT}(E)$ from Proposition 5.15.12 **(c)**. Applying Proposition 5.15.12 **(c)** to $D = Y(\lambda)$, we obtain that the map

$$\begin{aligned} \mathrm{SYT}(Y(\lambda)) &\to \mathrm{SYT}(\mathbf{r}(Y(\lambda))), \\ T &\mapsto T\mathbf{r} \end{aligned}$$

is a bijection. Hence, $|\mathrm{SYT}(Y(\lambda))| = |\mathrm{SYT}(\mathbf{r}(Y(\lambda)))|$. In other words, $|\mathrm{SYT}(Y(\lambda))| = \left|\mathrm{SYT}\left(Y\left(\lambda^t\right)\right)\right|$ (since the definition of $\lambda^t$ yields $Y\left(\lambda^t\right) = \mathbf{r}(Y(\lambda))$).

However, $f^\lambda$ is defined as the # of standard tableaux of shape $Y(\lambda)$. In other words, $f^\lambda$ is the # of standard $n$-tableaux of shape $Y(\lambda)$ (since the standard tableaux of shape $Y(\lambda)$ are precisely the standard $n$-tableaux of shape $Y(\lambda)$). In other words, $f^\lambda = |\mathrm{SYT}(Y(\lambda))|$ (since $\mathrm{SYT}(Y(\lambda))$ is defined as the set of all standard $n$-tableaux of shape $Y(\lambda)$). The same argument (applied to $\lambda^t$ instead of $\lambda$) yields $f^{\lambda^t} = \left|\mathrm{SYT}\left(Y\left(\lambda^t\right)\right)\right|$ (since Theorem 5.1.10 **(c)** yields $\left|\lambda^t\right| = |\lambda| = n$). Hence,

$$f^\lambda = |\mathrm{SYT}(Y(\lambda))| = \left|\mathrm{SYT}\left(Y\left(\lambda^t\right)\right)\right| = f^{\lambda^t}.$$

This proves Corollary 5.15.13. □

Now, recall two $\mathbf{k}$-linear maps from $\mathbf{k}[S_n]$ to $\mathbf{k}[S_n]$: the sign-twist $T_{\mathrm{sign}}$ defined in Definition 3.11.2, and the antipode $S$ defined in Definition 3.11.12. Using these maps, we can relate the Young symmetrizer of a given $n$-tableau $P$ with that of its transpose $P\mathbf{r}$:

**Proposition 5.15.14.** Let $D$ be any diagram. Let $P$ be any $n$-tableau of shape $D$. Then,

$$\nabla_{\mathrm{Row}(P\mathbf{r})} = T_{\mathrm{sign}}\left(\nabla^-_{\mathrm{Col}\,P}\right); \tag{275}$$

$$\nabla^-_{\mathrm{Col}(P\mathbf{r})} = T_{\mathrm{sign}}\left(\nabla_{\mathrm{Row}\,P}\right); \tag{276}$$

$$\mathbf{E}_{P\mathbf{r}} = T_{\mathrm{sign}}\left(S\left(\mathbf{E}_P\right)\right). \tag{277}$$

*Proof.* Proposition 5.15.12 **(b)** (applied to $T = P$) yields $(P\mathbf{r})\,\mathbf{r} = P$. Proposition 5.15.12 **(e)** (applied to $T = P$) yields $\mathcal{R}(P) = \mathcal{C}(P\mathbf{r})$ and $\mathcal{C}(P) = \mathcal{R}(P\mathbf{r})$.

Now, the definition of $\nabla_{\mathrm{Row}(P\mathbf{r})}$ yields

$$\nabla_{\mathrm{Row}(P\mathbf{r})} = \sum_{w \in \mathcal{R}(P\mathbf{r})} w = \sum_{w \in \mathcal{C}(P)} w \tag{278}$$

(since $\mathcal{R}(P\mathbf{r}) = \mathcal{C}(P)$). On the other hand, the definition of $\nabla^-_{\operatorname{Col} P}$ yields $\nabla^-_{\operatorname{Col} P} = \sum_{w \in \mathcal{C}(P)} (-1)^w w$. Applying the map $T_{\operatorname{sign}}$ to this equality, we obtain

$$\begin{aligned} T_{\operatorname{sign}}\left(\nabla^-_{\operatorname{Col} P}\right) &= T_{\operatorname{sign}}\left(\sum_{w \in \mathcal{C}(P)} (-1)^w w\right) \\ &= \sum_{w \in \mathcal{C}(P)} (-1)^w \underbrace{T_{\operatorname{sign}}(w)}_{\substack{=(-1)^w w \\ \text{(by the definition} \\ \text{of } T_{\operatorname{sign}})}} \qquad \left(\text{since the map } T_{\operatorname{sign}} \text{ is } \mathbf{k}\text{-linear}\right) \\ &= \sum_{w \in \mathcal{C}(P)} \underbrace{(-1)^w (-1)^w}_{\substack{=\left((-1)^w\right)^2 = 1 \\ \text{(since } (-1)^w \in \{1,-1\}\text{)}}} w = \sum_{w \in \mathcal{C}(P)} w. \end{aligned}$$

Comparing this with (278), we obtain

$$\nabla_{\operatorname{Row}(P\mathbf{r})} = T_{\operatorname{sign}}\left(\nabla^-_{\operatorname{Col} P}\right).$$

Thus, (275) is proved.

Now, observe that $P\mathbf{r}$ is an $n$-tableau of shape $\mathbf{r}(D)$. Hence, applying (275) to $\mathbf{r}(D)$ and $P\mathbf{r}$ instead of $D$ and $P$, we find

$$\nabla_{\operatorname{Row}((P\mathbf{r})\mathbf{r})} = T_{\operatorname{sign}}\left(\nabla^-_{\operatorname{Col}(P\mathbf{r})}\right).$$

In view of $(P\mathbf{r})\,\mathbf{r} = P$, we can rewrite this as

$$\nabla_{\operatorname{Row} P} = T_{\operatorname{sign}}\left(\nabla^-_{\operatorname{Col}(P\mathbf{r})}\right).$$

Applying the map $T_{\operatorname{sign}}$ to both sides of this equality, we obtain

$$\begin{aligned} T_{\operatorname{sign}}\left(\nabla_{\operatorname{Row} P}\right) &= T_{\operatorname{sign}}\left(T_{\operatorname{sign}}\left(\nabla^-_{\operatorname{Col}(P\mathbf{r})}\right)\right) = \underbrace{\left(T_{\operatorname{sign}} \circ T_{\operatorname{sign}}\right)}_{\substack{=\mathrm{id} \\ \text{(by Theorem 3.11.5 } \textbf{(b)}\text{)}}} \left(\nabla^-_{\operatorname{Col}(P\mathbf{r})}\right) \\ &= \mathrm{id}\left(\nabla^-_{\operatorname{Col}(P\mathbf{r})}\right) = \nabla^-_{\operatorname{Col}(P\mathbf{r})}. \end{aligned}$$

This proves (276).

Finally, Proposition 5.5.10 (applied to $T = P$) yields that the antipode $S$ satisfies

$$S\left(\nabla_{\operatorname{Row} P}\right) = \nabla_{\operatorname{Row} P} \qquad \text{and} \qquad S\left(\nabla^-_{\operatorname{Col} P}\right) = \nabla^-_{\operatorname{Col} P}.$$

However, we have $\mathbf{E}_P = \nabla^-_{\operatorname{Col} P} \nabla_{\operatorname{Row} P}$ (by the definition of $\mathbf{E}_P$). Applying the map $S$ to this equality, we obtain

$$S\left(\mathbf{E}_P\right) = S\left(\nabla^-_{\operatorname{Col} P} \nabla_{\operatorname{Row} P}\right) = S\left(\nabla_{\operatorname{Row} P}\right) S\left(\nabla^-_{\operatorname{Col} P}\right)$$

(since Theorem 3.11.14 **(a)** shows that $S$ is a $\mathbf{k}$-algebra anti-morphism). Thus,

$$S\left(\mathbf{E}_P\right) = \underbrace{S\left(\nabla_{\operatorname{Row} P}\right)}_{=\nabla_{\operatorname{Row} P}} \underbrace{S\left(\nabla^-_{\operatorname{Col} P}\right)}_{=\nabla^-_{\operatorname{Col} P}} = \nabla_{\operatorname{Row} P}\nabla^-_{\operatorname{Col} P}.$$

Applying the map $T_{\operatorname{sign}}$ to this equality, we obtain

$$\begin{aligned} T_{\operatorname{sign}}\left(S\left(\mathbf{E}_P\right)\right) &= T_{\operatorname{sign}}\left(\nabla_{\operatorname{Row} P}\nabla^-_{\operatorname{Col} P}\right) \\ &= \underbrace{T_{\operatorname{sign}}\left(\nabla_{\operatorname{Row} P}\right)}_{\substack{=\nabla^-_{\operatorname{Col}(P\mathbf{r})} \\ \text{(by (276))}}} \underbrace{T_{\operatorname{sign}}\left(\nabla^-_{\operatorname{Col} P}\right)}_{\substack{=\nabla_{\operatorname{Row}(P\mathbf{r})} \\ \text{(by (275))}}} \\ &\qquad\qquad \left(\begin{array}{c}\text{since Theorem 3.11.5 \textbf{(a)} shows} \\ \text{that } T_{\operatorname{sign}} \text{ is a } \mathbf{k}\text{-algebra morphism}\end{array}\right) \\ &= \nabla^-_{\operatorname{Col}(P\mathbf{r})}\nabla_{\operatorname{Row}(P\mathbf{r})} = \mathbf{E}_{P\mathbf{r}} \end{aligned}$$

(since $\mathbf{E}_{P\mathbf{r}}$ is defined to be $\nabla^-_{\operatorname{Col}(P\mathbf{r})}\nabla_{\operatorname{Row}(P\mathbf{r})}$). This proves (277), thus concluding the proof of Proposition 5.15.14. □

**Exercise 5.15.3.** [2] Let $\lambda$ be a partition of $n$. Prove that $T_{\operatorname{sign}}\left(\mathbf{E}_\lambda\right) = \mathbf{E}_{\lambda^t}$.

**Exercise 5.15.4.** [2] Let $D$ be a skew Young diagram. Let $T$ be an $n$-tableau of shape $D$. Let $\mathbf{F}_T := \nabla_{\operatorname{Row} T}\nabla^-_{\operatorname{Col} T} \in \mathcal{A}$. Prove that $\mathcal{A}\mathbf{E}_T \cong \mathcal{A}\mathbf{E}_{T\mathbf{r}} \cong \mathcal{A}\mathbf{F}_T$ as $\mathbf{k}$-modules. (But not generally as $\mathcal{A}$-modules, as we saw in Exercise 5.11.3 **(c)**.)

Next, we shall extend (277) to the generalization $\mathbf{E}_{P,Q}$ of the Young symmetrizers. First, however, we prove the following simple properties of the permutations $w_{P,Q}$ from Definition 5.15.2:

**Proposition 5.15.15.** Let $D$ be any diagram. Then:

**(a)** If $P$, $Q$ and $R$ are three $n$-tableaux of shape $D$, then $w_{P,Q}w_{Q,R} = w_{P,R}$.

**(b)** If $P$ and $Q$ are two $n$-tableaux of shape $D$, then $w_{Q,P} = w_{P,Q}^{-1}$.

**(c)** If $P$ and $Q$ are two $n$-tableaux of shape $D$, then $w_{Q,P} = w_{Q\mathbf{r},P\mathbf{r}}$.

**(d)** If $T$ is an $n$-tableau of shape $D$, and if $u \in S_n$, then $w_{uT,T} = u$.

*Proof.* **(a)** Let $P$, $Q$ and $R$ be three $n$-tableaux of shape $D$. Then,

$$w_{P,Q}\underbrace{w_{Q,R}R}_{\substack{=Q \\ \text{(by (269), applied to } Q \text{ and } R \\ \text{instead of } P \text{ and } Q)}} = w_{P,Q}Q = P \qquad \text{(by (269))}.$$

But $w_{P,R}$ is defined as the unique permutation $w \in S_n$ that satisfies $wR = P$. Hence, if a permutation $w \in S_n$ satisfies $wR = P$, then $w = w_{P,R}$. Applying this to $w = w_{P,Q}w_{Q,R}$, we obtain $w_{P,Q}w_{Q,R} = w_{P,R}$ (since $w_{P,Q}w_{Q,R}R = P$). This proves Proposition 5.15.15 **(a)**.

**(b)** Let $P$ and $Q$ be two $n$-tableaux of shape $D$. Then, Proposition 5.15.15 **(a)** (applied to $R = P$) yields $w_{P,Q}w_{Q,P} = w_{P,P} = \mathrm{id}$ (by Proposition 5.15.8). Hence, $w_{Q,P} = w_{P,Q}^{-1}$. This proves Proposition 5.15.15 **(b)**.

**(c)** Let $P$ and $Q$ be two $n$-tableaux of shape $D$. Then, (269) (applied to $Q$ and $P$ instead of $P$ and $Q$) yields $w_{Q,P}P = Q$. However, Proposition 5.15.12 **(d)** (applied to $P$ and $w_{Q,P}$ instead of $T$ and $w$) yields $\left(w_{Q,P}P\right)\mathbf{r} = w_{Q,P}\left(P\mathbf{r}\right)$. Hence, $w_{Q,P}\left(P\mathbf{r}\right) = \underbrace{\left(w_{Q,P}P\right)}_{=Q}\mathbf{r} = Q\mathbf{r}$.

However, $w_{Q\mathbf{r},P\mathbf{r}}$ is defined as the unique permutation $w \in S_n$ that satisfies $w\left(P\mathbf{r}\right) = Q\mathbf{r}$. Hence, if a permutation $w \in S_n$ satisfies $w\left(P\mathbf{r}\right) = Q\mathbf{r}$, then $w = w_{Q\mathbf{r},P\mathbf{r}}$. Applying this to $w = w_{Q,P}$, we obtain $w_{Q,P} = w_{Q\mathbf{r},P\mathbf{r}}$ (since $w_{Q,P}\left(P\mathbf{r}\right) = Q\mathbf{r}$). Proposition 5.15.15 **(c)** is now proved.

**(d)** Let $T$ be an $n$-tableau of shape $D$. Let $u \in S_n$. Then, $w_{uT,T}$ is defined as the unique permutation $w \in S_n$ that satisfies $wT = uT$. Hence, if a permutation $w \in S_n$ satisfies $wT = uT$, then $w = w_{uT,T}$. Applying this to $w = u$, we obtain $u = w_{uT,T}$ (since $uT = uT$). In other words, $w_{uT,T} = u$. This proves Proposition 5.15.15 **(d)**. □

We are now ready to prove the generalized version of (277):

**Proposition 5.15.16.** Let $D$ be any diagram. Let $P$ and $Q$ be two $n$-tableaux of shape $D$. Then,

$$(-1)^{w_{Q,P}}\mathbf{E}_{Q\mathbf{r},P\mathbf{r}} = T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{P,Q}\right)\right).$$

*Proof.* The definition of $S$ yields $S\left(w_{P,Q}\right) = w_{P,Q}^{-1}$. But Proposition 5.15.15 **(b)** yields $w_{Q,P} = w_{P,Q}^{-1}$. Comparing these two equalities, we find $S\left(w_{P,Q}\right) = w_{Q,P}$.

The definition of $\mathbf{E}_{P,Q}$ says that $\mathbf{E}_{P,Q} = \nabla^{-}_{\mathrm{Col}\,P}w_{P,Q}\nabla_{\mathrm{Row}\,Q}$. Applying the map $S$ to this equality, we obtain

$$\begin{aligned}
S\left(\mathbf{E}_{P,Q}\right) &= S\left(\nabla^{-}_{\mathrm{Col}\,P}w_{P,Q}\nabla_{\mathrm{Row}\,Q}\right) \\
&= \underbrace{S\left(\nabla_{\mathrm{Row}\,Q}\right)}_{\substack{=\nabla_{\mathrm{Row}\,Q} \\ \text{(by Proposition 5.5.10,} \\ \text{applied to } T=Q\text{)}}} \underbrace{S\left(w_{P,Q}\right)}_{=w_{Q,P}} \underbrace{S\left(\nabla^{-}_{\mathrm{Col}\,P}\right)}_{\substack{=\nabla^{-}_{\mathrm{Col}\,P} \\ \text{(by Proposition 5.5.10,} \\ \text{applied to } T=P\text{)}}} \\
&\qquad \left(\begin{array}{c}\text{since Theorem 3.11.14 \textbf{(a)} shows} \\ \text{that } S \text{ is a } \mathbf{k}\text{-algebra anti-morphism}\end{array}\right) \\
&= \nabla_{\mathrm{Row}\,Q}w_{Q,P}\nabla^{-}_{\mathrm{Col}\,P}.
\end{aligned}$$

Applying the map $T_{\mathrm{sign}}$ to this equality, we obtain

$$\begin{aligned}
&T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{P,Q}\right)\right)\\
&= T_{\mathrm{sign}}\left(\nabla_{\mathrm{Row}\,Q} w_{Q,P} \nabla^{-}_{\mathrm{Col}\,P}\right)\\
&= \underbrace{T_{\mathrm{sign}}\left(\nabla_{\mathrm{Row}\,Q}\right)}_{\substack{=\nabla^{-}_{\mathrm{Col}(Q\mathbf{r})}\\ \text{(since (276) (applied to } Q \text{ instead of } P\text{)}\\ \text{yields } \nabla^{-}_{\mathrm{Col}(Q\mathbf{r})} = T_{\mathrm{sign}}\left(\nabla_{\mathrm{Row}\,Q}\right)\text{)}}}
\underbrace{T_{\mathrm{sign}}\left(w_{Q,P}\right)}_{\substack{=(-1)^{w_{Q,P}} w_{Q,P}\\ \text{(by the definition of } T_{\mathrm{sign}}\text{)}}}
\underbrace{T_{\mathrm{sign}}\left(\nabla^{-}_{\mathrm{Col}\,P}\right)}_{\substack{=\nabla_{\mathrm{Row}(P\mathbf{r})}\\ \text{(by (275))}}}\\
&\qquad\qquad \left(\begin{array}{c}\text{since Theorem 3.11.5 \textbf{(a)} shows}\\ \text{that } T_{\mathrm{sign}} \text{ is a } \mathbf{k}\text{-algebra morphism}\end{array}\right)\\
&= \nabla^{-}_{\mathrm{Col}(Q\mathbf{r})} \left(-1\right)^{w_{Q,P}} \underbrace{w_{Q,P}}_{\substack{=w_{Q\mathbf{r},P\mathbf{r}}\\ \text{(by Proposition 5.15.15 \textbf{(c)})}}} \nabla_{\mathrm{Row}(P\mathbf{r})}\\
&= \nabla^{-}_{\mathrm{Col}(Q\mathbf{r})} \left(-1\right)^{w_{Q,P}} w_{Q\mathbf{r},P\mathbf{r}} \nabla_{\mathrm{Row}(P\mathbf{r})}\\
&= \left(-1\right)^{w_{Q,P}} \cdot \underbrace{\nabla^{-}_{\mathrm{Col}(Q\mathbf{r})} w_{Q\mathbf{r},P\mathbf{r}} \nabla_{\mathrm{Row}(P\mathbf{r})}}_{\substack{=\mathbf{E}_{Q\mathbf{r},P\mathbf{r}}\\ \text{(since } \mathbf{E}_{Q\mathbf{r},P\mathbf{r}} \text{ is defined to}\\ \text{be } \nabla^{-}_{\mathrm{Col}(Q\mathbf{r})} w_{Q\mathbf{r},P\mathbf{r}} \nabla_{\mathrm{Row}(P\mathbf{r})}\text{)}}} = \left(-1\right)^{w_{Q,P}} \mathbf{E}_{Q\mathbf{r},P\mathbf{r}}.
\end{aligned}$$

This proves Proposition 5.15.16. □

### 5.15.3. Straightening left and right

Combining the above properties of transposed tableaux with Proposition 5.15.9, we obtain the following:

**Lemma 5.15.17.** Let $D$ be a skew Young diagram. Let $Q$ be an $n$-tableau of shape $D$. Let $\mathrm{SYT}\left(D\right)$ be the set of all standard $n$-tableaux of shape $D$. Then,

$$\mathcal{A}\mathbf{E}_Q = \mathrm{span}_{\mathbf{k}}\left\{\mathbf{E}_{P,Q} \;\mid\; P \in \mathrm{SYT}\left(D\right)\right\} \tag{279}$$

and

$$\mathbf{E}_Q\mathcal{A} = \mathrm{span}_{\mathbf{k}}\left\{\mathbf{E}_{Q,P} \;\mid\; P \in \mathrm{SYT}\left(D\right)\right\}. \tag{280}$$

*Proof.* We know that $Q$ is an $n$-tableau of shape $D$, thus a bijection from $D$ to $[n]$. Hence, $|D| = |[n]| = n$. But $\mathbf{r}$ is a bijection; thus, $|\mathbf{r}\left(D\right)| = |D| = n$ as well.

Proposition 5.15.9 (applied to $T = Q$) yields that the family $\left(\mathbf{E}_{P,Q}\right)_{P\in\mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_Q$. Thus, this family spans $\mathcal{A}\mathbf{E}_Q$. In other words, $\mathcal{A}\mathbf{E}_Q = \mathrm{span}_{\mathbf{k}}\left\{\mathbf{E}_{P,Q} \;\mid\; P \in \mathrm{SYT}\left(D\right)\right\}$. This proves (279).

It remains to prove (280). To do so, we will essentially apply (279) to $\mathbf{r}\left(D\right)$ and $Q\mathbf{r}$ instead of $D$ and $Q$, and then transform the result using the antipode

$S$ (which is an anti-morphism and thus turns $\mathcal{A}\mathbf{E}_{Q\mathbf{r}}$ into $S\left(\mathbf{E}_{Q\mathbf{r}}\right)\mathcal{A}$) and the sign-twist $T_{\mathrm{sign}}$. Here are the details:

For any diagram $E$, let $\mathrm{SYT}\left(E\right)$ be the set of all standard $n$-tableaux of shape $E$. (This generalizes the notation $\mathrm{SYT}\left(D\right)$ defined in the lemma.) Proposition 5.15.12 **(c)** then says that the map

$$\begin{aligned}\mathrm{SYT}\left(D\right) &\to \mathrm{SYT}\left(\mathbf{r}\left(D\right)\right), \\ T &\mapsto T\mathbf{r}\end{aligned}$$

is a bijection.

But $D$ is a skew Young diagram. Hence, $\mathbf{r}\left(D\right)$ is a skew Young diagram as well[104]. Therefore, (279) (applied to $\mathbf{r}\left(D\right)$ and $Q\mathbf{r}$ instead of $D$ and $Q$) yields

$$\mathcal{A}\mathbf{E}_{Q\mathbf{r}} = \mathrm{span}_{\mathbf{k}}\left\{\mathbf{E}_{P,Q\mathbf{r}} \;\mid\; P \in \mathrm{SYT}\left(\mathbf{r}\left(D\right)\right)\right\}$$

(since $Q\mathbf{r}$ is an $n$-tableau of shape $\mathbf{r}\left(D\right)$). Applying the map $S$ to both sides of this equality, we obtain

$$\begin{aligned}S\left(\mathcal{A}\mathbf{E}_{Q\mathbf{r}}\right) &= S\left(\mathrm{span}_{\mathbf{k}}\left\{\mathbf{E}_{P,Q\mathbf{r}} \;\mid\; P \in \mathrm{SYT}\left(\mathbf{r}\left(D\right)\right)\right\}\right) \\ &= \mathrm{span}_{\mathbf{k}}\left\{S\left(\mathbf{E}_{P,Q\mathbf{r}}\right) \;\mid\; P \in \mathrm{SYT}\left(\mathbf{r}\left(D\right)\right)\right\} \qquad \left(\begin{array}{c}\text{since the map } S \\ \text{is } \mathbf{k}\text{-linear}\end{array}\right) \\ &= \mathrm{span}_{\mathbf{k}}\left\{S\left(\mathbf{E}_{T\mathbf{r},Q\mathbf{r}}\right) \;\mid\; T \in \mathrm{SYT}\left(D\right)\right\} \\ &\qquad\qquad \left(\begin{array}{c}\text{here, we have substituted } T\mathbf{r} \text{ for } P\text{, since} \\ \text{the map } \mathrm{SYT}\left(D\right) \to \mathrm{SYT}\left(\mathbf{r}\left(D\right)\right),\ T \mapsto T\mathbf{r} \\ \text{is a bijection}\end{array}\right) \\ &= \mathrm{span}_{\mathbf{k}}\left\{S\left(\mathbf{E}_{P\mathbf{r},Q\mathbf{r}}\right) \;\mid\; P \in \mathrm{SYT}\left(D\right)\right\}\end{aligned}$$

(here, we have renamed the index $T$ as $P$). Applying the map $T_{\mathrm{sign}}$ to this equality, we furthermore obtain

$$\begin{aligned}&T_{\mathrm{sign}}\left(S\left(\mathcal{A}\mathbf{E}_{Q\mathbf{r}}\right)\right) \\ &= T_{\mathrm{sign}}\left(\mathrm{span}_{\mathbf{k}}\left\{S\left(\mathbf{E}_{P\mathbf{r},Q\mathbf{r}}\right) \;\mid\; P \in \mathrm{SYT}\left(D\right)\right\}\right) \\ &= \mathrm{span}_{\mathbf{k}}\left\{T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{P\mathbf{r},Q\mathbf{r}}\right)\right) \;\mid\; P \in \mathrm{SYT}\left(D\right)\right\} \qquad (281)\end{aligned}$$

(since the map $T_{\mathrm{sign}}$ is $\mathbf{k}$-linear).

Now, let $P \in \mathrm{SYT}\left(D\right)$. Then, $P\mathbf{r}$ and $Q\mathbf{r}$ are two $n$-tableaux of shape $\mathbf{r}\left(D\right)$ (since $P$ and $Q$ are two $n$-tableaux of shape $D$). Thus, Proposition 5.15.16 (applied to $\mathbf{r}\left(D\right)$, $P\mathbf{r}$ and $Q\mathbf{r}$ instead of $D$, $P$ and $Q$) yields

$$\left(-1\right)^{w_{Q\mathbf{r},P\mathbf{r}}}\mathbf{E}_{\left(Q\mathbf{r}\right)\mathbf{r},\left(P\mathbf{r}\right)\mathbf{r}} = T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{P\mathbf{r},Q\mathbf{r}}\right)\right). \qquad (282)$$

[104] *Proof.* We assumed that $D$ is a skew Young diagram. In other words, $D = Y\left(\lambda/\mu\right)$ for some skew partition $\lambda/\mu$. Consider this $\lambda/\mu$. Now, Remark 5.1.12 shows that $\lambda^t/\mu^t$ is again a skew partition, and that we have $Y\left(\lambda^t/\mu^t\right) = \mathbf{r}\left(Y\left(\lambda/\mu\right)\right)$. But $D = Y\left(\lambda/\mu\right)$ and therefore $\mathbf{r}\left(D\right) = \mathbf{r}\left(Y\left(\lambda/\mu\right)\right) = Y\left(\lambda^t/\mu^t\right)$. This shows that $\mathbf{r}\left(D\right)$ is a skew Young diagram.

But Proposition 5.15.12 **(b)** yields $(P\mathbf{r})\,\mathbf{r} = P$ and $(Q\mathbf{r})\,\mathbf{r} = Q$. In view of this, we can rewrite (282) as

$$(-1)^{w_{Q\mathbf{r},P\mathbf{r}}}\,\mathbf{E}_{Q,P} = T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{P\mathbf{r},Q\mathbf{r}}\right)\right). \tag{283}$$

Forget that we fixed $P$. We thus have proved (283) for each $P \in \mathrm{SYT}\,(D)$. Thus, (281) becomes

$$\begin{aligned} T_{\mathrm{sign}}\left(S\left(\mathcal{A}\mathbf{E}_{Q\mathbf{r}}\right)\right) &= \mathrm{span}_{\mathbf{k}}\left\{ \underbrace{T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{P\mathbf{r},Q\mathbf{r}}\right)\right)}_{\substack{=(-1)^{w_{Q\mathbf{r},P\mathbf{r}}}\mathbf{E}_{Q,P} \\ \text{(by (283))}}} \;\mid\; P \in \mathrm{SYT}\,(D) \right\} \\ &= \mathrm{span}_{\mathbf{k}}\left\{(-1)^{w_{Q\mathbf{r},P\mathbf{r}}}\,\mathbf{E}_{Q,P} \;\mid\; P \in \mathrm{SYT}\,(D)\right\} \\ &= \mathrm{span}_{\mathbf{k}}\left\{\mathbf{E}_{Q,P} \;\mid\; P \in \mathrm{SYT}\,(D)\right\} \end{aligned} \tag{284}$$

(here, we have removed the $(-1)^{w_{Q\mathbf{r},P\mathbf{r}}}$ factors inside the span, since the span of a set of vectors does not change if we scale each of these vectors by a power of $-1$).

Proposition 5.15.12 **(b)** yields $(Q\mathbf{r})\,\mathbf{r} = Q$. But $Q\mathbf{r}$ is an $n$-tableau of shape $\mathbf{r}\,(D)$ (since $Q$ is an $n$-tableau of shape $D$). Thus, (277) (applied to $\mathbf{r}\,(D)$ and $Q\mathbf{r}$ instead of $D$ and $P$) yields $\mathbf{E}_{(Q\mathbf{r})\mathbf{r}} = T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{Q\mathbf{r}}\right)\right)$. In view of $(Q\mathbf{r})\,\mathbf{r} = Q$, we can rewrite this as

$$\mathbf{E}_Q = T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{Q\mathbf{r}}\right)\right). \tag{285}$$

However, $S : \mathcal{A} \to \mathcal{A}$ is an involution (by Theorem 3.11.14 **(b)**), and thus a bijection. Hence, $S\,(\mathcal{A}) = \mathcal{A}$. Furthermore, $S$ is an algebra anti-morphism (by Theorem 3.11.14 **(a)**). Therefore,

$$S\left(\mathcal{A}\mathbf{E}_{Q\mathbf{r}}\right) = S\left(\mathbf{E}_{Q\mathbf{r}}\right)\underbrace{S\,(\mathcal{A})}_{=\mathcal{A}} = S\left(\mathbf{E}_{Q\mathbf{r}}\right)\mathcal{A}. \tag{286}$$

Furthermore, $T_{\mathrm{sign}} : \mathcal{A} \to \mathcal{A}$ is an involution (by Theorem 3.11.5 **(b)**), and thus a bijection. Hence, $T_{\mathrm{sign}}\,(\mathcal{A}) = \mathcal{A}$. Furthermore, $T_{\mathrm{sign}}$ is an algebra morphism (by Theorem 3.11.5 **(a)**). Thus,

$$\begin{aligned} T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{Q\mathbf{r}}\right)\mathcal{A}\right) &= \underbrace{T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{Q\mathbf{r}}\right)\right)}_{\substack{=\mathbf{E}_Q \\ \text{(by (285))}}}\underbrace{T_{\mathrm{sign}}\,(\mathcal{A})}_{=\mathcal{A}} \\ &= \mathbf{E}_Q\mathcal{A}. \end{aligned} \tag{287}$$

Applying the map $T_{\mathrm{sign}}$ to both sides of (286), we obtain

$$T_{\mathrm{sign}}\left(S\left(\mathcal{A}\mathbf{E}_{Q\mathbf{r}}\right)\right) = T_{\mathrm{sign}}\left(S\left(\mathbf{E}_{Q\mathbf{r}}\right)\mathcal{A}\right) = \mathbf{E}_Q\mathcal{A} \qquad \text{(by (287))}.$$

Comparing this with (284), we obtain

$$\mathbf{E}_Q\mathcal{A} = \mathrm{span}_{\mathbf{k}}\left\{\mathbf{E}_{Q,P} \;\mid\; P \in \mathrm{SYT}\,(D)\right\}.$$

Thus, (280) is proved, and the proof of Lemma 5.15.17 is complete. □

**Lemma 5.15.18.** Let $D$ be a skew Young diagram. Let $\mathrm{SYT}(D)$ be the set of all standard $n$-tableaux of shape $D$. Let $P$ and $Q$ be two $n$-tableaux of shape $D$. Then,

$$\mathbf{E}_{P,Q} \in \mathrm{span}_{\mathbf{k}} \left\{ \mathbf{E}_{U,V} \mid U, V \in \mathrm{SYT}(D) \right\}.$$

*Proof.* Let

$$\Omega := \mathrm{span}_{\mathbf{k}} \left\{ \mathbf{E}_{U,V} \mid U, V \in \mathrm{SYT}(D) \right\}. \tag{288}$$

Thus, $\Omega$ is a $\mathbf{k}$-submodule of $\mathcal{A}$ and satisfies

$$\mathbf{E}_{U,V} \in \Omega \qquad \text{for each } U, V \in \mathrm{SYT}(D). \tag{289}$$

Forget that we fixed $P$ and $Q$. We must prove that $\mathbf{E}_{P,Q} \in \mathrm{span}_{\mathbf{k}} \left\{ \mathbf{E}_{U,V} \mid U, V \in \mathrm{SYT}(D) \right\}$ for any two $n$-tableaux $P, Q$ of shape $D$. Using (289), we can restate this as follows: We must prove that $\mathbf{E}_{P,Q} \in \Omega$ for any two $n$-tableaux $P, Q$ of shape $D$. Let us rename the indices $P$ and $Q$ as $X$ and $Y$ here. Thus, we must prove that $\mathbf{E}_{X,Y} \in \Omega$ for any two $n$-tableaux $X, Y$ of shape $D$.

Let us now prove this. Let $X$ and $Y$ be two $n$-tableaux of shape $D$. Then, (270) (applied to $P = X$ and $Q = Y$) yields

$$\begin{aligned}
\mathbf{E}_{X,Y} &= \underbrace{w_{X,Y}}_{\in \mathcal{A}} \mathbf{E}_Y \\
&\in \mathcal{A}\mathbf{E}_Y \\
&= \mathrm{span}_{\mathbf{k}} \left\{ \mathbf{E}_{P,Y} \mid P \in \mathrm{SYT}(D) \right\} \qquad \text{(by (279), applied to } Q = Y\text{)} \\
&= \mathrm{span}_{\mathbf{k}} \left\{ \mathbf{E}_{U,Y} \mid U \in \mathrm{SYT}(D) \right\} \qquad (290)
\end{aligned}$$

(here, we have renamed the index $P$ as $U$). Furthermore, for each $U \in \mathrm{SYT}(D)$, we have

$$\begin{aligned}
\mathbf{E}_{U,Y} &= \mathbf{E}_U \underbrace{w_{U,Y}}_{\in \mathcal{A}} \qquad \text{(by (271), applied to } P = U \text{ and } Q = Y\text{)} \\
&\in \mathbf{E}_U \mathcal{A} \\
&= \mathrm{span}_{\mathbf{k}} \underbrace{\left\{ \mathbf{E}_{U,P} \mid P \in \mathrm{SYT}(D) \right\}}_{\substack{= \{\mathbf{E}_{U,V} \mid V \in \mathrm{SYT}(D)\} \\ \subseteq \Omega \\ \text{(since (289) shows} \\ \text{that } \mathbf{E}_{U,V} \in \Omega \text{ for each } V \in \mathrm{SYT}(D)\text{)}}} \qquad \text{(by (280), applied to } Q = U\text{)} \\
&\subseteq \mathrm{span}_{\mathbf{k}} \Omega \subseteq \Omega \qquad \text{(since } \Omega \text{ is a } \mathbf{k}\text{-module)}.
\end{aligned}$$

In other words,

$$\left\{ \mathbf{E}_{U,Y} \mid U \in \mathrm{SYT}(D) \right\} \subseteq \Omega.$$

Thus, (290) becomes

$$\mathbf{E}_{X,Y} \in \mathrm{span}_{\mathbf{k}} \underbrace{\left\{ \mathbf{E}_{U,Y} \mid U \in \mathrm{SYT}(D) \right\}}_{\subseteq \Omega} \subseteq \mathrm{span}_{\mathbf{k}} \Omega \subseteq \Omega$$

(since $\Omega$ is a $\mathbf{k}$-module).

Forget that we fixed $X$ and $Y$. We thus have shown that $\mathbf{E}_{X,Y} \in \Omega$ for any two $n$-tableaux $X, Y$ of shape $D$. As we said above, this completes the proof of Lemma 5.15.18. $\square$

**Proposition 5.15.19.** Let $\lambda$ be a partition of $n$. Let $D = Y(\lambda)$. Let $\mathrm{SYT}(D)$ be the set of all standard $n$-tableaux of shape $D$. Then:

**(a)** If $T$ is any $n$-tableau of shape $D$, then

$$\mathcal{A}\mathbf{E}_T \subseteq \mathrm{span}_{\mathbf{k}} \{\mathbf{E}_{U,V} \mid U, V \in \mathrm{SYT}(D)\}.$$

**(b)** We have

$$\mathcal{A}\mathbf{E}_\lambda \subseteq \mathrm{span}_{\mathbf{k}} \{\mathbf{E}_{U,V} \mid U, V \in \mathrm{SYT}(D)\}.$$

*Proof.* Note that $D = Y(\lambda) = Y(\lambda/\varnothing)$ is a skew Young diagram.

Let $\Omega := \mathrm{span}_{\mathbf{k}} \{\mathbf{E}_{U,V} \mid U, V \in \mathrm{SYT}(D)\}$. Then, $\Omega$ is a $\mathbf{k}$-submodule of $\mathcal{A}$. Hence, $\mathrm{span}_{\mathbf{k}} \Omega \subseteq \Omega$.

**(a)** Let $T$ be any $n$-tableau of shape $D$. Let $u \in S_n$. We shall first show that $u\mathbf{E}_T \in \Omega$.

Indeed, Proposition 5.15.15 **(d)** yields $w_{uT,T} = u$. But (270) (applied to $P = uT$ and $Q = T$) yields $\mathbf{E}_{uT,T} = w_{uT,T}\mathbf{E}_T = u\mathbf{E}_T$ (since $w_{uT,T} = u$). Hence,

$$\begin{aligned} u\mathbf{E}_T = \mathbf{E}_{uT,T} &\in \mathrm{span}_{\mathbf{k}} \{\mathbf{E}_{U,V} \mid U, V \in \mathrm{SYT}(D)\} \\ &\qquad \text{(by Lemma 5.15.18, applied to } P = uT \text{ and } Q = T\text{)} \\ &= \Omega. \end{aligned}$$

Forget that we fixed $u$. We thus have shown that $u\mathbf{E}_T \in \Omega$ for each $u \in S_n$. In other words, $\{u\mathbf{E}_T \mid u \in S_n\} \subseteq \Omega$. However, $\mathcal{A} = \mathbf{k}[S_n] = \mathrm{span}_{\mathbf{k}} \{u \mid u \in S_n\}$ (since the permutations $u \in S_n$ form a basis of $\mathbf{k}[S_n]$), and thus

$$\begin{aligned} \mathcal{A}\mathbf{E}_T &= \mathrm{span}_{\mathbf{k}} \{u \mid u \in S_n\} \cdot \mathbf{E}_T \\ &= \mathrm{span}_{\mathbf{k}} \underbrace{\{u\mathbf{E}_T \mid u \in S_n\}}_{\subseteq \Omega} \\ &\subseteq \mathrm{span}_{\mathbf{k}} \Omega \subseteq \Omega = \mathrm{span}_{\mathbf{k}} \{\mathbf{E}_{U,V} \mid U, V \in \mathrm{SYT}(D)\}. \end{aligned}$$

This proves Proposition 5.15.19 **(a)**.

**(b)** We have

$$\begin{aligned}
\mathcal{A}\mathbf{E}_\lambda &= \mathcal{A}\left(\sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \mathbf{E}_T\right) && \text{(by (241))} \\
&\subseteq \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \underbrace{\mathcal{A}\mathbf{E}_T}_{\substack{\subseteq \operatorname{span}_{\mathbf{k}}\{\mathbf{E}_{U,V} \mid U,V \in \mathrm{SYT}(D)\} \\ \text{(by Proposition 5.15.19 \textbf{(a)})}}} && \text{(by distributivity)} \\
&\subseteq \sum_{T \text{ is an } n\text{-tableau of shape } \lambda} \operatorname{span}_{\mathbf{k}}\left\{\mathbf{E}_{U,V} \mid U,V \in \mathrm{SYT}(D)\right\} \\
&\subseteq \operatorname{span}_{\mathbf{k}}\left\{\mathbf{E}_{U,V} \mid U,V \in \mathrm{SYT}(D)\right\}
\end{aligned}$$

(since $\operatorname{span}_{\mathbf{k}}\left\{\mathbf{E}_{U,V} \mid U,V \in \mathrm{SYT}(D)\right\}$ is a $\mathbf{k}$-module). This proves Proposition 5.15.19 **(b)**. □

**Exercise 5.15.5.** [2] Let $D$ be a skew Young diagram. Let $T$ be an $n$-tableau of shape $D$. Let $\mathrm{SYT}(D)$ be the set of all standard $n$-tableaux of shape $D$. Prove that the family $\left(\mathbf{E}_{T,P}\right)_{P \in \mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathbf{E}_T\mathcal{A}$.

### 5.15.4. Decomposing $\mathcal{A}$ and $\mathcal{A}\mathbf{E}_\lambda$ as left $\mathcal{A}$-modules

We now take aim at the following crucial result:

**Theorem 5.15.20.** Let $\lambda$ be a partition of $n$. Let $D = Y(\lambda)$. Let $\mathrm{SYT}(D)$ be the set of all standard $n$-tableaux of shape $D$. Assume that $h^\lambda$ is invertible in $\mathbf{k}$. Then,

$$\mathcal{A}\mathbf{E}_\lambda = \bigoplus_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T \qquad \text{(an internal direct sum)}.$$

The proof will rely on the following three lemmas:

**Lemma 5.15.21.** Let $\lambda$ be a partition of $n$. Let $S$ and $T$ be two standard $n$-tableaux of shape $Y(\lambda)$. Assume that $\overline{S} > \overline{T}$ with respect to the Young last letter order (this is the total order constructed in Proposition 5.7.7). Then, $\mathbf{E}_S\mathbf{E}_T = 0$.

*Proof.* We have $\overline{S} > \overline{T}$, so that $\overline{T} < \overline{S}$. Moreover, the $n$-tableau $T$ is standard, hence column-standard. Thus, Lemma 5.10.9 **(a)** shows that there exist no permutations $r \in \mathcal{R}(S)$ and $c \in \mathcal{C}(T)$ such that $rS = cT$. Hence, Proposition 5.11.16 **(a)** yields $\mathbf{E}_S\mathbf{E}_T = 0$, so that Lemma 5.15.21 is proven. □

**Lemma 5.15.22.** Let $\lambda$ be a partition of $n$. Let $D = Y(\lambda)$. Let $\mathrm{SYT}(D)$ be the set of all standard $n$-tableaux of shape $D$. Assume that $h^\lambda$ is invertible in $\mathbf{k}$. Then, the sum $\sum_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T$ is direct (i.e., the submodules $\mathcal{A}\mathbf{E}_T$ for all $T \in \mathrm{SYT}(D)$ are linearly disjoint).

*Proof.* Let us first recall Proposition 5.3.8, which says that there is a bijection

$$\begin{aligned}&\text{from } \{\text{row-standard } n\text{-tableaux of shape } D\}\\ &\text{to } \{n\text{-tabloids of shape } D\}\end{aligned}$$

that sends each row-standard $n$-tableau $T$ to its tabloid $\overline{T}$. Thus, in particular, this bijection is injective. In other words, if two elements $T$ and $S$ of $\{\text{row-standard } n\text{-tableaux of shape } D\}$ are distinct, then their images under this bijection are also distinct, i.e., we have $\overline{T} \neq \overline{S}$. In other words, the following holds:

> *Observation 1:* If $T$ and $S$ are two distinct row-standard $n$-tableaux of shape $D$, then $\overline{T} \neq \overline{S}$.

Next, we recall a fundamental criterion for when a sum of submodules is direct:

> *Observation 2:* Let $M$ be a $\mathbf{k}$-module, and let $(N_T)_{T \in I}$ be a finite family of $\mathbf{k}$-submodules of $M$. Then, the sum $\sum_{T \in I} N_T$ is direct if and only if the only family $(\mathbf{a}_T)_{T \in I} \in \prod_{T \in I} N_T$ that satisfies $\sum_{T \in I} \mathbf{a}_T = 0$ is the family $(0)_{T \in I}$.

This observation is just part of Proposition 4.4.2, restated with a different indexing.

Now, Lemma 5.15.22 is claiming that the sum $\sum_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T$ is direct. This is what we need to prove. By Observation 2 (applied to $M = \mathcal{A}$ and $I = \mathrm{SYT}(D)$ and $N_T = \mathcal{A}\mathbf{E}_T$), it suffices to show that the only family $(\mathbf{a}_T)_{T \in \mathrm{SYT}(D)} \in \prod_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T$ that satisfies $\sum_{T \in \mathrm{SYT}(D)} \mathbf{a}_T = 0$ is the family $(0)_{T \in \mathrm{SYT}(D)}$. So let us prove this.

Let $(\mathbf{a}_T)_{T \in \mathrm{SYT}(D)} \in \prod_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T$ be a family that satisfies $\sum_{T \in \mathrm{SYT}(D)} \mathbf{a}_T = 0$. We must prove that $(\mathbf{a}_T)_{T \in \mathrm{SYT}(D)} = (0)_{T \in \mathrm{SYT}(D)}$.

Assume the contrary. Thus, $(\mathbf{a}_T)_{T \in \mathrm{SYT}(D)} \neq (0)_{T \in \mathrm{SYT}(D)}$. Renaming the index $T$ as $S$, we can rewrite this as follows: $(\mathbf{a}_S)_{S \in \mathrm{SYT}(D)} \neq (0)_{S \in \mathrm{SYT}(D)}$. In other words, there exists some $n$-tableau $S \in \mathrm{SYT}(D)$ such that $\mathbf{a}_S \neq 0$. Among all such $n$-tableaux $S$, let us choose one for which the $n$-tabloid $\overline{S}$ is smallest

with respect to the Young last letter order (this is the total order constructed in Proposition 5.7.7). Thus, $S \in \operatorname{SYT}(D)$ is an $n$-tableau such that $\mathbf{a}_S \neq 0$, and moreover, it has the smallest $n$-tabloid $\overline{S}$ among all such $n$-tableaux. The "moreover" part of the preceding sentence can be restated as follows:

$$\begin{aligned} &\text{Each } n\text{-tableau } T \in \operatorname{SYT}(D) \text{ that satisfies } \mathbf{a}_T \neq 0 \\ &\text{must satisfy } \overline{T} \geq \overline{S}. \end{aligned} \tag{291}$$

We recall that $(\mathbf{a}_T)_{T \in \operatorname{SYT}(D)} \in \prod_{T \in \operatorname{SYT}(D)} \mathcal{A}\mathbf{E}_T$. In other words,

$$\mathbf{a}_T \in \mathcal{A}\mathbf{E}_T \qquad \text{for each } T \in \operatorname{SYT}(D). \tag{292}$$

Now, we shall prove that each $n$-tableau $T \in \operatorname{SYT}(D)$ that is distinct from $S$ must satisfy

$$\mathbf{a}_T \mathbf{E}_S = 0. \tag{293}$$

*Proof of (293):* Let $T \in \operatorname{SYT}(D)$ be an $n$-tableau that is distinct from $S$. We must prove that $\mathbf{a}_T \mathbf{E}_S = 0$.

Assume the contrary. Thus, $\mathbf{a}_T \mathbf{E}_S \neq 0$, so that $\mathbf{a}_T \neq 0$. Hence, (291) yields $\overline{T} \geq \overline{S}$. However, both $n$-tableaux $T$ and $S$ are standard (since they belong to $\operatorname{SYT}(D)$) and therefore row-standard. Since $T$ and $S$ are furthermore distinct, we thus conclude (by Observation 1) that $\overline{T} \neq \overline{S}$. Combining this with $\overline{T} \geq \overline{S}$, we obtain $\overline{T} > \overline{S}$. Moreover, $T$ and $S$ are two standard $n$-tableaux of shape $D = Y(\lambda)$. Thus, Lemma 5.15.21 (applied to $T$ and $S$ instead of $S$ and $T$) yields $\mathbf{E}_T \mathbf{E}_S = 0$ (since $\overline{T} > \overline{S}$).

However, $\mathbf{a}_T \in \mathcal{A}\mathbf{E}_T$ (by (292)). In other words, $\mathbf{a}_T = \mathbf{b}\mathbf{E}_T$ for some $\mathbf{b} \in \mathcal{A}$. Consider this $\mathbf{b}$. Then, $\underbrace{\mathbf{a}_T}_{=\mathbf{b}\mathbf{E}_T} \mathbf{E}_S = \mathbf{b} \underbrace{\mathbf{E}_T \mathbf{E}_S}_{=0} = 0$. This contradicts $\mathbf{a}_T \mathbf{E}_S \neq 0$.

This contradiction shows that our assumption was false. Hence, $\mathbf{a}_T \mathbf{E}_S = 0$ is proved. In other words, (293) is proved. □

Furthermore, we claim that

$$\mathbf{a}_S \mathbf{E}_S = h^\lambda \mathbf{a}_S. \tag{294}$$

*Proof of (294):* We know that $S$ is an $n$-tableau of shape $D = Y(\lambda)$. Thus, Theorem 5.11.3 (applied to $T = S$) yields $\mathbf{E}_S^2 = \underbrace{\frac{n!}{f^\lambda}}_{=h^\lambda} \mathbf{E}_S = h^\lambda \mathbf{E}_S$.
(by the definition of $h^\lambda$)

However, (292) (applied to $T = S$) shows that $\mathbf{a}_S \in \mathcal{A}\mathbf{E}_S$. In other words, $\mathbf{a}_S = \mathbf{b}\mathbf{E}_S$ for some $\mathbf{b} \in \mathcal{A}$. Consider this $\mathbf{b}$. Thus,

$$\underbrace{\mathbf{a}_S}_{=\mathbf{b}\mathbf{E}_S} \mathbf{E}_S = \mathbf{b} \underbrace{\mathbf{E}_S \mathbf{E}_S}_{=\mathbf{E}_S^2 = h^\lambda \mathbf{E}_S} = \mathbf{b} h^\lambda \mathbf{E}_S = h^\lambda \underbrace{\mathbf{b}\mathbf{E}_S}_{=\mathbf{a}_S} = h^\lambda \mathbf{a}_S.$$

This proves (294). □

Now, recall that $\sum_{T \in \mathrm{SYT}(D)} \mathbf{a}_T = 0$. Multiplying this equality by $\mathbf{E}_S$ from the right, we obtain $\left( \sum_{T \in \mathrm{SYT}(D)} \mathbf{a}_T \right) \mathbf{E}_S = 0 \mathbf{E}_S = 0$. Thus,

$$\begin{aligned} 0 &= \left( \sum_{T \in \mathrm{SYT}(D)} \mathbf{a}_T \right) \mathbf{E}_S = \sum_{T \in \mathrm{SYT}(D)} \mathbf{a}_T \mathbf{E}_S \\ &= \underbrace{\mathbf{a}_S \mathbf{E}_S}_{\substack{= h^\lambda \mathbf{a}_S \\ \text{(by (294))}}} + \sum_{\substack{T \in \mathrm{SYT}(D); \\ T \neq S}} \underbrace{\mathbf{a}_T \mathbf{E}_S}_{\substack{=0 \\ \text{(by (293))}}} \qquad \left( \begin{array}{c} \text{here, we have split off the} \\ \text{addend for } T = S \text{ from the sum} \end{array} \right) \\ &= h^\lambda \mathbf{a}_S + \underbrace{\sum_{\substack{T \in \mathrm{SYT}(D); \\ T \neq S}} 0}_{=0} = h^\lambda \mathbf{a}_S. \end{aligned}$$

We can divide this equality by $h^\lambda$ (since $h^\lambda$ is invertible in $\mathbf{k}$), and thus obtain $0 = \mathbf{a}_S$. But this contradicts $\mathbf{a}_S \neq 0$.

This contradiction shows that our assumption was false. Hence, we have proved that $(\mathbf{a}_T)_{T \in \mathrm{SYT}(D)} = (0)_{T \in \mathrm{SYT}(D)}$. As explained above, this completes our proof of Lemma 5.15.22. □

**Lemma 5.15.23.** Let $\lambda$ be a partition of $n$. Assume that $h^\lambda$ is invertible in $\mathbf{k}$. Let $T$ be an $n$-tableau of shape $Y(\lambda)$. Then,

$$\mathcal{A} \mathbf{E}_T \subseteq \mathcal{A} \mathbf{E}_\lambda.$$

*Proof.* Proposition 5.12.15 yields $\mathbf{E}_\lambda \mathbf{E}_T = \left(h^\lambda\right)^2 \mathbf{E}_T$. But Proposition 5.12.3 yields that $\mathbf{E}_\lambda$ belongs to the center $Z(\mathbf{k}[S_n])$ of $\mathbf{k}[S_n]$. Hence, $\mathbf{E}_\lambda \mathbf{E}_T = \mathbf{E}_T \mathbf{E}_\lambda$. Comparing this with $\mathbf{E}_\lambda \mathbf{E}_T = \left(h^\lambda\right)^2 \mathbf{E}_T$, we find $\left(h^\lambda\right)^2 \mathbf{E}_T = \mathbf{E}_T \mathbf{E}_\lambda$. We can divide this equality by $\left(h^\lambda\right)^2$ (since the number $h^\lambda$ and thus also its square $\left(h^\lambda\right)^2$ is invertible in $\mathbf{k}$), and thus obtain

$$\mathbf{E}_T = \underbrace{\frac{1}{\left(h^\lambda\right)^2} \mathbf{E}_T}_{\in \mathcal{A}} \mathbf{E}_\lambda \in \mathcal{A} \mathbf{E}_\lambda.$$

Hence, $\mathcal{A} \underbrace{\mathbf{E}_T}_{\in \mathcal{A}\mathbf{E}_\lambda} \subseteq \underbrace{\mathcal{A}\mathcal{A}}_{\subseteq \mathcal{A}} \mathbf{E}_\lambda \subseteq \mathcal{A} \mathbf{E}_\lambda$. This proves Lemma 5.15.23. □

*Proof of Theorem 5.15.20. Step 1:* Lemma 5.15.22 shows that the sum $\sum_{T \in \mathrm{SYT}(D)} \mathcal{A} \mathbf{E}_T$ is direct. Thus, the internal direct sum $\bigoplus_{T \in \mathrm{SYT}(D)} \mathcal{A} \mathbf{E}_T$ is well-defined. Let us de-

note this internal direct sum by $\Omega$. Thus, $\Omega = \bigoplus_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T$ is a $\mathbf{k}$-submodule of $\mathcal{A}$.

*Step 2:* We shall now show that

$$\Omega \subseteq \mathcal{A}\mathbf{E}_\lambda. \tag{295}$$

Indeed, we have

$$\Omega = \bigoplus_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T = \sum_{T \in \mathrm{SYT}(D)} \underbrace{\mathcal{A}\mathbf{E}_T}_{\substack{\subseteq \mathcal{A}\mathbf{E}_\lambda \\ \text{(by Lemma 5.15.23,} \\ \text{since } T \in \mathrm{SYT}(D) \text{ shows} \\ \text{that } T \text{ is a standard } n\text{-tableau} \\ \text{of shape } D = Y(\lambda))}} \subseteq \sum_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_\lambda \subseteq \mathcal{A}\mathbf{E}_\lambda$$

(since $\mathcal{A}\mathbf{E}_\lambda$ is a $\mathbf{k}$-module). This proves (295).

*Step 3:* Let us now show that

$$\mathcal{A}\mathbf{E}_\lambda \subseteq \Omega. \tag{296}$$

This, too, is quite easy: Each $U, V \in \mathrm{SYT}(D)$ satisfy

$$\begin{aligned} \mathbf{E}_{U,V} &= \underbrace{w_{U,V}}_{\in \mathcal{A}} \mathbf{E}_V \qquad \text{(by (270), applied to } P = U \text{ and } Q = V\text{)} \\ &\in \mathcal{A}\mathbf{E}_V \subseteq \bigoplus_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T \end{aligned}$$

(since $\mathcal{A}\mathbf{E}_V$ is one of the addends of the internal direct sum $\bigoplus_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T$). In other words,

$$\{\mathbf{E}_{U,V} \mid U, V \in \mathrm{SYT}(D)\} \subseteq \bigoplus_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T = \Omega.$$

But Proposition 5.15.19 **(b)** yields

$$\mathcal{A}\mathbf{E}_\lambda \subseteq \mathrm{span}_{\mathbf{k}} \underbrace{\{\mathbf{E}_{U,V} \mid U, V \in \mathrm{SYT}(D)\}}_{\subseteq \Omega} \subseteq \mathrm{span}_{\mathbf{k}} \Omega \subseteq \Omega$$

(since $\Omega$ is a $\mathbf{k}$-module). This proves (296).

*Step 4:* Combining (296) with (295), we find

$$\mathcal{A}\mathbf{E}_\lambda = \Omega = \bigoplus_{T \in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T.$$

This proves Theorem 5.15.20. □

Theorem 5.15.20 allows us to describe $\mathcal{A}\mathbf{E}_\lambda$ as a direct sum of Specht modules:

**Corollary 5.15.24.** Let $\lambda$ be a partition of $n$. Assume that $h^\lambda$ is invertible in $\mathbf{k}$. Then,

$$\mathcal{A}\mathbf{E}_\lambda \cong \left(\mathcal{S}^\lambda\right)^{f^\lambda} \qquad \text{as an } S_n\text{-representation.}$$

Here, $M^k$ just means the external direct sum $\underbrace{M \oplus M \oplus \cdots \oplus M}_{k \text{ times}}$ whenever $M$ is an $S_n$-representation and $k \in \mathbb{N}$.

*Proof.* Let $D = Y(\lambda)$. Let $\operatorname{SYT}(D)$ be the set of all standard $n$-tableaux of shape $D$.

From $D = Y(\lambda)$, we obtain $|\operatorname{SYT}(D)| = |\operatorname{SYT}(Y(\lambda))| = f^\lambda$ (since we established the equality $f^\lambda = |\operatorname{SYT}(Y(\lambda))|$ during our proof of Corollary 5.15.13). From $D = Y(\lambda)$, we also obtain $\mathcal{S}^D = \mathcal{S}^{Y(\lambda)} = \mathcal{S}^\lambda$.

Let $T \in \operatorname{SYT}(D)$. Then, $T$ is a standard $n$-tableau of shape $D = Y(\lambda)$. Hence, (228) yields that $\mathcal{S}^D \cong \mathcal{A}\mathbf{E}_T$ as left $\mathbf{k}[S_n]$-modules. Hence, $\mathcal{A}\mathbf{E}_T \cong \mathcal{S}^D = \mathcal{S}^\lambda$ as left $\mathbf{k}[S_n]$-modules.

Forget that we fixed $T$. We thus have shown that

$$\mathcal{A}\mathbf{E}_T \cong \mathcal{S}^\lambda \text{ as left } \mathbf{k}[S_n]\text{-modules} \tag{297}$$

for each $T \in \operatorname{SYT}(D)$.

But Theorem 5.15.20 yields

$$\begin{aligned}
\mathcal{A}\mathbf{E}_\lambda &= \bigoplus_{T \in \operatorname{SYT}(D)} \underbrace{\mathcal{A}\mathbf{E}_T}_{\substack{\cong \mathcal{S}^\lambda \\ \text{(by (297))}}} \\
&\cong \bigoplus_{T \in \operatorname{SYT}(D)} \mathcal{S}^\lambda \qquad \text{(an external direct sum)} \\
&= \left(\mathcal{S}^\lambda\right)^{|\operatorname{SYT}(D)|} = \left(\mathcal{S}^\lambda\right)^{f^\lambda} \qquad \left(\text{since } |\operatorname{SYT}(D)| = f^\lambda\right)
\end{aligned}$$

as left $\mathbf{k}[S_n]$-modules. Hence, $\mathcal{A}\mathbf{E}_\lambda \cong \left(\mathcal{S}^\lambda\right)^{f^\lambda}$ as $S_n$-representations (since Corollary 4.3.19 easily yields that isomorphic left $\mathbf{k}[S_n]$-modules are isomorphic $S_n$-representations). This proves Corollary 5.15.24. □

We also obtain a decomposition of the left regular representation $\mathcal{A} = \mathbf{k}[S_n]$ of $S_n$:

**Corollary 5.15.25.** For any partition $\lambda$ of $n$, let $\operatorname{SYT}(\lambda)$ be the set of all standard $n$-tableaux of shape $Y(\lambda)$. Assume that $n!$ is invertible in $\mathbf{k}$. Then,

$$\begin{aligned}
\mathcal{A} &= \bigoplus_{\lambda \text{ is a partition of } n} \bigoplus_{T \in \operatorname{SYT}(\lambda)} \mathcal{A}\mathbf{E}_T \qquad \text{(an internal direct sum)} \\
&\cong \bigoplus_{\lambda \text{ is a partition of } n} \left(\mathcal{S}^\lambda\right)^{f^\lambda} \qquad \text{as an } S_n\text{-representation.}
\end{aligned}$$

(See Corollary 5.15.24 for the meaning of $\left(\mathcal{S}^\lambda\right)^{f^\lambda}$.)

*Proof.* Let $\lambda$ be any partition of $n$. Then, the number $h^\lambda$ is invertible in $\mathbf{k}$ (as we saw in the proof of Corollary 5.12.24 **(a)**).

Let $D = Y(\lambda)$. Then, $\operatorname{SYT}(\lambda)$ is the set of all standard $n$-tableaux of shape $D$ (since $\operatorname{SYT}(\lambda)$ is the set of all standard $n$-tableaux of shape $Y(\lambda)$). In other words, $\operatorname{SYT}(\lambda)$ is what we called $\operatorname{SYT}(D)$ in Theorem 5.15.20. Thus, Theorem 5.15.20 (with $\operatorname{SYT}(D)$ renamed as $\operatorname{SYT}(\lambda)$) yields that

$$\mathcal{A}\mathbf{E}_\lambda = \bigoplus_{T \in \operatorname{SYT}(\lambda)} \mathcal{A}\mathbf{E}_T \qquad \text{(an internal direct sum)}. \tag{298}$$

Moreover, Corollary 5.15.24 shows that

$$\mathcal{A}\mathbf{E}_\lambda \cong \left(\mathcal{S}^\lambda\right)^{f^\lambda} \qquad \text{as an } S_n\text{-representation.} \tag{299}$$

Forget that we fixed $\lambda$. We thus have proved (298) and (299) for each partition $\lambda$ of $n$. Now, Corollary 5.12.25 yields

$$\begin{aligned}
\mathcal{A} &= \bigoplus_{\lambda \text{ is a partition of } n} \underbrace{\mathcal{A}\mathbf{E}_\lambda}_{\substack{= \bigoplus_{T \in \operatorname{SYT}(\lambda)} \mathcal{A}\mathbf{E}_T \\ \text{(by (298))}}} \qquad \text{(an internal direct sum)} \\
&= \bigoplus_{\lambda \text{ is a partition of } n} \ \bigoplus_{T \in \operatorname{SYT}(\lambda)} \mathcal{A}\mathbf{E}_T \qquad \text{(an internal direct sum)}.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\mathcal{A} &= \bigoplus_{\lambda \text{ is a partition of } n} \underbrace{\mathcal{A}\mathbf{E}_\lambda}_{\substack{\cong \left(\mathcal{S}^\lambda\right)^{f^\lambda} \\ \text{(by (299))}}} \\
&\cong \bigoplus_{\lambda \text{ is a partition of } n} \left(\mathcal{S}^\lambda\right)^{f^\lambda} \qquad \text{(an external direct sum)}
\end{aligned}$$

as left $\mathcal{A}$-modules, i.e., as left $\mathbf{k}[S_n]$-modules (since $\mathcal{A} = \mathbf{k}[S_n]$), i.e., as $S_n$-representations (since Corollary 4.3.19 easily yields that isomorphic left $\mathbf{k}[S_n]$-modules are isomorphic $S_n$-representations). Thus, our proof of Corollary 5.15.25 is complete. □

If $\mathbf{k}$ is a field of characteristic 0, then Corollary 5.15.25 gives a decomposition of the left regular representation $\mathcal{A} = \mathbf{k}[S_n]$ of $S_n$ into a direct sum of irreducible representations (since Corollary 5.13.3 shows that the Specht modules $\mathcal{S}^\lambda$ are irreducible).

### 5.15.5. The Young natural bases of $\mathcal{A}\mathbf{E}_\lambda$ and $\mathcal{A}$

We can now find a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$ for any partition $\lambda$ for which $h^\lambda$ is invertible:

**Theorem 5.15.26.** Let $\lambda$ be a partition of $n$. Let $D = Y(\lambda)$. Let $\operatorname{SYT}(D)$ be the set of all standard $n$-tableaux of shape $D$. Assume that $h^\lambda$ is invertible in $\mathbf{k}$. Then, the family $(\mathbf{E}_{U,V})_{U,V \in \operatorname{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$.

**Remark 5.15.27.** This is not generally true if $h^\lambda$ is not invertible in $\mathbf{k}$. In fact, in this general case, the elements $\mathbf{E}_{U,V}$ might not even lie in the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$ in the first place. Moreover, they may fail to be $\mathbf{k}$-linearly independent. For an example of the latter situation, let us consider the case $n = 3$ and $\lambda = (2,1)$. The formulas from Example 5.15.5 easily entail

$$\begin{aligned}
&\mathbf{E}_{12\backslash\backslash 3,\ 12\backslash\backslash 3} - \mathbf{E}_{12\backslash\backslash 3,\ 13\backslash\backslash 2} + \mathbf{E}_{13\backslash\backslash 2,\ 12\backslash\backslash 3} - \mathbf{E}_{13\backslash\backslash 2,\ 13\backslash\backslash 2} \\
&= -3t_{1,3} + 3t_{1,2} - 3\operatorname{cyc}_{1,2,3} + 3\operatorname{cyc}_{1,3,2} \\
&= 0 \qquad \text{if } 3 = 0 \text{ in } \mathbf{k}.
\end{aligned}$$

Thus, the family $(\mathbf{E}_{U,V})_{U,V \in \operatorname{SYT}(D)}$ is not linearly independent if $\mathbf{k}$ is a field of characteristic 3.

*Proof of Theorem 5.15.26.* We first recall a basic property of $\mathbf{k}$-modules:

> *Observation 1:* Let $M$ be a $\mathbf{k}$-module, and let $(N_T)_{T \in I}$ be a finite family of $\mathbf{k}$-submodules of $M$. Assume that the sum $\sum_{T \in I} N_T$ is direct. Assume furthermore that each $\mathbf{k}$-module $N_T$ has a basis $\left(v_{T,j}\right)_{j \in J_T}$. Then, the internal direct sum $\bigoplus_{T \in I} N_T$ has a basis $\left(v_{T,j}\right)_{T \in I, \text{ and } j \in J_T}$.

Observation 1 is a basic fact in linear algebra and easy to prove. (For example, in the case when $I = \{1,2\}$, this fact takes the following simple form: If $N_1$ and $N_2$ are two $\mathbf{k}$-submodules of $M$ such that the sum $N_1 + N_2$ is direct, and if $(v_{1,1}, v_{1,2}, \ldots, v_{1,i})$ and $\left(v_{2,1}, v_{2,2}, \ldots, v_{2,j}\right)$ are bases of $N_1$ and $N_2$, respectively, then the internal direct sum $N_1 \oplus N_2$ has a basis $\left(v_{1,1}, v_{1,2}, \ldots, v_{1,i}, v_{2,1}, v_{2,2}, \ldots, v_{2,j}\right)$. The general case is no harder to prove.)

Let us note that $D = Y(\lambda) = Y(\lambda/\varnothing)$, so that $D$ is a skew Young diagram. Now, Theorem 5.15.20 yields

$$\mathcal{A}\mathbf{E}_\lambda = \bigoplus_{T \in \operatorname{SYT}(D)} \mathcal{A}\mathbf{E}_T \qquad \text{(an internal direct sum)}.$$

Thus, the sum $\sum_{T \in \operatorname{SYT}(D)} \mathcal{A}\mathbf{E}_T$ is direct. Moreover, each $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_T$ in this direct sum has a basis $(\mathbf{E}_{P,T})_{P \in \operatorname{SYT}(D)}$ (by Proposition 5.15.9). Thus, Observation 1 (applied to $M = \mathcal{A}$ and $I = \operatorname{SYT}(D)$ and $N_T = \mathcal{A}\mathbf{E}_T$ and $\left(v_{T,j}\right)_{j \in J_T} =$

$(\mathbf{E}_{P,T})_{P\in \mathrm{SYT}(D)}$) shows that the internal direct sum $\bigoplus_{T\in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T$ has a basis

$$(\mathbf{E}_{P,T})_{T\in \mathrm{SYT}(D),\text{ and } P\in \mathrm{SYT}(D)}\,.$$

Since $\mathcal{A}\mathbf{E}_\lambda = \bigoplus_{T\in \mathrm{SYT}(D)} \mathcal{A}\mathbf{E}_T$, we can rewrite this as follows: The $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$ has a basis

$$(\mathbf{E}_{P,T})_{T\in \mathrm{SYT}(D),\text{ and } P\in \mathrm{SYT}(D)} = (\mathbf{E}_{P,T})_{P,T\in \mathrm{SYT}(D)} = (\mathbf{E}_{U,V})_{U,V\in \mathrm{SYT}(D)}$$

(here, we have renamed the indices $P$ and $T$ as $U$ and $V$). This proves Theorem 5.15.26. □

**Corollary 5.15.28.** For any partition $\lambda$ of $n$, let $\mathrm{SYT}(\lambda)$ be the set of all standard $n$-tableaux of shape $Y(\lambda)$. Assume that $n!$ is invertible in $\mathbf{k}$. Then, the family $(\mathbf{E}_{U,V})_{\lambda \text{ is a partition of } n,\text{ and } U,V\in \mathrm{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}$.

*Proof.* Corollary 5.12.25 yields

$$\mathcal{A} = \bigoplus_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda \qquad \text{(an internal direct sum)}\,.$$

Thus, the sum $\sum_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda$ is direct.

Let $\lambda$ be any partition of $n$. Then, the number $h^\lambda$ is invertible in $\mathbf{k}$ (as we saw in the proof of Corollary 5.12.24 **(a)**).

Let $D = Y(\lambda)$. Then, $\mathrm{SYT}(\lambda)$ is the set of all standard $n$-tableaux of shape $D$ (since $\mathrm{SYT}(\lambda)$ is the set of all standard $n$-tableaux of shape $Y(\lambda)$). In other words, $\mathrm{SYT}(\lambda)$ is what we called $\mathrm{SYT}(D)$ in Theorem 5.15.26. Thus, Theorem 5.15.26 (with $\mathrm{SYT}(D)$ renamed as $\mathrm{SYT}(\lambda)$) yields that the family $(\mathbf{E}_{U,V})_{U,V\in \mathrm{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$.

Forget that we fixed $\lambda$. We thus have shown that the family $(\mathbf{E}_{U,V})_{U,V\in \mathrm{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$ whenever $\lambda$ is a partition of $n$.

Now, recall Observation 1 from our above proof of Theorem 5.15.26. Let us restate this observation with the index $T$ renamed as $\lambda$:

> *Observation 1:* Let $M$ be a $\mathbf{k}$-module, and let $(N_\lambda)_{\lambda\in I}$ be a finite family of $\mathbf{k}$-submodules of $M$. Assume that the sum $\sum_{\lambda\in I} N_\lambda$ is direct. Assume furthermore that each $\mathbf{k}$-module $N_\lambda$ has a basis $\left(v_{\lambda,j}\right)_{j\in J_\lambda}$. Then, the internal direct sum $\bigoplus_{\lambda\in I} N_\lambda$ has a basis $\left(v_{\lambda,j}\right)_{\lambda\in I,\text{ and } j\in J_\lambda}$.

Now, we can apply this observation to $M = \mathcal{A}$ and $I = \{\text{partitions of } n\}$ and $N_\lambda = \mathcal{A}\mathbf{E}_\lambda$ and $\left(v_{\lambda,j}\right)_{j\in J_\lambda} = (\mathbf{E}_{U,V})_{U,V\in \mathrm{SYT}(\lambda)}$ (since the sum $\sum_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda$

is direct, and since each $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$ has a basis $(\mathbf{E}_{U,V})_{U,V\in \mathrm{SYT}(\lambda)}$). Thus, we conclude that the internal direct sum $\bigoplus_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda$ has a basis $(\mathbf{E}_{U,V})_{\lambda \text{ is a partition of } n, \text{ and } U,V\in \mathrm{SYT}(\lambda)}$. In other words, the $\mathbf{k}$-module $\mathcal{A}$ has a basis $(\mathbf{E}_{U,V})_{\lambda \text{ is a partition of } n, \text{ and } U,V\in \mathrm{SYT}(\lambda)}$ (since $\mathcal{A} = \bigoplus_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda$). This proves Corollary 5.15.28. □

The basis $(\mathbf{E}_{U,V})_{\lambda \text{ is a partition of } n, \text{ and } U,V\in \mathrm{SYT}(\lambda)}$ in Corollary 5.15.28 is called (by myself at least) the *Young natural basis* of $\mathbf{k}[S_n]$ (due to its relation to what is called the *Young natural basis* of a Specht module $\mathcal{S}^\lambda$, which is just the standard basis consisting of the standard polytabloids $\mathbf{e}_T$).

One neat application of the Young natural basis is the following purely enumerative identity:

**Exercise 5.15.6.** [3] Prove that

$$\sum_{w\in S_n \text{ is an involution}} (-1)^w = \sum_{\substack{\lambda \text{ is a partition of } n; \\ \lambda^t=\lambda}} f^\lambda.$$

[**Hint:** Consider the endomorphism $T_{\mathrm{sign}} \circ S$ of the $\mathbf{k}$-module $\mathcal{A}$. Compute the trace of this endomorphism in two ways: once using the standard basis $(w)_{w\in S_n}$ and once using the Young natural basis $(\mathbf{E}_{U,V})_{\lambda \text{ is a partition of } n, \text{ and } U,V\in \mathrm{SYT}(\lambda)}$.]

The identity in Exercise 5.15.6 has a simpler-looking (and better-known) partner, which says that

$$(\#\text{ of involutions } w \in S_n) = \sum_{\lambda \text{ is a partition of } n} f^\lambda.$$

This, too, can be proved by computing the trace of an endomorphism of $\mathcal{A}$ using two different bases. For this, see Corollary 5.21.11 below.

For our proof of Theorem 5.14.9, we need one last lemma:

**Lemma 5.15.29.** Let $\lambda$ be a partition of $n$. Assume that $h^\lambda$ is invertible in $\mathbf{k}$. Then, $\mathcal{A}\mathbf{E}_\lambda$ is a free $\mathbf{k}$-module of rank $\left(f^\lambda\right)^2$.

*Proof.* From Lemma 5.9.17, we know that $\mathcal{S}^\lambda$ is a free $\mathbf{k}$-module of rank $f^\lambda$.

Moreover, if $M$ is a free $\mathbf{k}$-module of some rank $r \in \mathbb{N}$, and if $k \in \mathbb{N}$ is an arbitrary nonnegative integer, then the $k$-th power $M^k = \underbrace{M \oplus M \oplus \cdots \oplus M}_{k \text{ times}}$ (an external direct sum) is a free $\mathbf{k}$-module of rank $kr$ (this follows easily from Lemma 5.14.4, since a finite direct sum is the same as a finite direct product).

Applying this to $M = \mathcal{S}^\lambda$ and $r = f^\lambda$ and $k = f^\lambda$, we conclude that $\left(\mathcal{S}^\lambda\right)^{f^\lambda}$ is a free $\mathbf{k}$-module of rank $f^\lambda f^\lambda$ (since $\mathcal{S}^\lambda$ is a free $\mathbf{k}$-module of rank $f^\lambda$). In other words, $\left(\mathcal{S}^\lambda\right)^{f^\lambda}$ is a free $\mathbf{k}$-module of rank $\left(f^\lambda\right)^2$ (since $f^\lambda f^\lambda = \left(f^\lambda\right)^2$).

Furthermore, Corollary 5.15.24 yields $\mathcal{A}\mathbf{E}_\lambda \cong \left(\mathcal{S}^\lambda\right)^{f^\lambda}$ as $S_n$-representations and thus as $\mathbf{k}$-modules as well. Hence, $\mathcal{A}\mathbf{E}_\lambda$ is a free $\mathbf{k}$-module of rank $\left(f^\lambda\right)^2$ (since $\left(\mathcal{S}^\lambda\right)^{f^\lambda}$ is a free $\mathbf{k}$-module of rank $\left(f^\lambda\right)^2$). This proves Lemma 5.15.29. $\square$

### 5.15.6. Proof of Theorem 5.14.9

We shall now finally prove Theorem 5.14.9:

*Proof of Theorem 5.14.9.* This is rather similar to our proof of Theorem 5.14.1. The main difference is that we no longer take sums over all partitions $\lambda$ of $n$, but instead work with our chosen partition $\lambda$. Instead of the two maps $\Phi : \mathbf{k}\left[S_n\right] \to \prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ and $\Psi : \prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right) \to \mathbf{k}\left[S_n\right]$ (the second of which might fail to exist if $n!$ is not invertible), we consider two maps $\Phi_\lambda : \mathcal{A}\mathbf{E}_\lambda \to \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ and $\Psi_\lambda : \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right) \to \mathcal{A}\mathbf{E}_\lambda$ which are (in a sense) the "$\lambda$-parts" of $\Phi$ and $\Psi$. Here is the proof in detail:

*Step 1:* In the following, we shall abbreviate the summation sign "$\sum_{T \text{ is an } n\text{-tableau of shape } \lambda}$" as "$\sum_{T \text{ of shape } \lambda}$". Using this abbreviation, we can rewrite the equality (241) as

$$\mathbf{E}_\lambda = \sum_{T \text{ of shape } \lambda} \mathbf{E}_T. \tag{300}$$

*Step 2:* Corollary 5.12.16 **(a)** shows that $\mathcal{A}\mathbf{E}_\lambda$ is a nonunital subalgebra of $\mathcal{A}$. Furthermore, Corollary 5.12.16 **(b)** shows that this nonunital algebra $\mathcal{A}\mathbf{E}_\lambda$ has a unity, namely $\dfrac{1}{\left(h^\lambda\right)^2}\mathbf{E}_\lambda$, and thus is a unital $\mathbf{k}$-algebra. Let us denote this unity by $1_\lambda$. Thus, $1_\lambda = \dfrac{1}{\left(h^\lambda\right)^2}\mathbf{E}_\lambda$ is the unity of the $\mathbf{k}$-algebra $\mathcal{A}\mathbf{E}_\lambda$.

Define a map

$$\begin{aligned} \Phi_\lambda : \mathcal{A}\mathbf{E}_\lambda &\to \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right), \\ \mathbf{a} &\mapsto \rho_\lambda\left(\mathbf{a}\right). \end{aligned}$$

In other words, $\Phi_\lambda$ is the restriction of the map $\rho_\lambda$ to the $\mathbf{k}$-submodule $\mathcal{A}\mathbf{E}_\lambda$. One of our goals is to prove that this map $\Phi_\lambda$ is a $\mathbf{k}$-algebra isomorphism.

As in Step 2 of the proof of Theorem 5.14.1, we can see that $\rho_\lambda : \mathbf{k}\left[S_n\right] \to \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ is a $\mathbf{k}$-algebra morphism.

Now, we shall show that $\Phi_\lambda$ is a $\mathbf{k}$-algebra morphism. Indeed, the map $\Phi_\lambda$ is clearly $\mathbf{k}$-linear. Furthermore, for any $\mathbf{a}, \mathbf{b} \in \mathcal{A}\mathbf{E}_\lambda$, we have

$$\begin{aligned}
\Phi_\lambda(\mathbf{ab}) &= \rho_\lambda(\mathbf{ab}) \qquad \text{(by the definition of } \Phi_\lambda\text{)} \\
&= \underbrace{\rho_\lambda(\mathbf{a})}_{\substack{=\Phi_\lambda(\mathbf{a}) \\ \text{(since the definition of } \Phi_\lambda \\ \text{yields } \Phi_\lambda(\mathbf{a})=\rho_\lambda(\mathbf{a}))}} \underbrace{\rho_\lambda(\mathbf{b})}_{\substack{=\Phi_\lambda(\mathbf{b}) \\ \text{(since the definition of } \Phi_\lambda \\ \text{yields } \Phi_\lambda(\mathbf{b})=\rho_\lambda(\mathbf{b}))}} \qquad \left(\begin{array}{c}\text{since } \rho_\lambda \text{ is a } \mathbf{k}\text{-algebra} \\ \text{morphism}\end{array}\right) \\
&= \Phi_\lambda(\mathbf{a})\,\Phi_\lambda(\mathbf{b}).
\end{aligned}$$

In other words, the map $\Phi_\lambda$ respects multiplication. Now, let us compute the endomorphism $\Phi_\lambda(1_\lambda) \in \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$. Indeed, for each $\mathbf{v} \in \mathcal{S}^\lambda$, we have

$$\underbrace{\left(\Phi_\lambda(1_\lambda)\right)}_{\substack{=\rho_\lambda(1_\lambda) \\ \text{(by the definition of } \Phi_\lambda\text{)}}}(\mathbf{v}) = \left(\rho_\lambda(1_\lambda)\right)(\mathbf{v}) = 1_\lambda \mathbf{v} \qquad \left(\begin{array}{c}\text{by the definition of} \\ \text{the curried form } \rho_\lambda\end{array}\right)$$

$$= \frac{1}{\left(h^\lambda\right)^2} \underbrace{\mathbf{E}_\lambda \mathbf{v}}_{\substack{=\left(h^\lambda\right)^2 \mathbf{v} \\ \text{(by Proposition 5.14.8)}}} \qquad \left(\text{since } 1_\lambda = \frac{1}{\left(h^\lambda\right)^2}\mathbf{E}_\lambda\right)$$

$$= \underbrace{\frac{1}{\left(h^\lambda\right)^2}\left(h^\lambda\right)^2}_{=1} \mathbf{v} = \mathbf{v}.$$

In other words, $\Phi_\lambda(1_\lambda) = \mathrm{id}$. In other words, $\Phi_\lambda$ sends the unity of the $\mathbf{k}$-algebra $\mathcal{A}\mathbf{E}_\lambda$ to the unity of the $\mathbf{k}$-algebra $\mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ (since $1_\lambda$ is the unity of the $\mathbf{k}$-algebra $\mathcal{A}\mathbf{E}_\lambda$, whereas id is the unity of the $\mathbf{k}$-algebra $\mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$). In other words, the map $\Phi_\lambda$ respects the unity. Since $\Phi_\lambda$ also respects multiplication and is $\mathbf{k}$-linear, we thus conclude that the map $\Phi_\lambda$ is a $\mathbf{k}$-algebra morphism.

We shall now focus on proving that $\Phi_\lambda$ is invertible.

*Step 3:* Let $T$ be an $n$-tableau of shape $\lambda$. Then, there exists an $\mathcal{A}$-module isomorphism $\alpha_{\lambda,T} : \mathcal{S}^\lambda \to \mathcal{A}\mathbf{E}_T$ (as we have shown in Step 3 of the proof of Theorem 5.14.1). Fix such an isomorphism $\alpha_{\lambda,T}$ once and for all.

Forget that we fixed $T$. Thus, for any $n$-tableau $T$ of shape $\lambda$, we have defined an $\mathcal{A}$-module isomorphism $\alpha_{\lambda,T} : \mathcal{S}^\lambda \to \mathcal{A}\mathbf{E}_T$. In particular, $\alpha_{\lambda,T}$ is thus a $\mathbf{k}$-module isomorphism. Therefore, if $\zeta \in \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ is a $\mathbf{k}$-module endomorphism of $\mathcal{S}^\lambda$, then $\alpha_{\lambda,T} \circ \zeta \circ \alpha_{\lambda,T}^{-1} \in \mathrm{End}_{\mathbf{k}}\left(\mathcal{A}\mathbf{E}_T\right)$ is a $\mathbf{k}$-module endomorphism of $\mathcal{A}\mathbf{E}_T$.

*Step 4:* Now, define a map

$$\Psi_\lambda : \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right) \to \mathcal{A}\mathbf{E}_\lambda,$$

which sends each endomorphism $\zeta \in \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right)$ to

$$\frac{1}{\left(h^{\lambda}\right)^{2}} \sum_{T \text{ of shape } \lambda} \underbrace{\left(\alpha_{\lambda,T} \circ \zeta \circ \alpha_{\lambda,T}^{-1}\right)\left(\mathbf{E}_{T}\right)}_{\substack{\text{(this is well-defined,} \\ \text{since } \mathbf{E}_T = 1\mathbf{E}_T \in \mathcal{A}\mathbf{E}_T \\ \text{and } \alpha_{\lambda,T}\circ\zeta\circ\alpha_{\lambda,T}^{-1} \in \mathrm{End}_{\mathbf{k}}(\mathcal{A}\mathbf{E}_T))}} .$$

This map $\Psi_{\lambda}$ is well-defined (since each $\zeta \in \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right)$ satisfies

$$\frac{1}{\left(h^{\lambda}\right)^{2}} \sum_{T \text{ of shape } \lambda} \underbrace{\left(\alpha_{\lambda,T} \circ \zeta \circ \alpha_{\lambda,T}^{-1}\right)\left(\mathbf{E}_{T}\right)}_{\substack{\in \mathcal{A}\mathbf{E}_T \\ (\text{since } \alpha_{\lambda,T}\circ\zeta\circ\alpha_{\lambda,T}^{-1} \in \mathrm{End}_{\mathbf{k}}(\mathcal{A}\mathbf{E}_T))}} \in \frac{1}{\left(h^{\lambda}\right)^{2}} \sum_{T \text{ of shape } \lambda} \underbrace{\mathcal{A}\mathbf{E}_T}_{\substack{\subseteq \mathcal{A}\mathbf{E}_\lambda \\ \text{(by Lemma 5.15.23)}}}$$

$$\subseteq \frac{1}{\left(h^{\lambda}\right)^{2}} \sum_{T \text{ of shape } \lambda} \mathcal{A}\mathbf{E}_{\lambda} \subseteq \mathcal{A}\mathbf{E}_{\lambda}$$

(since $\mathcal{A}\mathbf{E}_{\lambda}$ is a $\mathbf{k}$-module)) and $\mathbf{k}$-linear (since each $\left(\alpha_{\lambda,T} \circ \zeta \circ \alpha_{\lambda,T}^{-1}\right)\left(\mathbf{E}_{T}\right)$ depends $\mathbf{k}$-linearly on $\zeta$).

*Step 5:* We shall next show that $\Psi_{\lambda} \circ \Phi_{\lambda} = \mathrm{id}$.

Indeed, let $\mathbf{a} \in \mathcal{A}\mathbf{E}_{\lambda}$. Then, the definition of $\Phi_{\lambda}$ yields

$$\Phi_{\lambda}\left(\mathbf{a}\right) = \rho_{\lambda}\left(\mathbf{a}\right).$$

Hence,

$$\Psi_{\lambda}\left(\Phi_{\lambda}\left(\mathbf{a}\right)\right) = \Psi_{\lambda}\left(\rho_{\lambda}\left(\mathbf{a}\right)\right) = \frac{1}{\left(h^{\lambda}\right)^{2}} \sum_{T \text{ of shape } \lambda} \left(\alpha_{\lambda,T} \circ \left(\rho_{\lambda}\left(\mathbf{a}\right)\right) \circ \alpha_{\lambda,T}^{-1}\right)\left(\mathbf{E}_{T}\right)$$

(by the definition of $\Psi_{\lambda}$).

However, we can simplify the addends in this sum significantly, due to the following claim:

*Claim 1:* Let $T$ be an $n$-tableau of shape $\lambda$. Then,

$$\left(\alpha_{\lambda,T} \circ \left(\rho_{\lambda}\left(\mathbf{a}\right)\right) \circ \alpha_{\lambda,T}^{-1}\right)\left(\mathbf{E}_{T}\right) = \mathbf{a}\mathbf{E}_{T}.$$

*Proof of Claim 1.* This is the same Claim 1 that we proved in Step 5 of our proof of Theorem 5.14.1. Thus, we need not prove it again. □

Now,

$$\begin{aligned}
\left(\Psi_\lambda \circ \Phi_\lambda\right)(\mathbf{a}) &= \Psi_\lambda\left(\Phi_\lambda(\mathbf{a})\right) \\
&= \frac{1}{\left(h^\lambda\right)^2} \sum_{T \text{ of shape } \lambda} \underbrace{\left(\alpha_{\lambda,T} \circ \left(\rho_\lambda(\mathbf{a})\right) \circ \alpha_{\lambda,T}^{-1}\right)(\mathbf{E}_T)}_{\substack{=\mathbf{a}\mathbf{E}_T \\ \text{(by Claim 1)}}} \\
&\qquad\qquad \text{(by our above computation of } \Psi_\lambda\left(\Phi_\lambda(\mathbf{a})\right)) \\
&= \frac{1}{\left(h^\lambda\right)^2} \sum_{T \text{ of shape } \lambda} \mathbf{a}\mathbf{E}_T = \mathbf{a} \frac{1}{\left(h^\lambda\right)^2} \underbrace{\sum_{T \text{ of shape } \lambda} \mathbf{E}_T}_{\substack{=\mathbf{E}_\lambda \\ \text{(by (300))}}} \\
&= \mathbf{a} \underbrace{\frac{1}{\left(h^\lambda\right)^2}\mathbf{E}_\lambda}_{\substack{=1_\lambda \\ \text{(by the definition of } 1_\lambda)}} = \mathbf{a} 1_\lambda = \mathbf{a}
\end{aligned}$$

(since $\mathbf{a} \in \mathcal{A}\mathbf{E}_\lambda$, whereas $1_\lambda$ is the unity of the $\mathbf{k}$-algebra $\mathcal{A}\mathbf{E}_\lambda$).

Forget that we fixed $\mathbf{a}$. We thus have shown that $\left(\Psi_\lambda \circ \Phi_\lambda\right)(\mathbf{a}) = \mathbf{a}$ for each $\mathbf{a} \in \mathcal{A}\mathbf{E}_\lambda$. In other words, $\Psi_\lambda \circ \Phi_\lambda = \mathrm{id}$.

*Step 6:* Next we want to show that $\Phi_\lambda \circ \Psi_\lambda = \mathrm{id}$ as well. We will derive this from $\Psi_\lambda \circ \Phi_\lambda = \mathrm{id}$ using Lemma 5.14.2. In order to do so, we need to check that $\mathcal{A}\mathbf{E}_\lambda$ and $\mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ are two free $\mathbf{k}$-modules of the same rank (namely, $\left(f^\lambda\right)^2$).

Let us now do this.

Lemma 5.9.17 tells us that $\mathcal{S}^\lambda$ is a free $\mathbf{k}$-module of rank $f^\lambda$. Hence, Lemma 5.14.5 (applied to $M = \mathcal{S}^\lambda$ and $r = f^\lambda$) shows that $\mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ is a free $\mathbf{k}$-module of rank $\left(f^\lambda\right)^2$. Moreover, Lemma 5.15.29 tells us that $\mathcal{A}\mathbf{E}_\lambda$ is a free $\mathbf{k}$-module of rank $\left(f^\lambda\right)^2$.

Thus, Lemma 5.14.2 (applied to $V = \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ and $W = \mathcal{A}\mathbf{E}_\lambda$ and $r = \left(f^\lambda\right)^2$ and $f = \Psi_\lambda$ and $g = \Phi_\lambda$) shows that $\Phi_\lambda \circ \Psi_\lambda = \mathrm{id}$ (since $\Psi_\lambda \circ \Phi_\lambda = \mathrm{id}$).

*Step 7:* Combining $\Phi_\lambda \circ \Psi_\lambda = \mathrm{id}$ with $\Psi_\lambda \circ \Phi_\lambda = \mathrm{id}$, we see that the maps $\Phi_\lambda$ and $\Psi_\lambda$ are mutually inverse. Hence, the map $\Phi_\lambda$ is invertible, and thus is a $\mathbf{k}$-algebra isomorphism (since it is a $\mathbf{k}$-algebra morphism).

In other words, the restriction of $\rho_\lambda$ to $\mathcal{A}\mathbf{E}_\lambda$ is a $\mathbf{k}$-algebra isomorphism (since $\Phi_\lambda$ is the restriction of $\rho_\lambda$ to $\mathcal{A}\mathbf{E}_\lambda$). This proves Theorem 5.14.9 **(b)**.

*Step 8:* It remains to prove Theorem 5.14.9 **(a)**.

But this is easy: We know that the map $\Phi_\lambda$ is invertible, thus bijective, thus surjective. In other words, $\Phi_\lambda\left(\mathcal{A}\mathbf{E}_\lambda\right) = \mathrm{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$. However, $\Phi_\lambda$ is the restriction of $\rho_\lambda$ to $\mathcal{A}\mathbf{E}_\lambda$. Thus, the map $\Phi_\lambda$ is identical with $\rho_\lambda$ on the submodule $\mathcal{A}\mathbf{E}_\lambda$. Hence, $\Phi_\lambda\left(\mathcal{A}\mathbf{E}_\lambda\right) = \rho_\lambda\left(\underbrace{\mathcal{A}\mathbf{E}_\lambda}_{\subseteq \mathcal{A}}\right) \subseteq \rho_\lambda(\mathcal{A})$. Therefore, $\rho_\lambda(\mathcal{A}) \supseteq$

$\Phi_\lambda(\mathcal{A}\mathbf{E}_\lambda) = \mathrm{End}_{\mathbf{k}}(\mathcal{S}^\lambda)$. This entails $\rho_\lambda(\mathcal{A}) = \mathrm{End}_{\mathbf{k}}(\mathcal{S}^\lambda)$ (since we obviously have $\rho_\lambda(\mathcal{A}) \subseteq \mathrm{End}_{\mathbf{k}}(\mathcal{S}^\lambda)$). In other words, the map $\rho_\lambda$ is surjective. Thus, Theorem 5.14.9 **(a)** is proved. □

We will learn more about the Young natural basis of $\mathcal{A}$ in Section 6.4.

## 5.16. The action of conjugacy class sums on Specht modules

### 5.16.1. A general formula

Let us next discuss how a conjugacy class sum of $S_n$ (see Definition 3.9.4 **(c)**) acts on a Specht module of the form $\mathcal{S}^\lambda$. The answer turns out to be pretty nice: it just scales every vector in $\mathcal{S}^\lambda$ by a certain number. Here is the precise answer:

**Theorem 5.16.1.** Let $\lambda$ be a partition of $n$. Let $T$ be any $n$-tableau of shape $\lambda$. Let $C$ be a conjugacy class of $S_n$. Let $\mathbf{z}_C$ denote the corresponding conjugacy class sum $\sum_{c \in C} c \in \mathcal{A}$ (as in Definition 3.9.4 **(c)**). Then,

$$\mathbf{z}_C \mathbf{v} = \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} (-1)^c \right) \mathbf{v} \qquad \text{for each } \mathbf{v} \in \mathcal{S}^\lambda.$$

*Proof.* Let $\mathbf{v} \in \mathcal{S}^\lambda$.

Theorem 3.9.5 (applied to $G = S_n$) shows that the center $Z(\mathbf{k}[S_n])$ of the group algebra $\mathbf{k}[S_n]$ is the $\mathbf{k}$-linear span of the conjugacy class sums. Hence, in particular, the conjugacy class sums belong to $Z(\mathbf{k}[S_n])$. Thus, $\mathbf{z}_C$ belongs to $Z(\mathbf{k}[S_n])$ (since $\mathbf{z}_C$ is a conjugacy class sum of $S_n$).

Recall that $T$ is an $n$-tableau of shape $\lambda$, that is, an $n$-tableau of shape $Y(\lambda)$. Thus, (228) (applied to $D = Y(\lambda)$) yields $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules. In other words, $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules (since $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)}$). In other words, there exists a left $\mathcal{A}$-module isomorphism $f : \mathcal{S}^\lambda \to \mathcal{A}\mathbf{E}_T$. Consider this $f$. Thus, $f(\mathbf{v}) \in \mathcal{A}\mathbf{E}_T$. In other words,

$$f(\mathbf{v}) = \mathbf{a}\mathbf{E}_T \qquad \text{for some } \mathbf{a} \in \mathcal{A}.$$

Consider this $\mathbf{a}$.

Now, $f$ is a left $\mathcal{A}$-module morphism. Hence,

$$\begin{aligned} f\left(\mathbf{z}_C \mathbf{v}\right) &= \mathbf{z}_C \underbrace{f\left(\mathbf{v}\right)}_{=\mathbf{a}\mathbf{E}_T} = \mathbf{z}_C \mathbf{a} \underbrace{\mathbf{E}_T}_{\substack{=\nabla^-_{\operatorname{Col} T}\nabla_{\operatorname{Row} T} \\ \text{(by the definition of } \mathbf{E}_T\text{)}}} = \mathbf{z}_C \mathbf{a} \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \\ &= \underbrace{\mathbf{z}_C \left(\mathbf{a}\nabla^-_{\operatorname{Col} T}\right)}_{\substack{=\left(\mathbf{a}\nabla^-_{\operatorname{Col} T}\right)\mathbf{z}_C \\ \text{(since } \mathbf{z}_C \text{ belongs to} \\ \text{the center } Z(\mathbf{k}[S_n])\text{)}}} \nabla_{\operatorname{Row} T} = \left(\mathbf{a}\nabla^-_{\operatorname{Col} T}\right) \mathbf{z}_C \nabla_{\operatorname{Row} T} \\ &= \mathbf{a}\nabla^-_{\operatorname{Col} T} \underbrace{\mathbf{z}_C}_{\substack{=\sum_{c\in C} c \\ =\sum_{w\in C} w}} \nabla_{\operatorname{Row} T} = \mathbf{a}\nabla^-_{\operatorname{Col} T} \left(\sum_{w\in C} w\right) \nabla_{\operatorname{Row} T} \\ &= \sum_{w\in C} \mathbf{a}\nabla^-_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T}. \end{aligned} \tag{301}$$

Define a subset $U$ of $S_n$ by

$$U := \left\{cr \mid (c,r) \in \mathcal{C}(T) \times \mathcal{R}(T)\right\}.$$

Then, obviously, $cr \in U$ for each $(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T)$. Hence, the map

$$\begin{aligned} \phi : \mathcal{C}(T) \times \mathcal{R}(T) &\to U, \\ (c,r) &\mapsto cr \end{aligned}$$

is well-defined. Consider this map $\phi$. We claim the following:

*Claim 1:* The map $\phi$ is a bijection.

*Proof of Claim 1.* The image of $\phi$ is

$$\begin{aligned} \phi\left(\mathcal{C}(T) \times \mathcal{R}(T)\right) &= \left\{ \underbrace{\phi(c,r)}_{\substack{=cr \\ \text{(by the definition of } \phi\text{)}}} \mid (c,r) \in \mathcal{C}(T) \times \mathcal{R}(T) \right\} \\ &= \left\{cr \mid (c,r) \in \mathcal{C}(T) \times \mathcal{R}(T)\right\} = U. \end{aligned}$$

Hence, the map $\phi$ is surjective.

Let us next show that $\phi$ is injective. Indeed, let $(c_1,r_1)$ and $(c_2,r_2)$ be two elements of $\mathcal{C}(T) \times \mathcal{R}(T)$ satisfying $\phi(c_1,r_1) = \phi(c_2,r_2)$. We shall show that $(c_1,r_1) = (c_2,r_2)$.

The definition of $\phi$ shows that $\phi(c_1,r_1) = c_1 r_1$ and $\phi(c_2,r_2) = c_2 r_2$, so that $c_1 r_1 = \phi(c_1,r_1) = \phi(c_2,r_2) = c_2 r_2$. Hence, $c_2^{-1} \underbrace{c_1 r_1}_{=c_2 r_2} r_2^{-1} = \underbrace{c_2^{-1} c_2}_{=\mathrm{id}} \underbrace{r_2 r_2^{-1}}_{=\mathrm{id}} = \mathrm{id}$.

Moreover, $c_1 \in \mathcal{C}(T)$ (since $(c_1, r_1) \in \mathcal{C}(T) \times \mathcal{R}(T)$) and $c_2 \in \mathcal{C}(T)$ (similarly), so that $c_2^{-1} c_1 \in \mathcal{C}(T)$ (since $\mathcal{C}(T)$ is a group). Similarly, $r_1 r_2^{-1} \in \mathcal{R}(T)$. Combining these two facts, we obtain $\left(c_2^{-1} c_1, r_1 r_2^{-1}\right) \in \mathcal{C}(T) \times \mathcal{R}(T)$.

But Proposition 5.3.5 **(b)** (applied to $D = Y(\lambda)$) yields that the only pair $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$ satisfying $cr = \mathrm{id}$ is the pair $(\mathrm{id}, \mathrm{id})$. Hence, any pair $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$ satisfying $cr = \mathrm{id}$ must equal $(\mathrm{id}, \mathrm{id})$. Applying this to $(c, r) = \left(c_2^{-1} c_1, r_1 r_2^{-1}\right)$, we obtain that $\left(c_2^{-1} c_1, r_1 r_2^{-1}\right)$ equals $(\mathrm{id}, \mathrm{id})$ (since $\left(c_2^{-1} c_1, r_1 r_2^{-1}\right) \in \mathcal{C}(T) \times \mathcal{R}(T)$ and $c_2^{-1} c_1 r_1 r_2^{-1} = \mathrm{id}$). In other words, $c_2^{-1} c_1 = \mathrm{id}$ and $r_1 r_2^{-1} = \mathrm{id}$. Hence, $c_1 = c_2$ (since $c_2^{-1} c_1 = \mathrm{id}$) and $r_1 = r_2$ (since $r_1 r_2^{-1} = \mathrm{id}$). Therefore, $(c_1, r_1) = (c_2, r_2)$.

Forget that we fixed $(c_1, r_1)$ and $(c_2, r_2)$. We have thus proved that if $(c_1, r_1)$ and $(c_2, r_2)$ are two elements of $\mathcal{C}(T) \times \mathcal{R}(T)$ satisfying $\phi(c_1, r_1) = \phi(c_2, r_2)$, then $(c_1, r_1) = (c_2, r_2)$. In other words, the map $\phi$ is injective.

Now, we know that the map $\phi$ is both injective and surjective. Hence, $\phi$ is bijective, i.e., is a bijection. This proves Claim 1. □

*Claim 2:* Let $w \in S_n$ be such that $w \notin U$. Then, $\nabla^-_{\mathrm{Col}\, T} w \nabla_{\mathrm{Row}\, T} = 0$.

*Proof of Claim 2.* We have

$$w \notin U = \{cr \mid (c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)\} = \{cr \mid c \in \mathcal{C}(T) \text{ and } r \in \mathcal{R}(T)\}.$$

In other words, there exist no two permutations $c \in \mathcal{C}(T)$ and $r \in \mathcal{R}(T)$ such that $w = cr$.

Using Theorem 5.10.6 **(b)**, we can thus see that there exist two distinct integers that lie in the same row of $wT$ and simultaneously lie in the same column of $T$ (since otherwise, Theorem 5.10.6 **(b)** would yield that there exist two permutations $c \in \mathcal{C}(T)$ and $r \in \mathcal{R}(T)$ such that $w = cr$ and $\nabla^-_{\mathrm{Col}\, T} w \nabla_{\mathrm{Row}\, T} = (-1)^c \nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}$; but this would contradict the fact that there exist no two permutations $c \in \mathcal{C}(T)$ and $r \in \mathcal{R}(T)$ such that $w = cr$). Hence, Theorem 5.10.6 **(a)** yields $\nabla^-_{\mathrm{Col}\, T} w \nabla_{\mathrm{Row}\, T} = 0$. This proves Claim 2. □

*Claim 3:* Let $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$. Then, $\nabla^-_{\mathrm{Col}\, T} cr \nabla_{\mathrm{Row}\, T} = (-1)^c \mathbf{E}_T$.

*Proof of Claim 3.* From $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$, we obtain $c \in \mathcal{C}(T)$ and $r \in \mathcal{R}(T)$. Thus, Proposition 5.5.5 **(a)** (applied to $r$ instead of $w$) yields $r \nabla_{\mathrm{Row}\, T} = \nabla_{\mathrm{Row}\, T} r = \nabla_{\mathrm{Row}\, T}$. Also, Proposition 5.5.5 **(b)** (applied to $c$ instead of $w$) yields $c \nabla^-_{\mathrm{Col}\, T} = \nabla^-_{\mathrm{Col}\, T} c = (-1)^c \nabla^-_{\mathrm{Col}\, T}$ (since $c \in \mathcal{C}(T)$). Using these facts, we find

$$\underbrace{\nabla^-_{\mathrm{Col}\, T} c}_{= (-1)^c \nabla^-_{\mathrm{Col}\, T}} \underbrace{r \nabla_{\mathrm{Row}\, T}}_{= \nabla_{\mathrm{Row}\, T}} = (-1)^c \underbrace{\nabla^-_{\mathrm{Col}\, T} \nabla_{\mathrm{Row}\, T}}_{= \mathbf{E}_T \text{ (by the definition of } \mathbf{E}_T)} = (-1)^c \mathbf{E}_T.$$

This proves Claim 3. □

Now, each element $w \in C$ satisfies either $w \in U$ or $w \notin U$ (but not both). Hence,

$$\begin{aligned}
&\sum_{w \in C} \mathbf{a} \nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} \\
&= \underbrace{\sum_{\substack{w \in C; \\ w \in U}}}_{= \sum_{\substack{w \in U; \\ w \in C}}} \mathbf{a} \nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} + \sum_{\substack{w \in C; \\ w \notin U}} \mathbf{a} \underbrace{\nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T}}_{\substack{=0 \\ \text{(by Claim 2)}}} \\
&= \sum_{\substack{w \in U; \\ w \in C}} \mathbf{a} \nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} + \underbrace{\sum_{\substack{w \in C; \\ w \notin U}} \mathbf{a} 0}_{=0} = \sum_{\substack{w \in U; \\ w \in C}} \mathbf{a} \nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} \\
&= \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ \phi(c,r) \in C}} \mathbf{a} \nabla^{-}_{\operatorname{Col} T} \left( \phi\left(c,r\right) \right) \nabla_{\operatorname{Row} T} \\
&\qquad \left( \begin{array}{c} \text{here, we have substituted } \phi\left(c,r\right) \text{ for } w \text{ in the sum,} \\ \text{since the map } \phi : \mathcal{C}\left(T\right) \times \mathcal{R}\left(T\right) \to U \text{ is a bijection} \end{array} \right) \\
&= \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} \mathbf{a} \underbrace{\nabla^{-}_{\operatorname{Col} T} cr \nabla_{\operatorname{Row} T}}_{\substack{=(-1)^c \mathbf{E}_T \\ \text{(by Claim 3)}}} \qquad \left( \begin{array}{c} \text{since } \phi\left(c,r\right) = cr \\ \text{for each } \left(c,r\right) \in \mathcal{C}\left(T\right) \times \mathcal{R}\left(T\right) \\ \text{(by the definition of } \phi\text{)} \end{array} \right) \\
&= \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} \mathbf{a} \left(-1\right)^c \mathbf{E}_T = \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} \left(-1\right)^c \right) \underbrace{\mathbf{a} \mathbf{E}_T}_{\substack{=f(\mathbf{v}) \\ \text{(since } f(\mathbf{v}) = \mathbf{a}\mathbf{E}_T\text{)}}} \\
&= \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} \left(-1\right)^c \right) f\left(\mathbf{v}\right) = f\left( \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} \left(-1\right)^c \right) \mathbf{v} \right)
\end{aligned}$$

(since the map $f$ is $\mathbf{k}$-linear). Thus, (301) becomes

$$f\left(\mathbf{z}_C \mathbf{v}\right) = \sum_{w \in C} \mathbf{a} \nabla^{-}_{\operatorname{Col} T} w \nabla_{\operatorname{Row} T} = f\left( \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} \left(-1\right)^c \right) \mathbf{v} \right).$$

Since the map $f$ is injective (because $f$ is an isomorphism), we can "un-apply" $f$ to this map, and thus obtain

$$\mathbf{z}_C \mathbf{v} = \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} \left(-1\right)^c \right) \mathbf{v}.$$

This proves Theorem 5.16.1. □

### 5.16.2. The sum of all transpositions

Recall that all transpositions in $S_n$ form a single conjugacy class. Thus, the sum of all transpositions in $S_n$ is a conjugacy class sum. We can therefore apply Theorem 5.16.1 to see how this sum acts on a Specht module $\mathcal{S}^\lambda$. What is interesting is that the result can be simplified, to yield a much nicer formula for the result.

To state this formula, we introduce a notation:

**Definition 5.16.2.** Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ be a partition. Then, we set

$$\mathrm{n}(\lambda) := \sum_{i=1}^{k} \underbrace{(i-1)}_{\in \mathbb{N}} \lambda_i \in \mathbb{N}.$$

For example, if $\lambda = (4,2,2)$, then $\mathrm{n}(\lambda) = (1-1)\cdot 4 + (2-1)\cdot 2 + (3-1)\cdot 2 = 6$.

Note that many authors (e.g., Macdonald in [Macdon95, §I.1, (1.5)]) write $n(\lambda)$ for $\mathrm{n}(\lambda)$, but I am using the Roman-style n to avoid confusion with the number $n$.

The following three propositions give alternative descriptions of $\mathrm{n}(\lambda)$:

**Proposition 5.16.3.** Let $\lambda$ be a partition. Then,

$$\mathrm{n}(\lambda) = \sum_{(i,j)\in Y(\lambda)} (i-1).$$

**Proposition 5.16.4.** Let $\lambda$ be a partition. Write the conjugate $\lambda^t$ of $\lambda$ as $\lambda^t = (\lambda_1^t, \lambda_2^t, \ldots, \lambda_m^t)$. Then,

$$\mathrm{n}(\lambda) = \sum_{i=1}^{m} \binom{\lambda_i^t}{2}.$$

**Proposition 5.16.5.** Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $\lambda$. Then, $\mathrm{n}(\lambda)$ is the number of transpositions in the subgroup $\mathcal{C}(T)$ of $S_n$.

These three propositions are all elementary and quite easy to prove; detailed proofs can be found in the appendix (Section A.3).

We are now ready to state how the sum of all transpositions acts on a Specht module $\mathcal{S}^\lambda$:

**Theorem 5.16.6.** Let $\lambda$ be a partition of $n$. Let $\mathbf{t} \in \mathbf{k}[S_n]$ be the sum of all transpositions in $S_n$. Then,

$$\mathbf{tv} = \left(\mathrm{n}\left(\lambda^t\right) - \mathrm{n}\left(\lambda\right)\right)\mathbf{v} \qquad \text{for each } \mathbf{v} \in \mathcal{S}^\lambda.$$

Before we can prove this, we must establish two combinatorial lemmas:

**Lemma 5.16.7.** Let $T$ be an $n$-tableau (of any shape $D$). Let $c \in \mathcal{C}(T)$ and $r \in \mathcal{R}(T)$. Let $k \in [n]$. Then:

**(a)** If $c(k) = r(k)$, then $c(k) = r(k) = k$.

**(b)** If $(cr)(k) = k$, then $c(k) = r(k) = k$.

*Proof.* **(a)** Assume that $c(k) = r(k)$. The tableau $T$ is an $n$-tableau, and thus contains each of the integers $1, 2, \ldots, n$ exactly once. Hence, there is a unique cell $p$ of $T$ that contains the number $k$, and there is a unique cell $q$ of $T$ that contains the number $c(k) = r(k)$. Consider these two cells $p$ and $q$. Thus, $T(p) = k$ (since the cell $p$ of $T$ contains the number $k$) and similarly $T(q) = c(k) = r(k)$.

From $r \in \mathcal{R}(T)$, we see that the permutation $r$ is horizontal for $T$ (by the definition of $\mathcal{R}(T)$). Hence, the number $r(k)$ lies in the same row of $T$ as $k$ does (by the definition of "horizontal"). In other words, the cell of $T$ that contains the number $k$ lies in the same row as the cell of $T$ that contains the number $r(k)$. Since the former cell is $p$, and since the latter cell is $q$, we can rewrite this as follows: The cell $p$ lies in the same row as the cell $q$.

Similarly, from $c \in \mathcal{C}(T)$, we conclude that the cell $p$ lies in the same column as the cell $q$.

However, if two cells lie in the same row and in the same column, then they are simply the same cell. Thus, $p$ and $q$ are the same cell (since $p$ lies in the same row and in the same column as $q$). In other words, $p = q$. Hence, $T(p) = T(q) = r(k)$, so that $r(k) = T(p) = k$. Combining this with $c(k) = r(k)$, we obtain $c(k) = r(k) = k$. Thus, Lemma 5.16.7 **(a)** is proved.

**(b)** Assume that $(cr)(k) = k$. Thus, $c(r(k)) = (cr)(k) = k$. Therefore, $r(k) = c^{-1}(k)$, so that $c^{-1}(k) = r(k)$. But $c \in \mathcal{C}(T)$ and thus $c^{-1} \in \mathcal{C}(T)$ (since $\mathcal{C}(T)$ is a group). Hence, Lemma 5.16.7 **(a)** (applied to $c^{-1}$ instead of $c$) shows that $c^{-1}(k) = r(k) = k$. Hence, $k = c(k)$, so that $c(k) = k = r(k) = k$. This proves Lemma 5.16.7 **(b)**. □

**Lemma 5.16.8.** Let $C$ be the set of all transpositions in $S_n$. Let $T$ be an $n$-tableau (of any shape $D$). Let $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$ be such that $cr \in C$. Then:

**(a)** If $(-1)^c = 1$, then $c = \mathrm{id}$ and $r \in C$.

**(b)** If $(-1)^c = -1$, then $r = \mathrm{id}$ and $c \in C$.

*Proof.* From $(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T)$, we obtain $c \in \mathcal{C}(T)$ and $r \in \mathcal{R}(T)$.

We have $cr \in C$ (by assumption). In other words, $cr$ is a transposition (since $C$ is the set of all transpositions in $S_n$). In other words, $cr = t_{u,v}$ for some two distinct numbers $u, v \in [n]$. Consider these $u, v$. Define a subset $X$ of $[n]$ by $X := \{u, v\}$. This subset $X$ has size 2 (since $u$ and $v$ are distinct). Thus, its symmetric group $S_X$ consists of two permutations (namely, $\mathrm{id}_X$ and $t_{u,v}$). In other words, $|S_X| = 2$.

Define the set

$$S_{n,X} := \{w \in S_n \mid w(i) = i \text{ for all } i \notin X\}$$

(where "for all $i \notin X$" is shorthand for "for all $i \in [n] \setminus X$"). Proposition 3.7.4 **(a)** shows that this set $S_{n,X}$ is a subgroup of $S_n$, and is isomorphic to $S_X$. Hence, $|S_{n,X}| = |S_X| = 2$. But the two permutations $\mathrm{id} \in S_n$ and $t_{u,v} \in S_n$ clearly belong to $S_{n,X}$ (since $u, v \in X$), and are distinct (since $t_{u,v}(u) = v \neq u = \mathrm{id}(u)$ and thus $t_{u,v} \neq \mathrm{id}$). This easily yields that $S_{n,X} = \{\mathrm{id}, t_{u,v}\}$ [105].

Note that $(-1)^{t_{u,v}} = -1$ (since any transposition has sign $-1$).

We now claim the following:

*Claim 1:* We have $c \in S_{n,X}$.

*Proof of Claim 1.* Let $i \in [n] \setminus X$. Then, $i \notin X = \{u, v\}$, so that $t_{u,v}(i) = i$. But $cr = t_{u,v}$, so that $(cr)(i) = t_{u,v}(i) = i$. Hence, Lemma 5.16.7 **(b)** (applied to $k = i$) yields $c(i) = r(i) = i$. In particular, $c(i) = i$.

Forget that we fixed $i$. We thus have shown that $c(i) = i$ for all $i \in [n] \setminus X$. In other words, $c(i) = i$ for all $i \notin X$. In other words, $c \in S_{n,X}$ (by the definition of $S_{n,X}$). This proves Claim 1. □

Now, we can easily prove both parts of the lemma:

**(a)** Assume that $(-1)^c = 1$. Hence, $c \neq t_{u,v}$ (since $(-1)^c = 1 \neq -1 = (-1)^{t_{u,v}}$). But Claim 1 yields $c \in S_{n,X} = \{\mathrm{id}, t_{u,v}\}$. Combining this with $c \neq t_{u,v}$, we find $c \in \{\mathrm{id}, t_{u,v}\} \setminus \{t_{u,v}\} \subseteq \{\mathrm{id}\}$, so that $c = \mathrm{id}$. Thus, $cr = \mathrm{id}\, r = r$, so that $r = cr \in C$. Thus, Lemma 5.16.8 **(a)** is proved.

**(b)** Assume that $(-1)^c = -1$. Hence, $c \neq \mathrm{id}$ (since $(-1)^c = -1 \neq 1 = (-1)^{\mathrm{id}}$). But Claim 1 yields $c \in S_{n,X} = \{\mathrm{id}, t_{u,v}\}$. Combining this with $c \neq \mathrm{id}$, we find $c \in \{\mathrm{id}, t_{u,v}\} \setminus \{\mathrm{id}\} \subseteq \{t_{u,v}\}$, so that $c = t_{u,v} = cr$. Cancelling $c$ from this equality, we find $\mathrm{id} = r$ (since $c$ is invertible in $S_n$), so that $r = \mathrm{id}$. Moreover, $c = cr \in C$. Thus, Lemma 5.16.8 **(b)** is proved. □

The above lemma lets us show the following:

[105] *Proof.* We know that id and $t_{u,v}$ are two distinct elements of $S_{n,X}$ (since id and $t_{u,v}$ belong to $S_{n,X}$). Moreover, the set $S_{n,X}$ has no further elements beyond id and $t_{u,v}$ (since $S_{n,X}$ has only two elements (because $|S_{n,X}| = 2$)). Thus, the set $S_{n,X}$ has the elements id and $t_{u,v}$ and no others. In other words, $S_{n,X} = \{\mathrm{id}, t_{u,v}\}$.

**Lemma 5.16.9.** Let $C$ be the set of all transpositions in $S_n$. Let $\lambda$ be a partition of $n$. Let $T$ be any $n$-tableau of shape $\lambda$. Then,

$$\sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T); \\ cr\in C}} (-1)^c = \mathrm{n}\left(\lambda^t\right) - \mathrm{n}\left(\lambda\right).$$

*Proof.* Define the two sets

$$U := \left\{(c,r) \in \mathcal{C}\left(T\right) \times \mathcal{R}\left(T\right) \;\mid\; cr \in C \text{ and } (-1)^c = 1\right\}$$

and

$$V := \left\{(c,r) \in \mathcal{C}\left(T\right) \times \mathcal{R}\left(T\right) \;\mid\; cr \in C \text{ and } (-1)^c = -1\right\}.$$

Then, it is easy to see that $(c, \mathrm{id}) \in V$ for each $c \in C \cap \mathcal{C}\left(T\right)$ [106]. Hence, the map

$$\begin{aligned} C \cap \mathcal{C}\left(T\right) &\to V, \\ c &\mapsto (c, \mathrm{id}) \end{aligned} \tag{302}$$

is well-defined. This map is furthermore injective[107] and surjective[108]. Hence, it is bijective, i.e., a bijection. Thus,

$$|V| = \left|C \cap \mathcal{C}\left(T\right)\right|. \tag{303}$$

Furthermore, it is easy to see that $(\mathrm{id}, r) \in U$ for each $r \in C \cap \mathcal{R}\left(T\right)$ [109].

[106] *Proof.* Let $c \in C \cap \mathcal{C}\left(T\right)$. We must show that $(c, \mathrm{id}) \in V$. According to the definition of $V$, this means that we must prove that $(c, \mathrm{id}) \in \mathcal{C}\left(T\right) \times \mathcal{R}\left(T\right)$ and $c\,\mathrm{id} \in C$ and $(-1)^c = -1$.

But this is pretty easy: We have $\mathrm{id} \in \mathcal{R}\left(T\right)$ (since $\mathcal{R}\left(T\right)$ is a subgroup of $S_n$). Combining $c \in C \cap \mathcal{C}\left(T\right) \subseteq \mathcal{C}\left(T\right)$ with $\mathrm{id} \in \mathcal{R}\left(T\right)$, we obtain $(c, \mathrm{id}) \in \mathcal{C}\left(T\right) \times \mathcal{R}\left(T\right)$. Furthermore, $c\,\mathrm{id} = c \in C \cap \mathcal{C}\left(T\right) \subseteq C$. Finally, $c \in C$, so that $c$ is a transposition (since $C$ is the set of all transpositions in $S_n$). Hence, $(-1)^c = -1$ (since any transposition has sign $-1$).

Altogether, we have shown that $(c, \mathrm{id}) \in \mathcal{C}\left(T\right) \times \mathcal{R}\left(T\right)$ and $c\,\mathrm{id} \in C$ and $(-1)^c = -1$. In other words, $(c, \mathrm{id}) \in V$ (by the definition of $V$). Qed.

[107] since any $c \in C \cap \mathcal{C}\left(T\right)$ can be reconstructed from the pair $(c, \mathrm{id})$ (simply by taking the first entry of the pair)

[108] *Proof.* Let $v \in V$. We must show that there exists some $c \in C \cap \mathcal{C}\left(T\right)$ such that $v = (c, \mathrm{id})$.

We have $v \in V$. In other words, $v$ is a pair $(c, r) \in \mathcal{C}\left(T\right) \times \mathcal{R}\left(T\right)$ satisfying $cr \in C$ and $(-1)^c = -1$ (by the definition of $V$). Consider this pair $(c, r)$. Then, Lemma 5.16.8 **(b)** (applied to $D = Y\left(\lambda\right)$) yields that $r = \mathrm{id}$ and $c \in C$. Also, $c \in \mathcal{C}\left(T\right)$ (since $(c, r) \in \mathcal{C}\left(T\right) \times \mathcal{R}\left(T\right)$). Combining $c \in C$ with $c \in \mathcal{C}\left(T\right)$, we obtain $c \in C \cap \mathcal{C}\left(T\right)$. Moreover, by the definition of $(c, r)$, we have $v = (c, r) = (c, \mathrm{id})$ (since $r = \mathrm{id}$).

Forget that we defined $c$. We thus have found a $c \in C \cap \mathcal{C}\left(T\right)$ such that $v = (c, \mathrm{id})$. Thus, there exists some $c \in C \cap \mathcal{C}\left(T\right)$ such that $v = (c, \mathrm{id})$.

Forget that we fixed $v$. We thus have shown that for each $v \in V$, there exists some $c \in C \cap \mathcal{C}\left(T\right)$ such that $v = (c, \mathrm{id})$. In other words, the map (302) is surjective.

[109] *Proof.* Let $r \in C \cap \mathcal{R}\left(T\right)$. We must show that $(\mathrm{id}, r) \in U$. According to the definition of $U$, this means that we must prove that $(\mathrm{id}, r) \in \mathcal{C}\left(T\right) \times \mathcal{R}\left(T\right)$ and $\mathrm{id}\, r \in C$ and $(-1)^{\mathrm{id}} = 1$.

Hence, the map

$$\begin{aligned} C \cap \mathcal{R}(T) &\to U, \\ r &\mapsto (\mathrm{id}, r) \end{aligned} \tag{304}$$

is well-defined. This map is furthermore injective[110] and surjective[111]. Hence, it is bijective, i.e., a bijection. Thus,

$$|U| = |C \cap \mathcal{R}(T)|. \tag{305}$$

Furthermore, we have

$$|C \cap \mathcal{C}(T)| = \mathrm{n}(\lambda) \tag{306}$$

[112] and

$$|C \cap \mathcal{R}(T)| = \mathrm{n}\left(\lambda^t\right) \tag{308}$$

---

But this is pretty easy: We have $\mathrm{id} \in \mathcal{C}(T)$ (since $\mathcal{C}(T)$ is a subgroup of $S_n$). Combining this with $r \in C \cap \mathcal{R}(T) \subseteq \mathcal{R}(T)$, we obtain $(\mathrm{id}, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$. Furthermore, $\mathrm{id}\, r = r \in C \cap \mathcal{R}(T) \subseteq C$. Finally, $(-1)^{\mathrm{id}} = 1$.

Altogether, we have shown that $(\mathrm{id}, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$ and $\mathrm{id}\, r \in C$ and $(-1)^{\mathrm{id}} = 1$. In other words, $(\mathrm{id}, r) \in U$ (by the definition of $U$). Qed.

[110] since any $r \in C \cap \mathcal{R}(T)$ can be reconstructed from the pair $(\mathrm{id}, r)$ (simply by taking the first entry of the pair)

[111] *Proof.* Let $u \in U$. We must show that there exists some $r \in C \cap \mathcal{R}(T)$ such that $u = (\mathrm{id}, r)$.

We have $u \in U$. In other words, $u$ is a pair $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$ satisfying $cr \in C$ and $(-1)^c = 1$ (by the definition of $U$). Consider this pair $(c, r)$. Then, Lemma 5.16.8 **(a)** (applied to $D = Y(\lambda)$) yields that $c = \mathrm{id}$ and $r \in C$. Also, $r \in \mathcal{R}(T)$ (since $(c, r) \in \mathcal{C}(T) \times \mathcal{R}(T)$). Combining $r \in C$ with $r \in \mathcal{R}(T)$, we obtain $r \in C \cap \mathcal{R}(T)$. Moreover, by the definition of $(c, r)$, we have $u = (c, r) = (\mathrm{id}, r)$ (since $c = \mathrm{id}$).

Forget that we defined $r$. We thus have found an $r \in C \cap \mathcal{R}(T)$ such that $u = (\mathrm{id}, r)$. Thus, there exists some $r \in C \cap \mathcal{R}(T)$ such that $u = (\mathrm{id}, r)$.

Forget that we fixed $u$. We thus have shown that for each $u \in U$, there exists some $r \in C \cap \mathcal{R}(T)$ such that $u = (\mathrm{id}, r)$. In other words, the map (304) is surjective.

[112] *Proof:* Proposition 5.16.5 shows that $\mathrm{n}(\lambda)$ is the number of transpositions in $\mathcal{C}(T)$. In other words,

$$\mathrm{n}(\lambda) = (\#\text{ of all transpositions in } \mathcal{C}(T)). \tag{307}$$

Recall that $C$ is the set of all transpositions in $S_n$. In other words, $C = \{\text{all transpositions in } S_n\}$. Hence,

$$\begin{aligned} C \cap \mathcal{C}(T) &= \{\text{all transpositions in } S_n\} \cap \mathcal{C}(T) \\ &= \{\text{all transpositions in } S_n \text{ that belong to } \mathcal{C}(T)\} \\ &= \{\text{all transpositions in } \mathcal{C}(T)\} \qquad (\text{since } \mathcal{C}(T) \subseteq S_n). \end{aligned}$$

Thus,

$$\begin{aligned} |C \cap \mathcal{C}(T)| &= |\{\text{all transpositions in } \mathcal{C}(T)\}| \\ &= (\#\text{ of all transpositions in } \mathcal{C}(T)) = \mathrm{n}(\lambda) \end{aligned}$$

(by (307)).

[113].

Now, for each pair $(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T)$, the sign $(-1)^c$ is either 1 or $-1$ (since the sign of any permutation is 1 or $-1$). Hence, we can split the sum

[113] *Proof:* Consider the $n$-tableau $T\mathbf{r}$ defined in Definition 5.15.10. This is an $n$-tableau of shape $\mathbf{r}(Y(\lambda))$ (since $T$ is an $n$-tableau of shape $Y(\lambda)$). But the definition of $\lambda^t$ shows that $Y(\lambda^t) = \mathbf{r}(Y(\lambda))$. Hence, $T\mathbf{r}$ is an $n$-tableau of shape $Y(\lambda^t)$ (since $T\mathbf{r}$ is an $n$-tableau of shape $\mathbf{r}(Y(\lambda))$).

Proposition 5.16.5 shows that $\mathrm{n}(\lambda)$ is the number of transpositions in $\mathcal{C}(T)$. The same reasoning (applied to $\lambda^t$ and $T\mathbf{r}$ instead of $\lambda$ and $T$) shows that $\mathrm{n}(\lambda^t)$ is the number of transpositions in $\mathcal{C}(T\mathbf{r})$ (since $T\mathbf{r}$ is an $n$-tableau of shape $Y(\lambda^t)$). In other words,

$$\mathrm{n}(\lambda^t) = (\#\text{ of all transpositions in } \mathcal{C}(T\mathbf{r})). \tag{309}$$

But Proposition 5.15.12 **(e)** (applied to $D = Y(\lambda)$) yields $\mathcal{R}(T) = \mathcal{C}(T\mathbf{r})$. Hence, we can rewrite (309) as

$$\mathrm{n}(\lambda^t) = (\#\text{ of all transpositions in } \mathcal{R}(T)). \tag{310}$$

Recall that $C$ is the set of all transpositions in $S_n$. In other words, $C = \{\text{all transpositions in } S_n\}$. Hence,

$$\begin{aligned} C \cap \mathcal{R}(T) &= \{\text{all transpositions in } S_n\} \cap \mathcal{R}(T) \\ &= \{\text{all transpositions in } S_n \text{ that belong to } \mathcal{R}(T)\} \\ &= \{\text{all transpositions in } \mathcal{R}(T)\} \qquad (\text{since } \mathcal{R}(T) \subseteq S_n). \end{aligned}$$

Thus,

$$\begin{aligned} |C \cap \mathcal{R}(T)| &= |\{\text{all transpositions in } \mathcal{R}(T)\}| \\ &= (\#\text{ of all transpositions in } \mathcal{R}(T)) = \mathrm{n}(\lambda^t) \end{aligned}$$

(by (310)).

$\sum\limits_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ cr\in C}} (-1)^c$ as follows:

$$\begin{aligned}
\sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ cr\in C}} (-1)^c &= \sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ cr\in C;\\ (-1)^c=1}} \underbrace{(-1)^c}_{=1} + \sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ cr\in C;\\ (-1)^c=-1}} \underbrace{(-1)^c}_{=-1} \\
&= \underbrace{\sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ cr\in C;\\ (-1)^c=1}} 1}_{\substack{=\sum\limits_{(c,r)\in U}\\ \text{(by the definition of } U\text{)}}} + \underbrace{\sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ cr\in C;\\ (-1)^c=-1}} (-1)}_{\substack{=\sum\limits_{(c,r)\in V}\\ \text{(by the definition of } V\text{)}}} \\
&= \underbrace{\sum_{(c,r)\in U} 1}_{\substack{=|U|\cdot 1=|U|\\ =|C\cap\mathcal{R}(T)|\\ \text{(by (305))}}} + \underbrace{\sum_{(c,r)\in V} (-1)}_{\substack{=|V|\cdot(-1)=-|V|\\ =-|C\cap\mathcal{C}(T)|\\ \text{(by (303))}}} \\
&= |C\cap\mathcal{R}(T)| + (-|C\cap\mathcal{C}(T)|) \\
&= \underbrace{|C\cap\mathcal{R}(T)|}_{\substack{=\mathrm{n}\left(\lambda^t\right)\\ \text{(by (308))}}} - \underbrace{|C\cap\mathcal{C}(T)|}_{\substack{=\mathrm{n}(\lambda)\\ \text{(by (306))}}} = \mathrm{n}\left(\lambda^t\right) - \mathrm{n}(\lambda).
\end{aligned}$$

This proves Lemma 5.16.9. □

Now we are ready to prove Theorem 5.16.6:

*Proof of Theorem 5.16.6.* Let $C$ be the set of all transpositions in $S_n$. Then, $C$ is a conjugacy class of $S_n$ (since all transpositions in $S_n$ form a single conjugacy class). Let $\mathbf{z}_C$ denote the corresponding conjugacy class sum $\sum\limits_{c\in C} c \in \mathcal{A}$ (as in Definition 3.9.4 **(c)**).

Recall that $\mathbf{t}$ is the sum of all transpositions in $S_n$. In other words, $\mathbf{t}$ is the sum of all elements of $C$ (since the transpositions in $S_n$ are precisely the elements of $C$ (since $C$ is the set of all transpositions in $S_n$)). In other words, $\mathbf{t} = \sum\limits_{c\in C} c = \mathbf{z}_C$ (by the definition of $\mathbf{z}_C$).

Let $v \in \mathcal{S}^\lambda$. Pick any $n$-tableau $T$ of shape $Y(\lambda)$ (such a $T$ exists, since $|Y(\lambda)| = |\lambda| = n$). Thus, $T$ is an $n$-tableau of shape $\lambda$. Hence, Theorem 5.16.1 yields

$$\mathbf{z}_C\mathbf{v} = \underbrace{\left(\sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ cr\in C}} (-1)^c\right)}_{\substack{=\mathrm{n}\left(\lambda^t\right)-\mathrm{n}(\lambda)\\ \text{(by Lemma 5.16.9)}}} \mathbf{v} = \left(\mathrm{n}\left(\lambda^t\right) - \mathrm{n}(\lambda)\right)\mathbf{v}.$$

In other words, $\mathbf{tv} = \left(\mathrm{n}\left(\lambda^t\right) - \mathrm{n}\left(\lambda\right)\right)\mathbf{v}$ (since $\mathbf{t} = \mathbf{z}_C$). This proves Theorem 5.16.6. $\square$

## 5.17. Characters

The *character* of a representation is one of the key tools in classical representation theory: It is a much simpler object than the representation itself, yet encodes many of its properties and – under reasonable assumptions – uniquely determines the representation (whence probably its name). It thus can be viewed as a "fingerprint" of a representation. The general theory of characters is explained in many places (e.g., [Artin10, Chapter 10], [EGHetc11, Chapters 3 and 4], [Sagan01, Chapter 1], [Serre77], [Waerde91, Chapter 14]), and the specific case of symmetric groups can be found in many places as well (e.g., [Sagan01, §4.6–§4.10], [CSScTo10], [Prasad15, §5.4 and §5.11], [JamKer81, §2.4], [BleSch05]). We shall here focus on the basics (once again trading some depth for generality) and on connecting the characters of the Specht modules of $S_n$ to the centralized Young symmetrizers $\mathbf{E}_\lambda$.

### 5.17.1. Definition of a character

As we know from Corollary 4.3.13, the representations of a group $G$ are just the left $\mathbf{k}\left[G\right]$-modules. Thus, we define characters not just for group representations but for left $R$-modules for any $\mathbf{k}$-algebra $R$, subject to a freeness property:[114]

**Definition 5.17.1.** Let $R$ be a $\mathbf{k}$-algebra, and let $V$ be a left $R$-module. Assume that $V$ is a free $\mathbf{k}$-module of finite rank (i.e., with a finite basis).[115] For each $r \in R$, we let $\tau_r : V \to V$ be the $\mathbf{k}$-linear map $v \mapsto rv$ (which we already used in Theorem 4.3.27). Then, the *character* of the left $R$-module $V$ is defined as the map

$$\begin{aligned} \chi_V : R &\to \mathbf{k}, \\ r &\mapsto \mathrm{Tr}\left(\tau_r\right). \end{aligned}$$

(The trace here is well-defined since $V$ is a free $\mathbf{k}$-module of finite rank.)

The first, very simple, property of characters is their linearity:

[114] Recall the notion of the trace of a $\mathbf{k}$-module endomorphism; we defined this in Subsection 5.11.2.

[115] We could weaken the freeness requirement in Definition 5.17.1: Instead of requiring $V$ to be a free $\mathbf{k}$-module of finite rank, it would suffice to assume that $V$ is "finite projective" (i.e., a direct addend of a free $\mathbf{k}$-module of finite rank), since the notion of a trace is still defined in this case (see, e.g., `https://math.stackexchange.com/questions/4394614` ). But the above definition already suffices for the modules we are planning to consider.

**Proposition 5.17.2.** Let $R$ be a $\mathbf{k}$-algebra, and let $V$ be a left $R$-module. Assume that $V$ is a free $\mathbf{k}$-module of finite rank (i.e., with a finite basis). Then, the character $\chi_V$ of $V$ is a $\mathbf{k}$-linear map.

*Proof.* We must prove that $\chi_V(\lambda a + \mu b) = \lambda\chi_V(a) + \mu\chi_V(b)$ for any $\lambda, \mu \in \mathbf{k}$ and any $a, b \in R$.

Let $\lambda, \mu \in \mathbf{k}$ and $a, b \in R$. For each $r \in R$, we let $\tau_r : V \to V$ be the $\mathbf{k}$-linear map $v \mapsto rv$. The definition of the character $\chi_V$ then shows that $\chi_V(r) = \operatorname{Tr}(\tau_r)$ for each $r \in R$. Thus, $\chi_V(\lambda a + \mu b) = \operatorname{Tr}(\tau_{\lambda a+\mu b})$ and $\chi_V(a) = \operatorname{Tr}(\tau_a)$ and $\chi_V(b) = \operatorname{Tr}(\tau_b)$.

Recall two simple properties of traces: We have $\operatorname{Tr}(f+g) = \operatorname{Tr} f + \operatorname{Tr} g$ for any two $\mathbf{k}$-linear maps $f, g : V \to V$, and we have $\operatorname{Tr}(\lambda f) = \lambda \operatorname{Tr} f$ for any $\lambda \in \mathbf{k}$ and any $\mathbf{k}$-linear map $f : V \to V$. (Both of these properties can be easily derived from the analogous properties of traces of matrices.) These two properties show that the map $\operatorname{Tr} : \operatorname{End}_{\mathbf{k}} V \to \mathbf{k}$ is $\mathbf{k}$-linear.

Now, let $v \in V$. Then, the definition of $\tau_a$ yields that $\tau_a(v) = av$. Similarly, $\tau_b(v) = bv$ and $\tau_{\lambda a+\mu b}(v) = (\lambda a + \mu b)\, v$. Thus,

$$\tau_{\lambda a+\mu b}(v) = (\lambda a + \mu b)\, v = \lambda \underbrace{av}_{=\tau_a(v)} + \mu \underbrace{bv}_{=\tau_b(v)} = \lambda\tau_a(v) + \mu\tau_b(v) = (\lambda\tau_a + \mu\tau_b)(v).$$

Forget that we fixed $v$. We thus have shown that $\tau_{\lambda a+\mu b}(v) = (\lambda\tau_a + \mu\tau_b)(v)$ for each $v \in V$. In other words, $\tau_{\lambda a+\mu b} = \lambda\tau_a + \mu\tau_b$.

Now,

$$\chi_V(\lambda a + \mu b) = \operatorname{Tr}\left(\underbrace{\tau_{\lambda a+\mu b}}_{=\lambda\tau_a+\mu\tau_b}\right) = \operatorname{Tr}(\lambda\tau_a + \mu\tau_b) = \lambda \operatorname{Tr}(\tau_a) + \mu \operatorname{Tr}(\tau_b)$$

(since the map $\operatorname{Tr} : \operatorname{End}_{\mathbf{k}} V \to \mathbf{k}$ is $\mathbf{k}$-linear). In view of $\chi_V(a) = \operatorname{Tr}(\tau_a)$ and $\chi_V(b) = \operatorname{Tr}(\tau_b)$, we can rewrite this as $\chi_V(\lambda a + \mu b) = \lambda\chi_V(a) + \mu\chi_V(b)$. This completes the proof of Proposition 5.17.2. □

A slightly less obvious yet crucial property of characters is their "pseudo-commutativity":

**Proposition 5.17.3.** Let $R$ be a $\mathbf{k}$-algebra, and let $V$ be a left $R$-module. Assume that $V$ is a free $\mathbf{k}$-module of finite rank (i.e., with a finite basis). Let $a, b \in R$. Then, $\chi_V(ab) = \chi_V(ba)$.

*Proof.* For each $r \in R$, we let $\tau_r : V \to V$ be the $\mathbf{k}$-linear map $v \mapsto rv$. The definition of the character $\chi_V$ then shows that $\chi_V(r) = \operatorname{Tr}(\tau_r)$ for each $r \in R$. Thus, $\chi_V(ab) = \operatorname{Tr}(\tau_{ab})$. On the other hand, it is easy to see the following:

*Claim 1:* We have $\tau_{ab} = \tau_a \circ \tau_b$.

*Proof of Claim 1.* Let $v \in V$. Then, the definition of $\tau_b$ yields that $\tau_b(v) = bv$. Similarly, $\tau_{ab}(v) = (ab)\, v$. Furthermore, the definition of $\tau_a$ yields $\tau_a(bv) = a(bv)$. Thus,

$\tau_{ab}(v) = (ab)v = abv = a(bv) = \tau_a(bv) = \tau_a(\tau_b(v))$ (since $bv = \tau_b(v)$), so that $\tau_{ab}(v) = \tau_a(\tau_b(v)) = (\tau_a \circ \tau_b)(v)$.

Forget that we fixed $v$. We thus have shown that $\tau_{ab}(v) = (\tau_a \circ \tau_b)(v)$ for each $v \in V$. In other words, $\tau_{ab} = \tau_a \circ \tau_b$. Thus, Claim 1 is proved. □

Now, recall that $\chi_V(ab) = \operatorname{Tr}(\tau_{ab}) = \operatorname{Tr}(\tau_a \circ \tau_b)$ (since Claim 1 shows that $\tau_{ab} = \tau_a \circ \tau_b$). The same argument (with the roles of $a$ and $b$ interchanged) shows that $\chi_V(ba) = \operatorname{Tr}(\tau_b \circ \tau_a)$. However, Lemma 5.11.9 (applied to $W = V$ and $f = \tau_a$ and $g = \tau_b$) shows that $\operatorname{Tr}(\tau_a \circ \tau_b) = \operatorname{Tr}(\tau_b \circ \tau_a)$. Altogether, we thus find

$$\chi_V(ab) = \operatorname{Tr}(\tau_a \circ \tau_b) = \operatorname{Tr}(\tau_b \circ \tau_a) = \chi_V(ba) \qquad (\text{since } \chi_V(ba) = \operatorname{Tr}(\tau_b \circ \tau_a)).$$

This proves Proposition 5.17.3. □

**Warning 5.17.4.** Proposition 5.17.3 does **not** mean that products inside characters can be arbitrarily reordered. For instance, when $a, b, c \in R$, the values $\chi_V(abc)$ and $\chi_V(bac)$ are usually not equal. (But the values $\chi_V(abc)$ and $\chi_V(bca)$ are equal, since we can apply Proposition 5.17.3 to $bc$ instead of $b$.)

Our main interest lies in representations of groups, which (as we know) are particular cases of modules over rings. Thus we observe the following:

**Definition 5.17.5.** Let $G$ be a group, and let $V$ be a representation of $G$ over $\mathbf{k}$. Assume that $V$ is a free $\mathbf{k}$-module of finite rank (i.e., with a finite basis). Then, $V$ is a left $\mathbf{k}[G]$-module, and thus has a character $\chi_V$ (as defined in Definition 5.17.1). This character $\chi_V$ is $\mathbf{k}$-linear (by Proposition 5.17.2), and thus is uniquely determined by its values on the standard basis vectors $g \in G$, that is, by its restriction $\chi_V|_G$ to the group $G$. This restriction $\chi_V|_G$ is called the *group character* of the $G$-representation $V$.

We will often blur the distinction between the character $\chi_V : \mathbf{k}[G] \to \mathbf{k}$ and its restriction $\chi_V|_G : G \to \mathbf{k}$, since we can easily compute one from the other.

A simple property of group characters is that they assign equal values to conjugate elements:

**Corollary 5.17.6.** Let $G$ be a group, and let $V$ be a representation of $G$ over $\mathbf{k}$. Assume that $V$ is a free $\mathbf{k}$-module of finite rank (i.e., with a finite basis). Let $w, g \in G$. Then, the character $\chi_V$ of $V$ satisfies

$$\chi_V\left(wgw^{-1}\right) = \chi_V(g). \tag{311}$$

*Proof.* Proposition 5.17.3 (applied to $a = wg$ and $b = w^{-1}$) yields $\chi_V\left(wgw^{-1}\right) = \chi_V\left(\underbrace{w^{-1}w}_{=1}g\right) = \chi_V(g)$. This proves Corollary 5.17.6. □

### 5.17.2. Examples

As a warm-up, let us compute some group characters:

**Example 5.17.7.** Let $G$ be any group. Let $\mathbf{k}_{\mathrm{triv}}$ be the trivial $G$-representation on the $\mathbf{k}$-module $\mathbf{k}$. (This is the $\mathbf{k}$-module $\mathbf{k}$, equipped with the trivial $G$-action, which is given by $g \rightharpoonup v = v$ for each $g \in G$ and $v \in \mathbf{k}$.) Let us denote the character $\chi_{\mathbf{k}_{\mathrm{triv}}}$ of $\mathbf{k}_{\mathrm{triv}}$ by $\chi_{\mathrm{triv}}$. Then:

**(a)** We have

$$\chi_{\mathrm{triv}}(g) = 1 \qquad \text{for each } g \in G.$$

**(b)** The character $\chi_{\mathrm{triv}}$ is precisely the $\mathbf{k}$-algebra morphism $\varepsilon : \mathbf{k}[G] \to \mathbf{k}$ that sends each $g \in G$ to 1 (as suggested in Remark 3.12.4).

*Proof.* **(a)** Let $g \in G$. We must prove that $\chi_{\mathrm{triv}}(g) = 1$.

Recall that $\chi_{\mathrm{triv}} = \chi_{\mathbf{k}_{\mathrm{triv}}}$ (by the definition of $\chi_{\mathrm{triv}}$). Thus, $\chi_{\mathrm{triv}}(g) = \chi_{\mathbf{k}_{\mathrm{triv}}}(g)$.

For each $r \in \mathbf{k}[G]$, we let $\tau_r : \mathbf{k}_{\mathrm{triv}} \to \mathbf{k}_{\mathrm{triv}}$ be the $\mathbf{k}$-linear map $v \mapsto rv$. Thus, for each $v \in V$, we have

$$\begin{aligned}\tau_g(v) &= gv = g \rightharpoonup v = v \qquad \text{(since the } G\text{-action on } \mathbf{k}_{\mathrm{triv}} \text{ is trivial)} \\ &= \mathrm{id}_{\mathbf{k}}(v).\end{aligned}$$

In other words, $\tau_g = \mathrm{id}_{\mathbf{k}}$. But the matrix that represents the $\mathbf{k}$-linear map $\mathrm{id}_{\mathbf{k}} : \mathbf{k} \to \mathbf{k}$ with respect to the basis $(1)$ of $\mathbf{k}$ is the $1 \times 1$-matrix $\begin{pmatrix} 1 \end{pmatrix}$, whose trace is clearly 1. Thus, $\mathrm{Tr}(\mathrm{id}_{\mathbf{k}}) = 1$. Altogether, we now have

$$\begin{aligned}\chi_{\mathrm{triv}}(g) &= \chi_{\mathbf{k}_{\mathrm{triv}}}(g) = \mathrm{Tr}(\tau_g) \qquad \text{(by the definition of the character } \chi_{\mathbf{k}_{\mathrm{triv}}}) \\ &= \mathrm{Tr}(\mathrm{id}_{\mathbf{k}}) \qquad (\text{since } \tau_g = \mathrm{id}_{\mathbf{k}}) \\ &= 1.\end{aligned}$$

This proves Example 5.17.7 **(a)**.

**(b)** Consider the $\mathbf{k}$-algebra morphism $\varepsilon : \mathbf{k}[G] \to \mathbf{k}$ that sends each $g \in G$ to 1 (as suggested in Remark 3.12.4). We must show that $\chi_{\mathrm{triv}} = \varepsilon$. Since both maps $\chi_{\mathrm{triv}}$ and $\varepsilon$ are $\mathbf{k}$-linear, it suffices to show that these two maps $\chi_{\mathrm{triv}}$ and $\varepsilon$ agree on all standard basis vectors $g$ of $\mathbf{k}[G]$ (since these standard basis vectors span $\mathbf{k}[G]$). In other words, it suffices to show that $\chi_{\mathrm{triv}}(g) = \varepsilon(g)$ for all $g \in G$. But this is easy: For each $g \in G$, we have $\chi_{\mathrm{triv}}(g) = 1$ (by Example 5.17.7 **(a)**) and $\varepsilon(g) = 1$ (by the definition of $\varepsilon$), so that $\chi_{\mathrm{triv}}(g) = 1 = \varepsilon(g)$. Thus, Example 5.17.7 **(b)** is proved. □

For symmetric groups, there is an analogue of Example 5.17.7 **(a)**:

**Example 5.17.8.** Let $\mathbf{k}_{\mathrm{sign}}$ be the sign representation of the symmetric group $S_n$. (This is the $\mathbf{k}$-module $\mathbf{k}$, equipped with the left $S_n$-action given by $g \rightharpoonup v = (-1)^g v$ for each $g \in S_n$ and $v \in \mathbf{k}$. See Example 4.2.18 for details.) Let us denote the character $\chi_{\mathbf{k}_{\mathrm{sign}}}$ of $\mathbf{k}_{\mathrm{sign}}$ by $\chi_{\mathrm{sign}}$. Then, we have

$$\chi_{\mathrm{sign}}(g) = (-1)^g \qquad \text{for each } g \in S_n.$$

*Proof sketch.* Similar to Example 5.17.7 **(a)**, but now involving a $(-1)^g$ factor. □

Example 5.17.7 **(a)** can be generalized:

**Exercise 5.17.1.** $\boxed{1}$ Let $G$ be any group. Let $V$ be a representation of $G$ over $\mathbf{k}$. Assume that $V$ is a free $\mathbf{k}$-module of rank $d$, where $d \in \mathbb{N}$. Let $g \in G$ be an element that acts on $V$ as scaling by a scalar $\lambda \in \mathbf{k}$ (that is, satisfies $gv = \lambda v$ for all $v \in V$). Prove that $\chi_V(g) = d\lambda$.

Slightly more interesting examples of characters come from the natural and left regular representations of $S_n$:

**Example 5.17.9.** Let $\mathbf{k}^n$ be the natural representation of the symmetric group $S_n$. (See Example 4.2.17 for its definition.) Let us denote the character $\chi_{\mathbf{k}^n}$ of this representation by $\chi_{\text{nat}}$. Then, we have

$$\chi_{\text{nat}}(g) = (\#\text{ of fixed points of } g) \qquad \text{for each } g \in S_n.$$

*Proof sketch.* For each $r \in \mathbf{k}[S_n]$, we let $\tau_r : \mathbf{k}^n \to \mathbf{k}^n$ be the $\mathbf{k}$-linear map $v \mapsto rv$. Thus, the definition of the character $\chi_{\mathbf{k}^n}$ shows that $\chi_{\mathbf{k}^n}(r) = \operatorname{Tr}(\tau_r)$ for each $r \in \mathbf{k}[S_n]$.

Now, let $g \in S_n$. Then, $g \in S_n \subseteq \mathbf{k}[S_n]$, so that $\chi_{\mathbf{k}^n}(g) = \operatorname{Tr}(\tau_g)$ (since $\chi_{\mathbf{k}^n}(r) = \operatorname{Tr}(\tau_r)$ for each $r \in \mathbf{k}[S_n]$). In other words,

$$\chi_{\text{nat}}(g) = \operatorname{Tr}(\tau_g) \tag{312}$$

(since we denote the character $\chi_{\mathbf{k}^n}$ by $\chi_{\text{nat}}$).

Let $e_1, e_2, \ldots, e_n$ be the standard basis vectors of $\mathbf{k}^n$ (so that each $e_i$ is the $n$-tuple $(0, 0, \ldots, 0, 1, 0, 0, \ldots, 0)$ that has the 1 in its $i$-th position). The definition of the place permutation action of $S_n$ on $\mathbf{k}^n$ easily shows that

$$g \rightharpoonup e_j = e_{g(j)} \qquad \text{for each } j \in [n].$$

Thus, for each $j \in [n]$, we have

$$\begin{aligned} \tau_g(e_j) &= ge_j \qquad (\text{by the definition of } \tau_g) \\ &= g \rightharpoonup e_j = e_{g(j)}. \end{aligned}$$

Hence, the matrix that represents the $\mathbf{k}$-linear map $\tau_g : \mathbf{k}^n \to \mathbf{k}^n$ with respect to the basis $(e_1, e_2, \ldots, e_n)$ of $\mathbf{k}^n$ is the $n \times n$-matrix $P$ whose $j$-th column is $e_{g(j)}$ for each $j \in [n]$. Explicitly, the $(i, j)$-th entry of this matrix $P$ is

$$\begin{cases} 1, & \text{if } i = g(j); \\ 0, & \text{if } i \neq g(j) \end{cases}$$

for all $(i, j) \in [n] \times [n]$. Hence, the diagonal entries of $P$ are the entries

$$\begin{cases} 1, & \text{if } i = g(i); \\ 0, & \text{if } i \neq g(i) \end{cases} \qquad \text{for all } i \in [n].$$

Thus, the trace of $P$ is

$$\begin{aligned}
\operatorname{Tr} P &= \sum_{i\in[n]} \begin{cases} 1, & \text{if } i = g(i); \\ 0, & \text{if } i \neq g(i) \end{cases} \\
&= \sum_{\substack{i\in[n]; \\ i=g(i)}} 1 + \underbrace{\sum_{\substack{i\in[n]; \\ i\neq g(i)}} 0}_{=0} = \sum_{\substack{i\in[n]; \\ i=g(i)}} 1 = (\#\text{ of } i \in [n] \text{ such that } i = g(i)) \cdot 1 \\
&= (\#\text{ of } i \in [n] \text{ such that } i = g(i)) \\
&= (\#\text{ of fixed points of } g).
\end{aligned}$$

But we know that $P$ is the matrix that represents the $\mathbf{k}$-linear map $\tau_g : \mathbf{k}^n \to \mathbf{k}^n$ with respect to the basis $(e_1, e_2, \ldots, e_n)$ of $\mathbf{k}^n$. Hence, $\operatorname{Tr}(\tau_g) = \operatorname{Tr} P = (\#\text{ of fixed points of } g)$. Thus, we can rewrite (312) as

$$\chi_{\text{nat}}(g) = (\#\text{ of fixed points of } g).$$

Thus, Example 5.17.9 is proved. □

**Example 5.17.10.** Let $G$ be a finite group. Consider the left regular representation $\mathbf{k}[G]$ of $G$. (Its definition was given in Definition 4.3.16.) Let us denote the character $\chi_{\mathbf{k}[G]}$ of this representation by $\chi_{\text{reg}}$. Then, we have

$$\chi_{\text{reg}}(g) = \begin{cases} |G|, & \text{if } g = 1; \\ 0, & \text{if } g \neq 1 \end{cases} \qquad \text{for each } g \in G.$$

*Proof.* Let us recall Definition 5.11.6. Thus, for any two elements $h$ and $h'$ of $G$, we have

$$[h](h') = \begin{cases} 1, & \text{if } h' = h; \\ 0, & \text{if } h' \neq h. \end{cases} \tag{313}$$

For each $r \in \mathbf{k}[G]$, we let $\tau_r : \mathbf{k}[G] \to \mathbf{k}[G]$ be the $\mathbf{k}$-linear map $v \mapsto rv$. Then, the definition of the character $\chi_{\mathbf{k}[G]}$ shows that $\chi_{\mathbf{k}[G]}(r) = \operatorname{Tr}(\tau_r)$ for each $r \in \mathbf{k}[G]$.

Let $g \in G$. Then, $g \in G \subseteq \mathbf{k}[G]$, and thus $\chi_{\mathbf{k}[G]}(g) = \operatorname{Tr}(\tau_g)$ (since $\chi_{\mathbf{k}[G]}(r) = \operatorname{Tr}(\tau_r)$ for each $r \in \mathbf{k}[G]$). In other words, $\chi_{\text{reg}}(g) = \operatorname{Tr}(\tau_g)$ (since we denote the character $\chi_{\mathbf{k}[G]}$ by $\chi_{\text{reg}}$).

Now let us recall the following fact:

*Claim 1:* Let $f$ be any $\mathbf{k}$-module endomorphism of $\mathbf{k}[G]$. Then,

$$\operatorname{Tr} f = \sum_{h\in G} [h](f(h)).$$

*Proof of Claim 1.* This is just Claim 1 in the proof of Proposition 5.11.7, except that we have renamed the summation index $g$ as $h$. □

Applying Claim 1 to $f = \tau_g$, we obtain

$$\operatorname{Tr}\left(\tau_g\right) = \sum_{h \in G} [h] \left( \underbrace{\tau_g\left(h\right)}_{\substack{=gh \\ \text{(by the definition of } \tau_g\text{)}}} \right) = \sum_{h \in G} \underbrace{[h]\left(gh\right)}_{\substack{= \begin{cases} 1, & \text{if } gh = h; \\ 0, & \text{if } gh \neq h \end{cases} \\ \text{(by (313), applied to } h' = gh\text{)}}}$$

$$= \sum_{h \in G} \underbrace{\begin{cases} 1, & \text{if } gh = h; \\ 0, & \text{if } gh \neq h \end{cases}}_{\substack{= \begin{cases} 1, & \text{if } g = 1; \\ 0, & \text{if } g \neq 1 \end{cases} \\ \text{(since the equation } gh = h \text{ is equivalent to } g = 1 \text{ (because } G \text{ is a group))}}} = \sum_{h \in G} \begin{cases} 1, & \text{if } g = 1; \\ 0, & \text{if } g \neq 1 \end{cases}$$

$$= |G| \cdot \begin{cases} 1, & \text{if } g = 1; \\ 0, & \text{if } g \neq 1 \end{cases} = \begin{cases} |G| \cdot 1, & \text{if } g = 1; \\ |G| \cdot 0, & \text{if } g \neq 1 \end{cases} = \begin{cases} |G|, & \text{if } g = 1; \\ 0, & \text{if } g \neq 1 \end{cases}$$

(since $|G| \cdot 1 = |G|$ and $|G| \cdot 0 = 0$). In view of $\chi_{\text{reg}}\left(g\right) = \operatorname{Tr}\left(\tau_g\right)$, we can rewrite this as follows:

$$\chi_{\text{reg}}\left(g\right) = \begin{cases} |G|, & \text{if } g = 1; \\ 0, & \text{if } g \neq 1. \end{cases}$$

This proves Example 5.17.10. □

Soon (in Example 5.17.21), we will compute the character of the zero-sum representation $R\left(\mathbf{k}^n\right)$ of $S_n$ (defined in Subsection 4.2.6) as well. But it is a nice exercise to compute it by hand for $n = 3$:

**Example 5.17.11.** Let $n = 3$. Consider the zero-sum representation $R\left(\mathbf{k}^n\right)$ of $S_n$ (defined in Subsection 4.2.6). Let us compute its character $\chi_{R\left(\mathbf{k}^n\right)}$.

For each $r \in \mathbf{k}\left[S_n\right]$, we let $\tau_r : R\left(\mathbf{k}^n\right) \to R\left(\mathbf{k}^n\right)$ be the $\mathbf{k}$-linear map $v \mapsto rv$. Then,

$$\chi_{R\left(\mathbf{k}^n\right)}\left(r\right) = \operatorname{Tr}\left(\tau_r\right) \tag{314}$$

for each $r \in \mathbf{k}\left[S_n\right]$ (by the definition of the character).

We know that the $\mathbf{k}$-module $R\left(\mathbf{k}^n\right)$ has a basis

$$\left(e_1 - e_n,\ e_2 - e_n,\ \ldots,\ e_{n-1} - e_n\right) = \left(e_1 - e_3,\ e_2 - e_3\right) \qquad \left(\text{since } n = 3\right).$$

Let us denote this basis by $\left(f_1, f_2\right)$, so that $f_1 = e_1 - e_3$ and $f_2 = e_2 - e_3$.

The action of the simple transposition $s_1 \in S_3$ on this basis $\left(f_1, f_2\right)$ is given

by

$$\tau_{s_1}(f_1) = s_1 \underbrace{f_1}_{=e_1-e_3} = s_1(e_1-e_3) = e_2-e_3 = f_2;$$
$$\tau_{s_1}(f_2) = s_1 \underbrace{f_2}_{=e_2-e_3} = s_1(e_2-e_3) = e_1-e_3 = f_1.$$

Thus, the matrix that represents the map $\tau_{s_1} : R(\mathbf{k}^n) \to R(\mathbf{k}^n)$ with respect to this basis is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Its trace is $0+0=0$, so that we have $\operatorname{Tr}(\tau_{s_1}) = 0$. Hence, (314) (applied to $r = s_1$) yields $\chi_{R(\mathbf{k}^n)}(s_1) = \operatorname{Tr}(\tau_{s_1}) = 0$.

Let us next compute $\chi_{R(\mathbf{k}^n)}(s_2)$. Indeed, the action of the simple transposition $s_2 \in S_3$ on our basis $(f_1, f_2)$ is given by

$$\begin{aligned}\tau_{s_2}(f_1) &= s_2 f_1 = s_2(e_1-e_3) = e_1-e_2 \\ &= f_1 - f_2 \qquad \text{(by a simple calculation)};\\ \tau_{s_2}(f_2) &= s_2 f_2 = s_2(e_2-e_3) = e_3-e_2 = -f_2.\end{aligned}$$

Thus, the matrix that represents the map $\tau_{s_2} : R(\mathbf{k}^n) \to R(\mathbf{k}^n)$ with respect to this basis is $\begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}$. Its trace is $1+(-1)=0$, so that we have $\operatorname{Tr}(\tau_{s_2}) = 0$. Hence, (314) (applied to $r = s_2$) yields $\chi_{R(\mathbf{k}^n)}(s_2) = \operatorname{Tr}(\tau_{s_2}) = 0$. (Alternatively, we could have derived this from (311) and $\chi_{R(\mathbf{k}^n)}(s_1) = 0$, since $s_2 = t_{1,3} s_1 t_{1,3}^{-1}$.)

Similarly, we can compute $\chi_{R(\mathbf{k}^n)}(g)$ for the other four elements $g \in S_3$. A full list of character values is

| $g$ | $1$ | $s_1$ | $s_2$ | $t_{1,3}$ | $\mathrm{cyc}_{1,2,3}$ | $\mathrm{cyc}_{1,3,2}$ |
|---|---|---|---|---|---|---|
| $\chi_{R(\mathbf{k}^n)}(g)$ | 2 | 0 | 0 | 0 | $-1$ | $-1$ |
| $\chi_{\mathrm{triv}}(g)$ | 1 | 1 | 1 | 1 | 1 | 1 |
| $\chi_{\mathrm{sign}}(g)$ | 1 | $-1$ | $-1$ | $-1$ | 1 | 1 |
| $\chi_{\mathrm{nat}}(g)$ | 3 | 1 | 1 | 1 | 0 | 0 |
| $\chi_{\mathrm{reg}}(g)$ | 6 | 0 | 0 | 0 | 0 | 0 |

.

(For comparison, we have also listed the values of the characters $\chi_{\mathrm{triv}}, \chi_{\mathrm{sign}}, \chi_{\mathrm{nat}}, \chi_{\mathrm{reg}}$ computed in Examples 5.17.7, 5.17.8, 5.17.9, 5.17.10.)

The following exercise generalizes both Example 5.17.9 and Example 5.17.10 (why?):

**Exercise 5.17.2.** [1] Let $G$ be a group, and let $X$ be a finite left $G$-set. Let $\mathbf{k}^{(X)}$ be the permutation module corresponding to the left $G$-set $X$. (See Example 4.3.9 for its definition.) Then, $\mathbf{k}^{(X)}$ is a left $\mathbf{k}[G]$-module, i.e., a $G$-representation. Let us denote its character $\chi_{\mathbf{k}^{(X)}}$ by $\chi_X$. Prove that each $g \in G$ satisfies

$$\chi_X(g) = (\#\text{ of } x \in X \text{ that satisfy } g \rightharpoonup x = x).$$

Another example concerns an arbitrary $R$-module, but a rather specific element of $R$:

**Example 5.17.12.** Let $R$ be a $\mathbf{k}$-algebra, and let $V$ be a left $R$-module. Assume that $V$ is a free $\mathbf{k}$-module of rank $d$, where $d \in \mathbb{N}$. Then, $\chi_V(1_R) = d \cdot 1_{\mathbf{k}}$.

*Proof.* For each $r \in R$, we let $\tau_r : V \to V$ be the $\mathbf{k}$-linear map $v \mapsto rv$. The definition of the character $\chi_V$ then shows that $\chi_V(r) = \operatorname{Tr}(\tau_r)$ for each $r \in R$. Hence, $\chi_V(1_R) = \operatorname{Tr}(\tau_{1_R})$.

But each $v \in V$ satisfies

$$\begin{aligned} \tau_{1_R}(v) &= 1_R v && \text{(by the definition of } \tau_{1_R}) \\ &= v && \text{(by one of the module axioms)} \\ &= \mathrm{id}_V(v). \end{aligned}$$

Hence, $\tau_{1_R} = \mathrm{id}_V$. Thus, $\operatorname{Tr}(\tau_{1_R}) = \operatorname{Tr}(\mathrm{id}_V)$.

But $V$ is a free $\mathbf{k}$-module of rank $d$, and thus has a basis consisting of $d$ vectors. The matrix that represents the $\mathbf{k}$-linear map $\mathrm{id}_V : V \to V$ with respect to this basis is the $d \times d$-identity matrix, and thus has trace $d \cdot 1_{\mathbf{k}}$. Therefore, the map $\mathrm{id}_V$ has trace $d \cdot 1_{\mathbf{k}}$. In other words, $\operatorname{Tr}(\mathrm{id}_V) = d \cdot 1_{\mathbf{k}}$. Combining all we have found, we obtain

$$\chi_V(1_R) = \operatorname{Tr}(\tau_{1_R}) = \operatorname{Tr}(\mathrm{id}_V) = d \cdot 1_{\mathbf{k}}.$$

This proves Example 5.17.12. □

### 5.17.3. Characters and isomorphism

Isomorphic $R$-modules have the same character:

**Proposition 5.17.13.** Let $R$ be a $\mathbf{k}$-algebra, and let $V$ and $W$ be two left $R$-modules. Assume that $V$ and $W$ are free $\mathbf{k}$-modules of finite ranks. If $V \cong W$ as left $R$-modules, then $\chi_V = \chi_W$.

*Proof.* Assume that $V \cong W$ as left $R$-modules. Thus, there exists a left $R$-module isomorphism $f : V \to W$. Consider this $f$.

Let $r \in R$.

Let $\tau_r : V \to V$ be the $\mathbf{k}$-linear map $v \mapsto rv$. The definition of the character $\chi_V$ then shows that $\chi_V(r) = \operatorname{Tr}(\tau_r)$.

Let $\sigma_r : W \to W$ be the $\mathbf{k}$-linear map $v \mapsto rv$. Thus, $\sigma_r : W \to W$ is the $\mathbf{k}$-linear map that is defined just like $\tau_r$ but using $W$ instead of $V$. Hence, the definition of the character $\chi_W$ shows that $\chi_W(r) = \operatorname{Tr}(\sigma_r)$.

The map $f : V \to W$ is a left $R$-module isomorphism, thus left $R$-linear. Hence, $f(rv) = rf(v)$ for each $v \in V$. Thus, we easily obtain $f \circ \tau_r = \sigma_r \circ f$ [116]. Hence, $\sigma_r \circ f = f \circ \tau_r$.

But Lemma 5.11.9 (applied to $g = f^{-1} \circ \sigma_r$) yields $\operatorname{Tr}\left(f \circ f^{-1} \circ \sigma_r\right) = \operatorname{Tr}\left(f^{-1} \circ \sigma_r \circ f\right)$. In view of $\underbrace{f \circ f^{-1}}_{=\mathrm{id}} \circ \sigma_r = \sigma_r$ and $f^{-1} \circ \underbrace{\sigma_r \circ f}_{=f \circ \tau_r} = \underbrace{f^{-1} \circ f}_{=\mathrm{id}} \circ \tau_r = \tau_r$, we can rewrite this as $\operatorname{Tr}(\sigma_r) = \operatorname{Tr}(\tau_r)$. Hence, $\chi_W(r) = \operatorname{Tr}(\sigma_r) = \operatorname{Tr}(\tau_r) = \chi_V(r)$, so that $\chi_V(r) = \chi_W(r)$.

Forget that we fixed $r$. We thus have shown that $\chi_V(r) = \chi_W(r)$ for each $r \in R$. In other words, $\chi_V = \chi_W$. This proves Proposition 5.17.13. □

What is more surprising is that under certain "fair-weather" conditions – specifically, in the classical representation theory of a finite group – the converse of Proposition 5.17.13 holds as well:

**Theorem 5.17.14.** Assume that $\mathbf{k}$ is a field of characteristic 0. Let $G$ be a finite group. Let $V$ and $W$ be two finite-dimensional representations of $G$ over $\mathbf{k}$. Then, $V \cong W$ if and only if $\chi_V = \chi_W$.

We will not use this theorem here, so we will not prove it either (but we will sketch a proof of the particular case $G = S_n$ further below: see Proposition 5.17.42). Proofs of Theorem 5.17.14 under various additional assumptions can be easily found in the literature (e.g., in [EGHetc11, Corollary 4.2.4] when $\mathbf{k}$ is algebraically closed, or in [Webb16, Corollary 3.3.3] or [Serre77, §2.3, Corollary 2] when $\mathbf{k} = \mathbb{C}$). A fully general proof appears in [Lorenz18, Theorem 1.45].

### 5.17.4. The central element corresponding to a character

Characters of a finite group $G$ give rise to certain elements of $\mathbf{k}[G]$:

**Definition 5.17.15.** Let $G$ be a finite group. Let $V$ be a representation of $G$. Assume that $V$ is a free $\mathbf{k}$-module of finite rank. Then, we define an element $\mathbf{X}_V \in \mathbf{k}[G]$ by

$$\mathbf{X}_V := \sum_{g \in G} \chi_V(g)\, g.$$

(Here, of course, $\chi_V$ means the character of $V$.)

[116] *Proof.* Let $v \in V$. Then, $\tau_r(v) = rv$ (by the definition of $\tau_r$) and $\sigma_r(f(v)) = rf(v)$ (by the definition of $\sigma_r$). But $f(rv) = rf(v)$ (since the map $f$ is left $R$-linear). Hence,

$$(f \circ \tau_r)(v) = f\left(\underbrace{\tau_r(v)}_{=rv}\right) = f(rv) = rf(v) = \sigma_r(f(v)) = (\sigma_r \circ f)(v).$$

Forget that we fixed $v$. We thus have shown that $(f \circ \tau_r)(v) = (\sigma_r \circ f)(v)$ for each $v \in V$. In other words, $f \circ \tau_r = \sigma_r \circ f$.

The element $\mathbf{X}_V$ clearly encodes the character $\chi_V$, since each value $\chi_V(g)$ (and thus, by linearity, all values of $\chi_V$) can be reconstructed from it by taking the coefficient of $g$.

**Example 5.17.16.** Let $G$ be any finite group. Let $\mathbf{k}_{\text{triv}}$ be the trivial $G$-representation on the $\mathbf{k}$-module $\mathbf{k}$ (as in Example 5.17.7). Then, the element $\mathbf{X}_{\mathbf{k}_{\text{triv}}}$ (as defined in Definition 5.17.15) is

$$\mathbf{X}_{\mathbf{k}_{\text{triv}}} = \sum_{g \in G} \underbrace{\chi_{\mathbf{k}_{\text{triv}}}(g)}_{\substack{=1 \\ \text{(by Example 5.17.7 \textbf{(a)})}}} g = \sum_{g \in G} g.$$

This is precisely the element $\nabla_G$ from Remark 3.3.5.

**Proposition 5.17.17.** Let $G$ be a finite group. Let $V$ be a representation of $G$. Assume that $V$ is a free $\mathbf{k}$-module of finite rank. Then, $\mathbf{X}_V$ belongs to the center $Z(\mathbf{k}[G])$ of the ring $\mathbf{k}[G]$.

*Proof.* Let $\mathbf{a} := \mathbf{X}_V$. Now, let $w \in G$. But $G$ is a group. Hence, the map

$$\begin{aligned} G &\to G, \\ g &\mapsto wgw^{-1} \end{aligned}$$

is a bijection (and its inverse is the map $G \to G$, $g \mapsto w^{-1}gw$).

Now, recall that $\mathbf{a} = \mathbf{X}_V = \sum_{g \in G} \chi_V(g)\, g$ (by Definition 5.17.15). Thus,

$$\begin{aligned} w\mathbf{a}w^{-1} &= w\left(\sum_{g \in G} \chi_V(g)\, g\right) w^{-1} = \sum_{g \in G} \underbrace{\chi_V(g)}_{\substack{=\chi_V\left(wgw^{-1}\right) \\ \text{(by (311))}}} wgw^{-1} \\ &= \sum_{g \in G} \chi_V\left(wgw^{-1}\right) wgw^{-1} \\ &= \sum_{g \in G} \chi_V(g)\, g \qquad \left(\begin{array}{c} \text{here, we have substituted } g \text{ for } wgw^{-1} \text{ in} \\ \text{the sum, since the map } G \to G,\ g \mapsto wgw^{-1} \\ \text{is a bijection} \end{array}\right) \\ &= \mathbf{a}. \end{aligned}$$

Forget that we fixed $w$. We thus have shown that each $w \in G$ satisfies $w\mathbf{a}w^{-1} = \mathbf{a}$. Hence, Lemma 3.9.3 reveals that $\mathbf{a} \in Z(\mathbf{k}[G])$. In other words, $\mathbf{X}_V \in Z(\mathbf{k}[G])$ (since $\mathbf{a} = \mathbf{X}_V$). In other words, $\mathbf{X}_V$ belongs to the center $Z(\mathbf{k}[G])$ of the ring $\mathbf{k}[G]$. This proves Proposition 5.17.17. □

**Corollary 5.17.18.** Let $G$ be a finite group. Let $V$ and $W$ be two representations of $G$. Assume that $V$ and $W$ are free $\mathbf{k}$-modules of finite ranks. If $V \cong W$, then $\mathbf{X}_V = \mathbf{X}_W$.

*Proof.* Assume that $V \cong W$. Then, Proposition 5.17.13 (applied to $R = \mathbf{k}[G]$) yields $\chi_V = \chi_W$. But $\mathbf{X}_V$ is defined by the equality $\mathbf{X}_V = \sum_{g \in G} \chi_V(g)\, g$, whereas $\mathbf{X}_W$ is defined by the equality $\mathbf{X}_W = \sum_{g \in G} \chi_W(g)\, g$. The right hand sides of these two equalities are equal (since $\chi_V = \chi_W$). Hence, so are the left hand sides. In other words, $\mathbf{X}_V = \mathbf{X}_W$. This proves Corollary 5.17.18. □

The above facts are the tip of the iceberg, but they are the few things that can be said about the elements $\mathbf{X}_V$ in full generality. In the classical setup, more is known:

**Theorem 5.17.19.** Let $\mathbf{k}$ be a field of characteristic 0. Let $G$ be a finite group. Let $J_1, J_2, \ldots, J_m$ be all the irreducible representations of $G$ over $\mathbf{k}$, each listed exactly once (up to isomorphism). Then:

**(a)** The elements $\mathbf{X}_{J_1}, \mathbf{X}_{J_2}, \ldots, \mathbf{X}_{J_m}$ of $Z(\mathbf{k}[G])$ are linearly independent.

**(b)** If $\mathbf{k}$ is algebraically closed, then $(\mathbf{X}_{J_1}, \mathbf{X}_{J_2}, \ldots, \mathbf{X}_{J_m})$ is a basis of $Z(\mathbf{k}[G])$.

Theorem 5.17.19 **(b)** is proved in most texts on representation theory – e.g., in [EGHetc11, Theorem 4.2.1]. A proof of the whole Theorem 5.17.19 (and, in fact, of a more general fact) can be found in [Lorenz18, Theorem 1.44]. (Note that both of these sources work not with the elements $\mathbf{X}_{J_k}$ but rather with the characters $\chi_{J_k}$ themselves, which belong to the vector space of class functions[117] on $\mathbf{k}[G]$. This has the advantage of working in a more general setting than just group algebras. But this can easily be translated into our language when dealing with a group algebra.)

In the case when $G = S_n$, Theorem 5.17.19 **(b)** holds even without the assumption that $\mathbf{k}$ be algebraically closed; this follows from Corollary 5.17.31 **(d)** (which we will prove further below).

### 5.17.5. Quotients and direct sums

The next theorem is a highly useful tool in computing characters:

**Theorem 5.17.20.** Let $R$ be a $\mathbf{k}$-algebra. Let $M$ be a left $R$-module. Let $N$ be a left $R$-submodule of $M$. Assume that, as $\mathbf{k}$-modules, both $N$ and $M/N$ are free of finite ranks (not necessarily of the same rank). Then:

**(a)** The $\mathbf{k}$-module $M$ is free of finite rank as well.

**(b)** We have $\chi_{M/N} = \chi_M - \chi_N$. (Here, the subtraction is pointwise, as usual for $\mathbf{k}$-linear maps from $R$ to $\mathbf{k}$.)

[117] In [Lorenz18], class functions are called "trace forms".

*Proof.* For any vector $m \in M$, we let $\overline{m}$ denote its projection onto the quotient $M/N$ (that is, the coset $m + N$ of $N$ in $M$).

We assumed that both $\mathbf{k}$-modules $N$ and $M/N$ are free of finite ranks. Thus, they have bases $(n_1, n_2, \ldots, n_p)$ and $(c_1, c_2, \ldots, c_q)$, respectively. Consider these bases. For each $i \in [q]$, we pick some vector $m_i \in M$ satisfying

$$\overline{m_i} = c_i \tag{315}$$

(such a vector exists, since $c_i \in M/N$). Thus, we have defined $q$ vectors $m_1, m_2, \ldots, m_q$ in $M$.

Now, we shall show that the list $(n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q)$ is a basis of the $\mathbf{k}$-module $M$. This is a classical fact in linear algebra, but let us nevertheless outline the proof. We must prove the following two claims:

> *Claim 1:* The list $(n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q)$ is $\mathbf{k}$-linearly independent.
>
> *Claim 2:* The list $(n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q)$ spans the $\mathbf{k}$-module $M$.

*Proof of Claim 1.* Let $\nu_1, \nu_2, \ldots, \nu_p, \mu_1, \mu_2, \ldots, \mu_q \in \mathbf{k}$ be scalars satisfying

$$\nu_1 n_1 + \nu_2 n_2 + \cdots + \nu_p n_p + \mu_1 m_1 + \mu_2 m_2 + \cdots + \mu_q m_q = 0. \tag{316}$$

We shall show that all these scalars $\nu_1, \nu_2, \ldots, \nu_p, \mu_1, \mu_2, \ldots, \mu_q$ are 0.

Indeed, (316) shows that

$$\mu_1 m_1 + \mu_2 m_2 + \cdots + \mu_q m_q = -\left(\nu_1 n_1 + \nu_2 n_2 + \cdots + \nu_p n_p\right) \in N$$

(since all the vectors $n_1, n_2, \ldots, n_p$ belong to the $\mathbf{k}$-module $N$). In other words,

$$\overline{\mu_1 m_1 + \mu_2 m_2 + \cdots + \mu_q m_q} = 0$$

in $M/N$. In view of

$$\begin{aligned}\overline{\mu_1 m_1 + \mu_2 m_2 + \cdots + \mu_q m_q} &= \mu_1 \underbrace{\overline{m_1}}_{\substack{=c_1 \\ \text{(by (315))}}} + \mu_2 \underbrace{\overline{m_2}}_{\substack{=c_2 \\ \text{(by (315))}}} + \cdots + \mu_q \underbrace{\overline{m_q}}_{\substack{=c_q \\ \text{(by (315))}}} \\ &= \mu_1 c_1 + \mu_2 c_2 + \cdots + \mu_q c_q,\end{aligned}$$

we can rewrite this as $\mu_1 c_1 + \mu_2 c_2 + \cdots + \mu_q c_q = 0$. Since the list $(c_1, c_2, \ldots, c_q)$ is $\mathbf{k}$-linearly independent (because it is a basis of $M/N$), this shows that

$$\mu_1 = \mu_2 = \cdots = \mu_q = 0. \tag{317}$$

However, (316) also shows that

$$\begin{aligned}\nu_1 n_1 + \nu_2 n_2 + \cdots + \nu_p n_p &= -\left(\mu_1 m_1 + \mu_2 m_2 + \cdots + \mu_q m_q\right) \\ &= -\underbrace{\mu_1}_{\substack{=0 \\ \text{(by (317))}}} m_1 - \underbrace{\mu_2}_{\substack{=0 \\ \text{(by (317))}}} m_2 - \cdots - \underbrace{\mu_q}_{\substack{=0 \\ \text{(by (317))}}} m_q \\ &= -0m_1 - 0m_2 - \cdots - 0m_q = 0.\end{aligned}$$

Since the list $\left(n_1, n_2, \ldots, n_p\right)$ is $\mathbf{k}$-linearly independent (because it is a basis of $N$), this shows that

$$\nu_1 = \nu_2 = \cdots = \nu_p = 0.$$

Combining this with (317), we conclude that all the scalars $\nu_1, \nu_2, \ldots, \nu_p, \mu_1, \mu_2, \ldots, \mu_q$ are 0.

Forget that we fixed $\nu_1, \nu_2, \ldots, \nu_p, \mu_1, \mu_2, \ldots, \mu_q$. We thus have shown that if $\nu_1, \nu_2, \ldots, \nu_p, \mu_1, \mu_2, \ldots, \mu_q \in \mathbf{k}$ are scalars satisfying

$$\nu_1 n_1 + \nu_2 n_2 + \cdots + \nu_p n_p + \mu_1 m_1 + \mu_2 m_2 + \cdots + \mu_q m_q = 0,$$

then all the scalars $\nu_1, \nu_2, \ldots, \nu_p, \mu_1, \mu_2, \ldots, \mu_q$ are 0. In other words, the list $\left(n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q\right)$ is $\mathbf{k}$-linearly independent. This proves Claim 1. □

*Proof of Claim 2.* Let $v \in M$ be arbitrary. Then, the vector $\overline{v} \in M/N$ can be written as a $\mathbf{k}$-linear combination of $c_1, c_2, \ldots, c_q$ (since $\left(c_1, c_2, \ldots, c_q\right)$ is a basis of $M/N$). In other words, $\overline{v}$ can be written as

$$\overline{v} = \alpha_1 c_1 + \alpha_2 c_2 + \cdots + \alpha_q c_q \tag{318}$$

for some scalars $\alpha_1, \alpha_2, \ldots, \alpha_q \in \mathbf{k}$. Consider these scalars $\alpha_1, \alpha_2, \ldots, \alpha_q$.

Let $w := \alpha_1 m_1 + \alpha_2 m_2 + \cdots + \alpha_q m_q$. Thus,

$$\begin{aligned}\overline{w} &= \overline{\alpha_1 m_1 + \alpha_2 m_2 + \cdots + \alpha_q m_q} = \alpha_1 \underbrace{\overline{m_1}}_{\substack{=c_1 \\ \text{(by (315))}}} + \alpha_2 \underbrace{\overline{m_2}}_{\substack{=c_2 \\ \text{(by (315))}}} + \cdots + \alpha_q \underbrace{\overline{m_q}}_{\substack{=c_q \\ \text{(by (315))}}} \\ &= \alpha_1 c_1 + \alpha_2 c_2 + \cdots + \alpha_q c_q = \overline{v} \qquad \left(\text{by (318)}\right).\end{aligned}$$

Thus, $\overline{v} = \overline{w}$. In other words, $v - w \in N$. Hence, the vector $v - w$ can be written as a $\mathbf{k}$-linear combination of $n_1, n_2, \ldots, n_p$ (since $\left(n_1, n_2, \ldots, n_p\right)$ is a basis of $N$). In other words, $v - w$ can be written as

$$v - w = \beta_1 n_1 + \beta_2 n_2 + \cdots + \beta_p n_p \tag{319}$$

for some scalars $\beta_1, \beta_2, \ldots, \beta_p \in \mathbf{k}$. Consider these scalars $\beta_1, \beta_2, \ldots, \beta_p$.

From (319), we obtain

$$\begin{aligned}v &= \beta_1 n_1 + \beta_2 n_2 + \cdots + \beta_p n_p + \underbrace{w}_{=\alpha_1 m_1 + \alpha_2 m_2 + \cdots + \alpha_q m_q} \\ &= \beta_1 n_1 + \beta_2 n_2 + \cdots + \beta_p n_p + \alpha_1 m_1 + \alpha_2 m_2 + \cdots + \alpha_q m_q.\end{aligned}$$

Hence, $v$ is a $\mathbf{k}$-linear combination of the vectors $n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q$.

Forget that we fixed $v$. We thus have shown that each $v \in M$ is a $\mathbf{k}$-linear combination of the vectors $n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q$. In other words, the list $\left(n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q\right)$ spans the $\mathbf{k}$-module $M$. This proves Claim 2. $\square$

Combining Claim 1 with Claim 2, we see that the list $\left(n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q\right)$ is a basis of the $\mathbf{k}$-module $M$.

**(a)** We have shown above that the list $\left(n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q\right)$ (which has size $p+q$) is a basis of $M$. Hence, the $\mathbf{k}$-module $M$ has a finite basis (of size $p+q$). In other words, the $\mathbf{k}$-module $M$ is free of finite rank. This proves Theorem 5.17.20 **(a)**.

**(b)** Recall that the $p$ vectors $n_1, n_2, \ldots, n_p$ belong to $N$ and thus to $M$ (since $N \subseteq M$). Let us define $q$ further vectors $n_{p+1}, n_{p+2}, \ldots, n_{p+q}$ in $M$ by setting

$$n_i := m_{i-p} \qquad \text{for each } i \in \{p+1, p+2, \ldots, p+q\}. \tag{320}$$

Thus, we have defined altogether $p+q$ vectors $n_1, n_2, \ldots, n_{p+q}$ in $M$. Note that the $q$ new vectors $n_{p+1}, n_{p+2}, \ldots, n_{p+q}$ do not belong to $N$ any more.

For each $j \in [q]$, we have $p+j \in \{p+1, p+2, \ldots, p+q\}$ and thus

$$\begin{aligned} n_{p+j} &= m_{p+j-p} \qquad \left(\text{by (320), applied to } i = p+j\right) \\ &= m_j \qquad \left(\text{since } p+j-p = j\right). \end{aligned} \tag{321}$$

In other words,

$$\left(n_{p+1}, n_{p+2}, \ldots, n_{p+q}\right) = \left(m_1, m_2, \ldots, m_q\right).$$

Moreover,

$$\begin{aligned} \left(n_1, n_2, \ldots, n_{p+q}\right) &= \Bigg(n_1, n_2, \ldots, n_p, \underbrace{n_{p+1}}_{\substack{=m_1 \\ \text{(by (321))}}}, \underbrace{n_{p+2}}_{\substack{=m_2 \\ \text{(by (321))}}}, \ldots, \underbrace{n_{p+q}}_{\substack{=m_q \\ \text{(by (321))}}}\Bigg) \\ &= \left(n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q\right). \end{aligned}$$

Hence, the list $\left(n_1, n_2, \ldots, n_{p+q}\right)$ is a basis of the $\mathbf{k}$-module $M$ (since the list $\left(n_1, n_2, \ldots, n_p, m_1, m_2, \ldots, m_q\right)$ is a basis of the $\mathbf{k}$-module $M$).

Fix $r \in R$. For any left $R$-module $V$, we let $\tau_{r,V} : V \to V$ be the $\mathbf{k}$-linear map $v \mapsto rv$. (This is the map denoted by $\tau_r$ in Definition 5.17.1.) The definition of the character $\chi_V$ yields that

$$\chi_V(r) = \operatorname{Tr}\left(\tau_{r,V}\right) \tag{322}$$

whenever $V$ is a left $R$-module that is free of finite rank as a $\mathbf{k}$-module.

We shall use the notation $\left(a_{i,j}\right)_{i,j\in[k]}$ for the $k \times k$-matrix whose $(i,j)$-th entry is $a_{i,j}$ for all $(i,j) \in [k] \times [k]$. Clearly, the trace of this matrix is $\sum_{i=1}^{k} a_{i,i}$.

Consider the $\mathbf{k}$-linear map $\tau_{r,M} : M \to M$. Let $\left(a_{i,j}\right)_{i,j\in[p+q]}$ be the matrix that represents this map $\tau_{r,M}$ with respect to the basis $\left(n_1, n_2, \ldots, n_{p+q}\right)$ of $M$. Thus, we have

$$\tau_{r,M}\left(n_j\right) = \sum_{i=1}^{p+q} a_{i,j} n_i \tag{323}$$

for each $j \in [p+q]$. Hence, for each $j \in [p+q]$, we have

$$\begin{aligned}
rn_j &= \tau_{r,M}\left(n_j\right) \qquad \left( \begin{array}{c} \text{since the definition of } \tau_{r,M} \\ \text{shows that } \tau_{r,M}\left(n_j\right) = rn_j \end{array} \right) \\
&= \sum_{i=1}^{p+q} a_{i,j} n_i \qquad \text{(by (323))} \\
&= \sum_{i=1}^{p} a_{i,j} n_i + \sum_{i=p+1}^{p+q} a_{i,j} \underbrace{n_i}_{\substack{=m_{i-p} \\ \text{(by (320))}}} \\
&= \sum_{i=1}^{p} a_{i,j} n_i + \sum_{i=p+1}^{p+q} a_{i,j} m_{i-p} \\
&= \sum_{i=1}^{p} a_{i,j} n_i + \sum_{i=1}^{q} a_{i+p,j} m_i
\end{aligned} \tag{324}$$

(here, we have substituted $i+p$ for $i$ in the second sum).

Now we claim the following:

*Claim 3:* We have $\tau_{r,N}\left(n_j\right) = \sum_{i=1}^{p} a_{i,j} n_i$ for each $j \in [p]$.

*Proof of Claim 3.* Let $j \in [p]$. Then, $n_j \in N$ (since the $p$ vectors $n_1, n_2, \ldots, n_p$ belong to $N$). Hence, $\tau_{r,N}\left(n_j\right) = rn_j$ (by the definition of $\tau_{r,N}$). Thus,

$$\tau_{r,N}\left(n_j\right) = rn_j = \sum_{i=1}^{p} a_{i,j} n_i + \sum_{i=1}^{q} a_{i+p,j} m_i \tag{325}$$

(by (324)). Solving this for $\sum_{i=1}^{q} a_{i+p,j} m_i$, we obtain

$$\sum_{i=1}^{q} a_{i+p,j} m_i = \tau_{r,N}\left(n_j\right) - \sum_{i=1}^{p} a_{i,j} n_i \in N$$

(since the vector $\tau_{r,N}\left(n_j\right)$ as well as the $p$ vectors $n_1, n_2, \ldots, n_p$ belong to $N$). Hence, in $M/N$, we have

$$\overline{\sum_{i=1}^{q} a_{i+p,j} m_i} = 0.$$

In view of

$$\overline{\sum_{i=1}^{q} a_{i+p,j} m_i} = \sum_{i=1}^{q} a_{i+p,j} \underbrace{\overline{m_i}}_{\substack{=c_i \\ \text{(by (315))}}} = \sum_{i=1}^{q} a_{i+p,j} c_i = a_{1+p,j} c_1 + a_{2+p,j} c_2 + \cdots + a_{q+p,j} c_q,$$

we can rewrite this as $a_{1+p,j} c_1 + a_{2+p,j} c_2 + \cdots + a_{q+p,j} c_q = 0$. Since the list $\left(c_1, c_2, \ldots, c_q\right)$ is $\mathbf{k}$-linearly independent (because it is a basis of $M/N$), this shows that all the coefficients $a_{1+p,j}, a_{2+p,j}, \ldots, a_{q+p,j}$ are 0. In other words,

$$a_{i+p,j} = 0 \qquad \text{for each } i \in [q]. \tag{326}$$

Hence, (325) becomes

$$\tau_{r,N}\left(n_j\right) = \sum_{i=1}^{p} a_{i,j} n_i + \sum_{i=1}^{q} \underbrace{a_{i+p,j}}_{\substack{=0 \\ \text{(by (326))}}} m_i = \sum_{i=1}^{p} a_{i,j} n_i + \underbrace{\sum_{i=1}^{q} 0 m_i}_{=0} = \sum_{i=1}^{p} a_{i,j} n_i.$$

This proves Claim 3. □

*Claim 4:* We have $\tau_{r,M/N}\left(c_j\right) = \sum_{i=1}^{q} a_{i+p,j+p} c_i$ for each $j \in [q]$.

*Proof of Claim 4.* Let $j \in [q]$. Then, (321) yields $n_{p+j} = m_j$. In other words, $n_{j+p} = m_j$ (since $p + j = j + p$). Hence, $r n_{j+p} = r m_j$, so that

$$r m_j = r n_{j+p} = \sum_{i=1}^{p} a_{i,j+p} n_i + \sum_{i=1}^{q} a_{i+p,j+p} m_i \tag{327}$$

(by (324), applied to $j + p$ instead of $j$).

Furthermore, $\overline{m_j} = c_j$ (by (315), applied to $i = j$). Hence, $r\overline{m_j} = r c_j$, so that

$$\begin{aligned} r c_j = r\overline{m_j} = \overline{r m_j} &= \overline{\sum_{i=1}^{p} a_{i,j+p} n_i + \sum_{i=1}^{q} a_{i+p,j+p} m_i} \qquad \text{(by (327))} \\ &= \sum_{i=1}^{p} a_{i,j+p} \underbrace{\overline{n_i}}_{\substack{=0 \\ \text{(since } n_i \in N \\ \text{(because the } p \\ \text{vectors } n_1, n_2, \ldots, n_p \\ \text{belong to } N\text{))}}} + \sum_{i=1}^{q} a_{i+p,j+p} \underbrace{\overline{m_i}}_{\substack{=c_i \\ \text{(by (315))}}} \\ &= \underbrace{\sum_{i=1}^{p} a_{i,j+p} 0}_{=0} + \sum_{i=1}^{q} a_{i+p,j+p} c_i = \sum_{i=1}^{q} a_{i+p,j+p} c_i. \end{aligned}$$

But the definition of $\tau_{r,M/N}$ yields $\tau_{r,M/N}\left(c_j\right) = rc_j = \sum_{i=1}^{q} a_{i+p,j+p}c_i$. This proves Claim 4. $\square$

We are now ready to compute the character values $\chi_M\left(r\right)$, $\chi_N\left(r\right)$ and $\chi_{M/N}\left(r\right)$:

- Applying (322) to $V = M$, we obtain $\chi_M\left(r\right) = \operatorname{Tr}\left(\tau_{r,M}\right)$. But the matrix that represents the $\mathbf{k}$-linear map $\tau_{r,M} : M \to M$ with respect to the basis $\left(n_1, n_2, \ldots, n_{p+q}\right)$ of $M$ is the matrix $\left(a_{i,j}\right)_{i,j\in[p+q]}$. Hence,

$$\operatorname{Tr}\left(\tau_{r,M}\right) = \operatorname{Tr}\left(\left(a_{i,j}\right)_{i,j\in[p+q]}\right) = \sum_{i=1}^{p+q} a_{i,i} = \sum_{i=1}^{p} a_{i,i} + \sum_{i=p+1}^{p+q} a_{i,i}.$$

  In view of $\chi_M\left(r\right) = \operatorname{Tr}\left(\tau_{r,M}\right)$, we can rewrite this as

$$\chi_M\left(r\right) = \sum_{i=1}^{p} a_{i,i} + \sum_{i=p+1}^{p+q} a_{i,i}. \tag{328}$$

- Applying (322) to $V = N$, we obtain $\chi_N\left(r\right) = \operatorname{Tr}\left(\tau_{r,N}\right)$. But the matrix that represents the $\mathbf{k}$-linear map $\tau_{r,N} : N \to N$ with respect to the basis $\left(n_1, n_2, \ldots, n_p\right)$ of $N$ is the matrix $\left(a_{i,j}\right)_{i,j\in[p]}$, since each $j \in [p]$ satisfies

$$\tau_{r,N}\left(n_j\right) = \sum_{i=1}^{p} a_{i,j}n_i \qquad \text{(by Claim 3)}.$$

  Hence,

$$\operatorname{Tr}\left(\tau_{r,N}\right) = \operatorname{Tr}\left(\left(a_{i,j}\right)_{i,j\in[p]}\right) = \sum_{i=1}^{p} a_{i,i}.$$

  In view of $\chi_N\left(r\right) = \operatorname{Tr}\left(\tau_{r,N}\right)$, we can rewrite this as

$$\chi_N\left(r\right) = \sum_{i=1}^{p} a_{i,i}. \tag{329}$$

- Applying (322) to $V = M/N$, we obtain $\chi_{M/N}\left(r\right) = \operatorname{Tr}\left(\tau_{r,M/N}\right)$. But the matrix that represents the $\mathbf{k}$-linear map $\tau_{r,M/N} : M/N \to M/N$ with respect to the basis $\left(c_1, c_2, \ldots, c_q\right)$ of $M/N$ is the matrix $\left(a_{i+p,j+p}\right)_{i,j\in[q]}$, since each $j \in [q]$ satisfies

$$\tau_{r,M/N}\left(c_j\right) = \sum_{i=1}^{q} a_{i+p,j+p}c_i \qquad \text{(by Claim 4)}.$$

  Hence,

$$\operatorname{Tr}\left(\tau_{r,M/N}\right) = \operatorname{Tr}\left(\left(a_{i+p,j+p}\right)_{i,j\in[q]}\right) = \sum_{i=1}^{q} a_{i+p,i+p} = \sum_{i=p+1}^{p+q} a_{i,i}$$

(here, we have substituted $i$ for $i+p$ in the sum). In view of $\chi_{M/N}(r) = \operatorname{Tr}(\tau_{r,M/N})$, we can rewrite this as

$$\chi_{M/N}(r) = \sum_{i=p+1}^{p+q} a_{i,i}. \tag{330}$$

Now,

$$\begin{aligned}
(\chi_M - \chi_N)(r) &= \chi_M(r) - \chi_N(r) \\
&= \left(\sum_{i=1}^{p} a_{i,i} + \sum_{i=p+1}^{p+q} a_{i,i}\right) - \left(\sum_{i=1}^{p} a_{i,i}\right) \qquad \text{(by (328) and (329))} \\
&= \sum_{i=p+1}^{p+q} a_{i,i} = \chi_{M/N}(r) \qquad \text{(by (330))}.
\end{aligned}$$

Forget that we fixed $r$. We thus have shown that $(\chi_M - \chi_N)(r) = \chi_{M/N}(r)$ for each $r \in R$. In other words, $\chi_M - \chi_N = \chi_{M/N}$. This proves Theorem 5.17.20 **(b)**. □

As an example for the use of Theorem 5.17.20, we can now easily compute the character of the zero-sum representation $R(\mathbf{k}^n)$ of $S_n$ defined in Subsection 4.2.6:

**Example 5.17.21.** Let $n \geq 1$. Consider the natural representation $\mathbf{k}^n$ of the symmetric group $S_n$ (defined in Example 4.2.17), and its two subrepresentations $D(\mathbf{k}^n)$ and $R(\mathbf{k}^n)$ (defined in Subsection 4.2.6). Then, for each $g \in S_n$, we have

$$\chi_{D(\mathbf{k}^n)}(g) = 1 \qquad \text{and} \qquad \chi_{R(\mathbf{k}^n)}(g) = (\#\text{ of fixed points of } g) - 1.$$

*Proof.* Consider the trivial representation $\mathbf{k}_{\text{triv}}$ of $S_n$ (defined in Example 4.2.16). Let us denote its character $\chi_{\mathbf{k}_{\text{triv}}}$ by $\chi_{\text{triv}}$. Thus, $\chi_{\mathbf{k}_{\text{triv}}} = \chi_{\text{triv}}$.

Let us furthermore denote the character $\chi_{\mathbf{k}^n}$ of the natural representation $\mathbf{k}^n$ of $S_n$ by $\chi_{\text{nat}}$. Thus, $\chi_{\mathbf{k}^n} = \chi_{\text{nat}}$.

Let $g \in S_n$. We already know (from Subsection 4.2.6) that as an $S_n$-representation, $D(\mathbf{k}^n)$ is isomorphic to the trivial representation $\mathbf{k}_{\text{triv}}$. Thus, Proposition 5.17.13 (applied to $R = \mathbf{k}[S_n]$ and $V = D(\mathbf{k}^n)$ and $W = \mathbf{k}_{\text{triv}}$) shows that $\chi_{D(\mathbf{k}^n)} = \chi_{\mathbf{k}_{\text{triv}}} = \chi_{\text{triv}}$. But Example 5.17.7 **(a)** (applied to $G = S_n$) yields that $\chi_{\text{triv}}(g) = 1$. In view of $\chi_{D(\mathbf{k}^n)} = \chi_{\text{triv}}$, we can rewrite this as $\chi_{D(\mathbf{k}^n)}(g) = 1$.

It remains to show that $\chi_{R(\mathbf{k}^n)}(g) = (\#\text{ of fixed points of } g) - 1$. For this, we recall that Example 5.17.9 yields

$$\chi_{\text{nat}}(g) = (\#\text{ of fixed points of } g).$$

But Theorem 4.2.30 **(d)** yields that $\mathbf{k}^n / R(\mathbf{k}^n) \cong D(\mathbf{k}^n)$ as $S_n$-representations. Hence, $\mathbf{k}^n / R(\mathbf{k}^n)$ is a free $\mathbf{k}$-module of finite rank (since $D(\mathbf{k}^n)$ is a free $\mathbf{k}$-module of finite rank). Thus, Proposition 5.17.13 (applied to $R = \mathbf{k}[S_n]$ and $V = \mathbf{k}^n / R(\mathbf{k}^n)$ and $W = D(\mathbf{k}^n)$) shows that

$$\chi_{\mathbf{k}^n / R(\mathbf{k}^n)} = \chi_{D(\mathbf{k}^n)} \tag{331}$$

(since $\mathbf{k}^n / R(\mathbf{k}^n) \cong D(\mathbf{k}^n)$ as $S_n$-representations).

But Theorem 5.17.20 **(b)** (applied to $M = \mathbf{k}^n$ and $V = R(\mathbf{k}^n)$) yields $\chi_{\mathbf{k}^n / R(\mathbf{k}^n)} = \chi_{\mathbf{k}^n} - \chi_{R(\mathbf{k}^n)}$. In view of (331), we can rewrite this as $\chi_{D(\mathbf{k}^n)} = \chi_{\mathbf{k}^n} - \chi_{R(\mathbf{k}^n)}$. Solving this for $\chi_{R(\mathbf{k}^n)}$, we obtain

$$\chi_{R(\mathbf{k}^n)} = \underbrace{\chi_{\mathbf{k}^n}}_{=\chi_{\mathrm{nat}}} - \underbrace{\chi_{D(\mathbf{k}^n)}}_{=\chi_{\mathrm{triv}}} = \chi_{\mathrm{nat}} - \chi_{\mathrm{triv}}.$$

Thus,

$$\begin{aligned}\chi_{R(\mathbf{k}^n)}(g) &= (\chi_{\mathrm{nat}} - \chi_{\mathrm{triv}})(g) = \underbrace{\chi_{\mathrm{nat}}(g)}_{=(\#\text{ of fixed points of } g)} - \underbrace{\chi_{\mathrm{triv}}(g)}_{=1} \\ &= (\#\text{ of fixed points of } g) - 1.\end{aligned}$$

This completes the proof of Example 5.17.21. □

Another consequence of Theorem 5.17.20 is the following formula for characters of direct sums:

**Corollary 5.17.22.** Let $R$ be a $\mathbf{k}$-algebra. Let $V$ and $W$ be two left $R$-modules. Assume that, as $\mathbf{k}$-modules, both $V$ and $W$ are free of finite ranks (not necessarily of the same rank). Then,

$$\chi_{V \oplus W} = \chi_V + \chi_W.$$

(Here, the addition is pointwise, as usual for $\mathbf{k}$-linear maps from $R$ to $\mathbf{k}$.)

*Proof.* It is well-known that the $\mathbf{k}$-module $V \oplus W$ is free of finite rank (since $V$ and $W$ are).

Consider the left $R$-submodule

$$V \oplus 0 := \{(v, 0) \mid v \in V\}$$

of $V \oplus W$. It is well-known that this $R$-submodule $V \oplus 0$ is isomorphic to $V$ (via the isomorphism $V \oplus 0 \to V$ that sends each $(v, 0)$ to $v$). Thus, the $\mathbf{k}$-module $V \oplus 0$ is free of finite rank (since $V$ is). Hence, Proposition 5.17.13 (applied to $V \oplus 0$ and $V$ instead of $V$ and $W$) yields $\chi_{V \oplus 0} = \chi_V$ (since $V \oplus 0 \cong V$ as left $R$-modules).

But it is also well-known that the quotient $R$-module $(V \oplus W) / (V \oplus 0)$ is isomorphic to the $R$-module $W$ (via the isomorphism $(V \oplus W) / (V \oplus 0) \to W$ that sends each

$\overline{(v,w)} \in (V \oplus W) \,/\, (V \oplus 0)$ to $w$). Thus, the $\mathbf{k}$-module $(V \oplus W) \,/\, (V \oplus 0)$ is free of finite rank (since $W$ is). Hence, Proposition 5.17.13 (applied to $(V \oplus W) \,/\, (V \oplus 0)$ instead of $V$) yields $\chi_{(V \oplus W)/(V \oplus 0)} = \chi_W$ (since $(V \oplus W) \,/\, (V \oplus 0) \cong W$ as left $R$-modules).

Now, Theorem 5.17.20 **(b)** (applied to $M = V \oplus W$ and $N = V \oplus 0$) yields $\chi_{(V \oplus W)/(V \oplus 0)} = \chi_{V \oplus W} - \chi_{V \oplus 0}$. In view of $\chi_{(V \oplus W)/(V \oplus 0)} = \chi_W$ and $\chi_{V \oplus 0} = \chi_V$, we can rewrite this as $\chi_W = \chi_{V \oplus W} - \chi_V$. Thus, $\chi_{V \oplus W} = \chi_V + \chi_W$. This proves Corollary 5.17.22. ☐

We can easily extend Corollary 5.17.22 to direct sums of multiple $R$-modules:

**Corollary 5.17.23.** Let $R$ be a $\mathbf{k}$-algebra. Let $V_1, V_2, \ldots, V_m$ be finitely many left $R$-modules. Assume that, as a $\mathbf{k}$-module, each $V_i$ is free of finite rank. Then,

$$\chi_{V_1 \oplus V_2 \oplus \cdots \oplus V_m} = \chi_{V_1} + \chi_{V_2} + \cdots + \chi_{V_m}.$$

(Here, the addition is pointwise, as usual for $\mathbf{k}$-linear maps from $R$ to $\mathbf{k}$.)

*Proof sketch.* Induct on $m$. The *base case* $m = 0$ follows from the easy fact that the character of a trivial $R$-module is 0. The *induction step* uses Corollary 5.17.22 and the observation that $V_1 \oplus V_2 \oplus \cdots \oplus V_m \cong (V_1 \oplus V_2 \oplus \cdots \oplus V_{m-1}) \oplus V_m$. ☐

### 5.17.6. The quasiidempotent character formula

We shall now state an important formula for the characters of certain representations of a finite group $G$: specifically, for those representations that can be written as $\mathcal{A}\mathbf{a}$ for some quasi-idempotent element $\mathbf{a} \in \mathbf{k}[G]$. This class of representations includes the Specht modules $\mathcal{S}^\lambda$ for partitions $\lambda$ of $n$, since Theorem 5.11.3 shows that the respective Young symmetrizers $\mathbf{E}_T$ are quasi-idempotent.

Before we can state our formula, we recall Definition 5.11.6. The notation $[w]\,\mathbf{a}$ introduced in this definition has the following simple but important property:

**Proposition 5.17.24.** Let $G$ be a group. Let $u, v \in G$ and $\mathbf{b} \in \mathbf{k}[G]$. Then,

$$[uv]\,\mathbf{b} = [v]\left(u^{-1}\mathbf{b}\right) \tag{332}$$

$$= [u]\left(\mathbf{b}v^{-1}\right). \tag{333}$$

*Proof.* Recall that $G$ is a group. Thus, the map $G \to G,\ g \mapsto ug$ is a bijection (and its inverse is the map $G \to G,\ g \mapsto u^{-1}g$). For similar reasons, the map $G \to G,\ g \mapsto gv$ is a bijection.

Write the element $\mathbf{b} \in \mathbf{k}[G]$ as a $\mathbf{k}$-linear combination of the standard basis vectors $g \in G$ as follows:

$$\mathbf{b} = \sum_{g \in G} \beta_g g, \tag{334}$$

where the coefficients $\beta_g$ belong to $\mathbf{k}$. Then, $[g]\,\mathbf{b} = \beta_g$ for each $g \in G$ (by Definition 5.11.6). Applying this to $g = uv$, we obtain $[uv]\,\mathbf{b} = \beta_{uv}$.

On the other hand, multiplying the equality (334) by $u^{-1}$ from the left, we obtain

$$u^{-1}\mathbf{b} = u^{-1}\sum_{g\in G}\beta_g g = \sum_{g\in G}\beta_g u^{-1}g = \sum_{g\in G}\beta_{ug}\underbrace{u^{-1}u}_{=1}g$$
$$\left(\begin{array}{c}\text{here, we have substituted } ug \text{ for } g \text{ in the sum, since}\\ \text{the map } G\to G,\ g\mapsto ug \text{ is a bijection}\end{array}\right)$$
$$= \sum_{g\in G}\beta_{ug}g.$$

Hence, $[g]\left(u^{-1}\mathbf{b}\right) = \beta_{ug}$ for each $g\in G$ (by Definition 5.11.6). Applying this to $g = v$, we obtain $[v]\left(u^{-1}\mathbf{b}\right) = \beta_{uv}$. Comparing this with $[uv]\,\mathbf{b} = \beta_{uv}$, we find $[uv]\,\mathbf{b} = [v]\left(u^{-1}\mathbf{b}\right)$. This proves (332).

Furthermore, multiplying the equality (334) by $v^{-1}$ from the right, we obtain

$$\mathbf{b}v^{-1} = \left(\sum_{g\in G}\beta_g g\right)v^{-1} = \sum_{g\in G}\beta_g g v^{-1} = \sum_{g\in G}\beta_{gv}g\underbrace{vv^{-1}}_{=1}$$
$$\left(\begin{array}{c}\text{here, we have substituted } gv \text{ for } g \text{ in the sum, since}\\ \text{the map } G\to G,\ g\mapsto gv \text{ is a bijection}\end{array}\right)$$
$$= \sum_{g\in G}\beta_{gv}g.$$

Hence, $[g]\left(\mathbf{b}v^{-1}\right) = \beta_{gv}$ for each $g\in G$ (by Definition 5.11.6). Applying this to $g = u$, we obtain $[u]\left(\mathbf{b}v^{-1}\right) = \beta_{uv}$. Comparing this with $[uv]\,\mathbf{b} = \beta_{uv}$, we find $[uv]\,\mathbf{b} = [u]\left(\mathbf{b}v^{-1}\right)$. This proves (333).

Since we have proved both (332) and (333), we thus know that Proposition 5.17.24 holds. □

We are now ready to state our formula for characters:

**Theorem 5.17.25.** Let $G$ be a finite group. Let $\mathcal{A} = \mathbf{k}[G]$. (We do **not** follow the convention that $\mathcal{A} = \mathbf{k}[S_n]$ in this theorem.) Let $\mathbf{a}\in\mathcal{A}$ and $\kappa\in\mathbf{k}$ be such that $\mathbf{a}^2 = \kappa\mathbf{a}$. Assume that the $\mathbf{k}$-module $\mathcal{A}\mathbf{a}$ is free of finite rank. Then, the character $\chi_{\mathcal{A}\mathbf{a}}$ of the left $\mathbf{k}[G]$-module $\mathcal{A}\mathbf{a}$ satisfies

$$\kappa\cdot\chi_{\mathcal{A}\mathbf{a}}(h) = \sum_{w\in G}\left[h^{-1}\right]\left(w\mathbf{a}w^{-1}\right) \qquad \text{for each } h\in G.$$

*Proof.* The following proof is somewhat similar to the proof of Lemma 5.11.11 above.

First, we observe that $\mathcal{A}\mathbf{a}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$, thus a left $\mathcal{A}$-module, i.e., a left $\mathbf{k}[G]$-module (since $\mathcal{A} = \mathbf{k}[G]$). Thus, its character $\chi_{\mathcal{A}\mathbf{a}}$ is well-defined (since the $\mathbf{k}$-module $\mathcal{A}\mathbf{a}$ is free of finite rank).

We will need the following two facts:

*Claim 1:* Let $f$ be any $\mathbf{k}$-module endomorphism of $\mathbf{k}[G]$. Then,

$$\operatorname{Tr} f = \sum_{w \in G} [w] (f(w)).$$

*Proof of Claim 1.* This is just Claim 1 in the proof of Proposition 5.11.7, except that we have renamed the summation index $g$ as $w$. □

*Claim 2:* Let $h, w \in G$. Then,

$$[w] (hw\mathbf{a}) = \left[h^{-1}\right] \left(w\mathbf{a}w^{-1}\right).$$

*Proof of Claim 2.* Applying (333) to $u = 1_G$ and $v = w$ and $\mathbf{b} = hw\mathbf{a}$, we obtain $[1_G w] (hw\mathbf{a}) = [1_G] \left(hw\mathbf{a}w^{-1}\right)$. In other words, $[w] (hw\mathbf{a}) = [1_G] \left(hw\mathbf{a}w^{-1}\right)$ (since $1_G w = w$).

On the other hand, applying (332) to $u = h$ and $v = h^{-1}$ and $\mathbf{b} = hw\mathbf{a}w^{-1}$, we obtain $\left[hh^{-1}\right] \left(hw\mathbf{a}w^{-1}\right) = \left[h^{-1}\right] \left(\underbrace{h^{-1}h}_{=1} w\mathbf{a}w^{-1}\right) = \left[h^{-1}\right] \left(w\mathbf{a}w^{-1}\right)$. In view of $hh^{-1} = 1_G$, we can rewrite this as $[1_G] \left(hw\mathbf{a}w^{-1}\right) = \left[h^{-1}\right] \left(w\mathbf{a}w^{-1}\right)$.

Altogether, we now have $[w] (hw\mathbf{a}) = [1_G] \left(hw\mathbf{a}w^{-1}\right) = \left[h^{-1}\right] \left(w\mathbf{a}w^{-1}\right)$. This proves Claim 2. □

Now, fix $h \in G$. Let $f : \mathcal{A} \to \mathcal{A}\mathbf{a}$ be the map that sends each $\mathbf{b} \in \mathcal{A}$ to $h\mathbf{b}\mathbf{a}$. (This is well-defined, since each $\mathbf{b} \in \mathcal{A}$ satisfies $\underbrace{h\mathbf{b}}_{\in \mathcal{A}} \mathbf{a} \in \mathcal{A}\mathbf{a}$.) Let $g : \mathcal{A}\mathbf{a} \to \mathcal{A}$ be the inclusion map (i.e., the map that sends each $\mathbf{b} \in \mathbf{a}\mathcal{A}$ to $\mathbf{b}$ itself). Both maps $f$ and $g$ are clearly $\mathbf{k}$-linear. Hence, Lemma 5.11.9 (applied to $V = \mathcal{A}$ and $W = \mathcal{A}\mathbf{a}$) yields

$$\operatorname{Tr} (f \circ g) = \operatorname{Tr} (g \circ f). \tag{335}$$

However, the map $g \circ f : \mathcal{A} \to \mathcal{A}$ is $\mathbf{k}$-linear (since $g$ and $f$ are), and thus is a $\mathbf{k}$-module endomorphism of $\mathcal{A} = \mathbf{k}[G]$. Hence, Claim 1 (applied to $g \circ f$ instead of $f$) yields

$$\operatorname{Tr} (g \circ f) = \sum_{w \in G} [w] ((g \circ f)(w)). \tag{336}$$

However, each $w \in G$ satisfies

$$\begin{aligned} (g \circ f)(w) &= g(f(w)) = f(w) && \text{(by the definition of } g\text{)} \\ &= hw\mathbf{a} && \text{(by the definition of } f\text{)} \end{aligned}$$

and therefore

$$[w] \left(\underbrace{(g \circ f)(w)}_{=hw\mathbf{a}}\right) = [w] (hw\mathbf{a}) = \left[h^{-1}\right] \left(w\mathbf{a}w^{-1}\right) \tag{337}$$

(by Claim 2). Hence, (336) becomes

$$\begin{aligned}\operatorname{Tr}(g \circ f) &= \sum_{w \in G} \underbrace{[w]\left((g \circ f)(w)\right)}_{\substack{=\left[h^{-1}\right]\left(w\mathbf{a}w^{-1}\right) \\ \text{(by (337))}}} \\ &= \sum_{w \in G} \left[h^{-1}\right]\left(w\mathbf{a}w^{-1}\right). \end{aligned} \tag{338}$$

On the other hand, for each $r \in \mathcal{A}$, we let $\tau_r : \mathcal{A}\mathbf{a} \to \mathcal{A}\mathbf{a}$ be the $\mathbf{k}$-linear map $v \mapsto rv$. Then, the definition of the character $\chi_{\mathcal{A}\mathbf{a}}$ yields

$$\chi_{\mathcal{A}\mathbf{a}}(r) = \operatorname{Tr}(\tau_r) \qquad \text{for each } r \in \mathcal{A}.$$

Applying this to $r = \kappa h$, we obtain

$$\chi_{\mathcal{A}\mathbf{a}}(\kappa h) = \operatorname{Tr}(\tau_{\kappa h}). \tag{339}$$

Next, we claim that

$$\tau_{\kappa h} = f \circ g. \tag{340}$$

*Proof of (340):* Let $\mathbf{b} \in \mathcal{A}\mathbf{a}$. Thus, $\mathbf{b} = \mathbf{ca}$ for some $\mathbf{c} \in \mathcal{A}$. Consider this $\mathbf{c}$. Then, the definition of $\tau_{\kappa h}$ yields $\tau_{\kappa h}(\mathbf{b}) = \kappa h \underbrace{\mathbf{b}}_{=\mathbf{ca}} = \kappa h \mathbf{ca}$. On the other hand, the definition of $g$ yields $g(\mathbf{b}) = \mathbf{b}$. Applying the map $f$ to this equality, we find $f(g(\mathbf{b})) = f(\mathbf{b}) = h\mathbf{ba}$ (by the definition of $f$). Hence,

$$(f \circ g)(\mathbf{b}) = f(g(\mathbf{b})) = h \underbrace{\mathbf{b}}_{=\mathbf{ca}} \mathbf{a} = h\mathbf{c} \underbrace{\mathbf{aa}}_{=\mathbf{a}^2 = \kappa\mathbf{a}} = h\mathbf{c} \cdot \kappa\mathbf{a} = \kappa h \mathbf{ca}.$$

Comparing this with $\tau_{\kappa h}(\mathbf{b}) = \kappa h \mathbf{ca}$, we obtain $\tau_{\kappa h}(\mathbf{b}) = (f \circ g)(\mathbf{b})$.

Forget that we fixed $\mathbf{b}$. We thus have shown that $\tau_{\kappa h}(\mathbf{b}) = (f \circ g)(\mathbf{b})$ for each $\mathbf{b} \in \mathcal{A}\mathbf{a}$. In other words, $\tau_{\kappa h} = f \circ g$. This proves (340). $\square$

Now, (335) yields

$$\begin{aligned}\operatorname{Tr}(g \circ f) &= \operatorname{Tr}\underbrace{(f \circ g)}_{\substack{=\tau_{\kappa h} \\ \text{(by (340))}}} = \operatorname{Tr}(\tau_{\kappa h}) = \chi_{\mathcal{A}\mathbf{a}}(\kappa h) \qquad \text{(by (339))} \\ &= \kappa \cdot \chi_{\mathcal{A}\mathbf{a}}(h)\end{aligned}$$

(since Proposition 5.17.2 shows that $\chi_{\mathcal{A}\mathbf{a}}$ is a $\mathbf{k}$-linear map). Comparing this with (338), we obtain

$$\kappa \cdot \chi_{\mathcal{A}\mathbf{a}}(h) = \sum_{w \in G} \left[h^{-1}\right]\left(w\mathbf{a}w^{-1}\right).$$

This proves Theorem 5.17.25. $\square$

**Corollary 5.17.26.** Let $G$ be a finite group. Define the $\mathbf{k}$-linear map

$$\begin{aligned} S : \mathbf{k}[G] &\to \mathbf{k}[G], \\ w &\mapsto w^{-1} \qquad \text{for all } w \in G. \end{aligned}$$

This is understood as in Definition 3.11.2. (Note that this map $S$ generalizes the antipode $S$ of $\mathbf{k}[S_n]$, and was already mentioned in Remark 3.11.15.)

Let $\mathcal{A} = \mathbf{k}[G]$. (We do **not** follow the convention that $\mathcal{A} = \mathbf{k}[S_n]$ in this theorem.) Let $\mathbf{a} \in \mathcal{A}$ and $\kappa \in \mathbf{k}$ be such that $\mathbf{a}^2 = \kappa \mathbf{a}$. Assume that the $\mathbf{k}$-module $\mathcal{A}\mathbf{a}$ is free of finite rank. Then,

$$\kappa \cdot S\left(\mathbf{X}_{\mathcal{A}\mathbf{a}}\right) = \sum_{w \in G} w \mathbf{a} w^{-1}.$$

(See Definition 5.17.15 for the meaning of $\mathbf{X}_{\mathcal{A}\mathbf{a}}$.)

*Proof.* First, we observe that $\mathcal{A}\mathbf{a}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$, thus a left $\mathcal{A}$-module, i.e., a left $\mathbf{k}[G]$-module (since $\mathcal{A} = \mathbf{k}[G]$). Thus, $\mathbf{X}_{\mathcal{A}\mathbf{a}}$ is well-defined (since the $\mathbf{k}$-module $\mathcal{A}\mathbf{a}$ is free of finite rank).

Note that the map $G \to G$, $g \mapsto g^{-1}$ is a bijection (since it is inverse to itself).

The definition of $\mathbf{X}_{\mathcal{A}\mathbf{a}}$ yields

$$\mathbf{X}_{\mathcal{A}\mathbf{a}} = \sum_{g \in G} \chi_{\mathcal{A}\mathbf{a}}(g)\, g.$$

Applying the map $S$ to this equality, we obtain

$$\begin{aligned} S\left(\mathbf{X}_{\mathcal{A}\mathbf{a}}\right) &= S\left(\sum_{g \in G} \chi_{\mathcal{A}\mathbf{a}}(g)\, g\right) \\ &= \sum_{g \in G} \chi_{\mathcal{A}\mathbf{a}}(g) \underbrace{S(g)}_{\substack{= g^{-1} \\ \text{(by the definition of } S\text{)}}} \qquad \text{(since the map } S \text{ is } \mathbf{k}\text{-linear)} \\ &= \sum_{g \in G} \chi_{\mathcal{A}\mathbf{a}}(g)\, g^{-1}. \end{aligned}$$

Multiplying both sides of this equality by $\kappa$, we find

$$\begin{aligned} \kappa \cdot S\left(\mathbf{X}_{\mathcal{A}\mathbf{a}}\right) &= \kappa \cdot \sum_{g \in G} \chi_{\mathcal{A}\mathbf{a}}(g)\, g^{-1} = \sum_{g \in G} \underbrace{\kappa \cdot \chi_{\mathcal{A}\mathbf{a}}(g)}_{\substack{= \sum_{w \in G} \left[g^{-1}\right]\left(w\mathbf{a}w^{-1}\right) \\ \text{(by Theorem 5.17.25,} \\ \text{applied to } h=g\text{)}}} g^{-1} \\ &= \sum_{g \in G} \left( \sum_{w \in G} \left[g^{-1}\right]\left(w\mathbf{a}w^{-1}\right) \right) g^{-1} = \sum_{g \in G} \sum_{w \in G} \left( \left[g^{-1}\right]\left(w\mathbf{a}w^{-1}\right) \right) \cdot g^{-1} \\ &= \sum_{g \in G} \sum_{w \in G} \left( [g]\left(w\mathbf{a}w^{-1}\right) \right) \cdot g \end{aligned}$$

(here, we have substituted $g$ for $g^{-1}$ in the outer sum, since the map $G \to G$, $g \mapsto g^{-1}$ is a bijection). Thus,

$$\begin{aligned}\kappa \cdot S\left(\mathbf{X}_{\mathcal{A}\mathbf{a}}\right) &= \underbrace{\sum_{g\in G} \sum_{w\in G}}_{= \sum_{w\in G} \sum_{g\in G}} \left([g]\left(w\mathbf{a}w^{-1}\right)\right) \cdot g \\ &= \sum_{w\in G} \sum_{g\in G} \left([g]\left(w\mathbf{a}w^{-1}\right)\right) \cdot g. \end{aligned} \tag{341}$$

However, it is easy to see that each $\mathbf{b} \in \mathcal{A}$ satisfies

$$\sum_{g\in G} \left([g]\,\mathbf{b}\right) \cdot g = \mathbf{b} \tag{342}$$

[118]. Hence, (341) becomes

$$\kappa \cdot S\left(\mathbf{X}_{\mathcal{A}\mathbf{a}}\right) = \sum_{w\in G} \underbrace{\sum_{g\in G} \left([g]\left(w\mathbf{a}w^{-1}\right)\right) \cdot g}_{\substack{= w\mathbf{a}w^{-1} \\ \text{(by (342), applied to } \mathbf{b}=w\mathbf{a}w^{-1}\text{)}}} = \sum_{w\in G} w\mathbf{a}w^{-1}.$$

This proves Corollary 5.17.26. □

When $G$ is the symmetric group $S_n$, we can simplify Corollary 5.17.26 by removing the map $S$:

**Corollary 5.17.27.** Let $\mathcal{A} = \mathbf{k}\left[S_n\right]$. Let $\mathbf{a} \in \mathcal{A}$ and $\kappa \in \mathbf{k}$ be such that $\mathbf{a}^2 = \kappa\mathbf{a}$. Assume that the $\mathbf{k}$-module $\mathcal{A}\mathbf{a}$ is free of finite rank. Then,

$$\kappa \cdot \mathbf{X}_{\mathcal{A}\mathbf{a}} = \sum_{w\in S_n} w\mathbf{a}w^{-1}.$$

(See Definition 5.17.15 for the meaning of $\mathbf{X}_{\mathcal{A}\mathbf{a}}$.)

*Proof.* Let $G = S_n$. Then, the map $S$ defined in Corollary 5.17.26 is precisely the antipode $S$ of $\mathbf{k}\left[S_n\right]$. Hence, Corollary 5.17.26 yields

$$\kappa \cdot S\left(\mathbf{X}_{\mathcal{A}\mathbf{a}}\right) = \sum_{w\in G} w\mathbf{a}w^{-1}. \tag{343}$$

[118] *Proof:* Let $\mathbf{b} \in \mathcal{A}$. Write the element $\mathbf{b} \in \mathcal{A} = \mathbf{k}\left[G\right]$ as a $\mathbf{k}$-linear combination of the standard basis vectors $g \in G$ as follows:

$$\mathbf{b} = \sum_{g\in G} \beta_g g,$$

where the coefficients $\beta_g$ belong to $\mathbf{k}$. Then, $[g]\,\mathbf{b} = \beta_g$ for each $g \in G$ (by Definition 5.11.6). Hence, $\sum_{g\in G} \underbrace{\left([g]\,\mathbf{b}\right)}_{=\beta_g} \cdot g = \sum_{g\in G} \beta_g g = \mathbf{b}$. This proves (342).

But Proposition 5.17.17 (applied to $V = \mathcal{A}\mathbf{a}$) yields that $\mathbf{X}_{\mathcal{A}\mathbf{a}}$ belongs to the center $Z(\mathbf{k}[S_n])$ of the ring $\mathbf{k}[S_n]$. In other words, $\mathbf{X}_{\mathcal{A}\mathbf{a}} \in Z(\mathbf{k}[S_n])$. Hence, Proposition 3.13.3 (applied to $\mathbf{X}_{\mathcal{A}\mathbf{a}}$ instead of $\mathbf{a}$) yields $S(\mathbf{X}_{\mathcal{A}\mathbf{a}}) = \mathbf{X}_{\mathcal{A}\mathbf{a}}$. Hence, we can rewrite (343) as

$$\kappa \cdot \mathbf{X}_{\mathcal{A}\mathbf{a}} = \sum_{w \in G} w\mathbf{a}w^{-1} = \sum_{w \in S_n} w\mathbf{a}w^{-1} \qquad (\text{since } G = S_n).$$

This proves Corollary 5.17.27. □

### 5.17.7. Characters and centralized Young symmetrizers

We are now ready to state the main theorem about characters of Specht modules:

**Theorem 5.17.28.** Let $\lambda$ be a partition of $n$. Consider the Specht module $\mathcal{S}^\lambda$, and let $\mathbf{X}_\lambda$ denote the element $\mathbf{X}_{\mathcal{S}^\lambda}$ corresponding to it (defined as in Definition 5.17.15). Then,

$$\mathbf{E}_\lambda = h^\lambda \mathbf{X}_\lambda.$$

(See Definition 5.12.11 **(b)** for the meaning of $h^\lambda$, and Definition 5.12.1 for the meaning of $\mathbf{E}_\lambda$.)

*Proof.* Recall that $\mathcal{A} = \mathbf{k}[S_n]$. Pick any $n$-tableau $T$ of shape $Y(\lambda)$ (clearly, such an $n$-tableau exists, since $|Y(\lambda)| = |\lambda| = n$). Then, (228) (applied to $D = Y(\lambda)$) yields $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$ (as left $\mathcal{A}$-modules). This is an isomorphism of left $\mathcal{A}$-modules, i.e., of left $\mathbf{k}[S_n]$-modules (since $\mathcal{A} = \mathbf{k}[S_n]$), thus an isomorphism of representations of $S_n$, and therefore also an isomorphism of $\mathbf{k}$-modules.

Moreover, $Y(\lambda) = Y(\lambda/\varnothing)$ is a skew Young diagram. Hence, Theorem 5.9.1 (applied to $D = Y(\lambda)$) shows that the standard polytabloids form a basis of the $\mathbf{k}$-module $\mathcal{S}^{Y(\lambda)}$. Hence, $\mathcal{S}^{Y(\lambda)}$ is a free $\mathbf{k}$-module of finite rank. Thus, $\mathcal{A}\mathbf{E}_T$ is a free $\mathbf{k}$-module of finite rank as well (since $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$).

Hence, Corollary 5.17.18 (applied to $G = S_n$, $V = \mathcal{S}^{Y(\lambda)}$ and $W = \mathcal{A}\mathbf{E}_T$) yields $\mathbf{X}_{\mathcal{S}^{Y(\lambda)}} = \mathbf{X}_{\mathcal{A}\mathbf{E}_T}$ (since $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$ as representations of $S_n$). However, the definition of $\mathbf{X}_\lambda$ shows that $\mathbf{X}_\lambda = \mathbf{X}_{\mathcal{S}^\lambda} = \mathbf{X}_{\mathcal{S}^{Y(\lambda)}}$ (since $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)}$). Thus,

$$\mathbf{X}_\lambda = \mathbf{X}_{\mathcal{S}^{Y(\lambda)}} = \mathbf{X}_{\mathcal{A}\mathbf{E}_T}. \tag{344}$$

On the other hand, Theorem 5.11.3 yields $\mathbf{E}_T^2 = \dfrac{n!}{f^\lambda}\mathbf{E}_T$. In other words, $\mathbf{E}_T^2 = h^\lambda \mathbf{E}_T$ (since Definition 5.12.11 **(b)** yields $\dfrac{n!}{f^\lambda} = h^\lambda$). Hence, Corollary 5.17.27 (applied to $\mathbf{a} = \mathbf{E}_T$ and $\kappa = h^\lambda$) yields that

$$h^\lambda \cdot \mathbf{X}_{\mathcal{A}\mathbf{E}_T} = \sum_{w \in S_n} w\mathbf{E}_T w^{-1}.$$

On the other hand, Proposition 5.12.5 (applied to $P = T$) shows that

$$\mathbf{E}_\lambda = \sum_{w \in S_n} w \mathbf{E}_T w^{-1}.$$

Comparing these two equalities, we obtain $\mathbf{E}_\lambda = h^\lambda \cdot \underbrace{\mathbf{X}_{\mathcal{A}\mathbf{E}_T}}_{\substack{=\mathbf{X}_\lambda \\ \text{(by (344))}}} = h^\lambda \mathbf{X}_\lambda$. This proves Theorem 5.17.28. □

**Corollary 5.17.29.** Let $\lambda$ be a partition of $n$. Then, all coefficients of the element $\mathbf{E}_\lambda \in \mathbf{k}[S_n]$ are multiples of $h^\lambda$ (inside $\mathbf{k}$).

*Proof.* Theorem 5.17.28 yields that $\mathbf{E}_\lambda = h^\lambda \mathbf{X}_\lambda$ (where $\mathbf{X}_\lambda$ is as defined in Theorem 5.17.28). Thus, each coefficient of $\mathbf{E}_\lambda$ is $h^\lambda$ times the corresponding coefficient of $\mathbf{X}_\lambda$. Since the latter coefficient belongs to $\mathbf{k}$, we thus conclude that the former coefficient is a multiple of $h^\lambda$. This proves Corollary 5.17.29. □

**Remark 5.17.30.** Let $\lambda$ be a partition of $n$. Then, [Stanle23, Chapter 7, supplementary problem 101] shows that there exists a permutation $w \in S_n$ such that $\chi_{\mathcal{S}^\lambda}(w) = \pm 1$. Consider this $w$. Thus, the element $\mathbf{X}_\lambda$ of $\mathbf{k}[S_n]$ (defined in Theorem 5.17.28) has the coefficient $\pm 1$ in front of $w$. Consequently, the element $\mathbf{E}_\lambda$ of $\mathbf{k}[S_n]$ has the coefficient $\pm h^\lambda$ in front of $w$ (because Theorem 5.17.28 yields that $\mathbf{E}_\lambda = h^\lambda \mathbf{X}_\lambda$). Combined with Corollary 5.17.29, this easily yields that $h^\lambda$ is the greatest common divisor of all coefficients of $\mathbf{E}_\lambda$. (The notion of "greatest common divisor" here should be understood appropriately – after all, $\mathbf{k}$ is not necessarily a unique factorization domain; but we can take it literally when $\mathbf{k} = \mathbb{Z}$.)

If $h^\lambda$ is invertible in $\mathbf{k}$, then the equality $\mathbf{E}_\lambda = h^\lambda \mathbf{X}_\lambda$ in Theorem 5.17.28 can be solved for $\mathbf{X}_\lambda$, yielding $\mathbf{X}_\lambda = \dfrac{1}{h^\lambda} \mathbf{E}_\lambda$. This provides a good way of computing $\mathbf{X}_\lambda$ and thus the character $\chi_{\mathcal{S}^\lambda}$.

Using Theorem 5.17.28, we can furthermore derive some properties of the elements $\mathbf{X}_\lambda$ from analogous properties of the $\mathbf{E}_\lambda$:

**Corollary 5.17.31.** For each partition $\lambda$ of $n$, we let $\mathbf{X}_\lambda$ denote the element $\mathbf{X}_{\mathcal{S}^\lambda}$ corresponding to the Specht module $\mathcal{S}^\lambda$. Then:

**(a)** We have $\mathbf{X}_\lambda^2 = h^\lambda \mathbf{X}_\lambda$ for each partition $\lambda$ of $n$.

**(b)** We have $\mathbf{X}_\lambda \mathbf{X}_\mu = 0$ whenever $\lambda$ and $\mu$ are two distinct partitions of $n$.

**(c)** For each $\mathbf{a} \in Z(\mathbf{k}[S_n])$, we have

$$n! \cdot \mathbf{a} = \sum_{\lambda \text{ is a partition of } n} [1](\mathbf{X}_\lambda \mathbf{a}) \cdot \mathbf{X}_\lambda.$$

**(d)** If $n!$ is invertible in $\mathbf{k}$, then the family $(\mathbf{X}_\lambda)_{\lambda \text{ is a partition of } n}$ is a basis of the $\mathbf{k}$-module $Z(\mathbf{k}[S_n])$.

This corollary is easy to prove in the case when $n!$ is invertible in $\mathbf{k}$, because in this case we can use the above-mentioned equality $\mathbf{X}_\lambda = \dfrac{1}{h^\lambda}\mathbf{E}_\lambda$ to reduce everything to the properties of centralized Young symmetrizers (Theorem 5.12.14, Proposition 5.12.6, Proposition 5.12.22 and Theorem 5.12.23). Fortunately, the general case can be reduced to this easy case using a base change argument similar to the one we used back in the proof of Proposition 5.12.10. This reduction relies on the following two lemmas:

**Lemma 5.17.32.** Let $\lambda$ be a partition of $n$. Let us denote the Specht module $\mathcal{S}^\lambda$ (defined using the base ring $\mathbf{k}$) by $\mathcal{S}^\lambda_\mathbf{k}$ in order to stress its dependence on $\mathbf{k}$. Let $g \in S_n$. Then,

$$\chi_{\mathcal{S}^\lambda_\mathbf{k}}(g) = \chi_{\mathcal{S}^\lambda_\mathbb{Z}}(g) \cdot 1_\mathbf{k}.$$

*Proof.* Before I give a rigorous proof, let me explain why this lemma is obvious (even though the formal explanation of this obviousness will take us a couple boring pages). In a nutshell, the argument goes as follows: The Specht module $\mathcal{S}^\lambda$ has the same basis for any base ring $\mathbf{k}$ (namely, the basis described in Theorem 5.9.1). The matrix that represents the action of $g \in S_n$ with the respect to this basis is independent of $\mathbf{k}$ as well (more precisely, the entries of this matrix are the same integers for every $\mathbf{k}$, at least if you equate each integer $m \in \mathbb{Z}$ with the corresponding element $m \cdot 1_\mathbf{k}$ of $\mathbf{k}$). Thus, the trace of this matrix does not depend on $\mathbf{k}$ either. But this trace is precisely $\chi_{\mathcal{S}^\lambda_\mathbf{k}}(g)$ (by the definition of a character). Thus, $\chi_{\mathcal{S}^\lambda_\mathbf{k}}(g)$ is independent on $\mathbf{k}$. In particular, $\chi_{\mathcal{S}^\lambda_\mathbf{k}}(g)$ is the same as $\chi_{\mathcal{S}^\lambda_\mathbb{Z}}(g)$, or, to be more precise, we have $\chi_{\mathcal{S}^\lambda_\mathbf{k}}(g) = \chi_{\mathcal{S}^\lambda_\mathbb{Z}}(g) \cdot 1_\mathbf{k}$. Thus, Lemma 5.17.32 is proved, as long as you believe the above less-than-rigorous argument.

Let me now give a rigorous version of this proof.

Let $D$ be the Young diagram $Y(\lambda)$. Let $\mathrm{SYT}(D)$ be the set of all standard tableaux of shape $D$.

Let us denote the polytabloids $\mathbf{e}_T$ introduced in Definition 5.4.1 **(a)** by $\mathbf{e}_{T,\mathbf{k}}$ in order to stress their dependence on $\mathbf{k}$. Thus, $\mathbf{e}_T = \mathbf{e}_{T,\mathbf{k}}$ for any $n$-tableau $T$ of shape $D$.

From $D = Y(\lambda)$, we obtain $\mathcal{S}^D = \mathcal{S}^{Y(\lambda)} = \mathcal{S}^\lambda = \mathcal{S}^\lambda_\mathbf{k}$ (since we are denoting $\mathcal{S}^\lambda$ by $\mathcal{S}^\lambda_\mathbf{k}$). Moreover, $D$ is a skew Young diagram (since $D = Y(\lambda) = Y(\lambda/\varnothing)$). Thus, Theorem 5.9.1 shows that the standard polytabloids $\mathbf{e}_T$ form a basis of the $\mathbf{k}$-module $\mathcal{S}^D$. In other words, the family $(\mathbf{e}_T)_{T \in \mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{S}^D$ (since the standard polytabloids $\mathbf{e}_T$ are precisely the $\mathbf{e}_T$ for $T \in \mathrm{SYT}(D)$). In other words, the family $(\mathbf{e}_{T,\mathbf{k}})_{T \in \mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{S}^\lambda_\mathbf{k}$ (since $\mathcal{S}^D = \mathcal{S}^\lambda_\mathbf{k}$, and since $\mathbf{e}_T = \mathbf{e}_{T,\mathbf{k}}$ for any $n$-tableau $T$ of shape $D$). The same argument (applied to $\mathbb{Z}$ instead of $\mathbf{k}$) shows that the family $(\mathbf{e}_{T,\mathbb{Z}})_{T \in \mathrm{SYT}(D)}$ is a basis of the $\mathbb{Z}$-module $\mathcal{S}^\lambda_\mathbb{Z}$.

In particular, this family spans this $\mathbb{Z}$-module $\mathcal{S}^\lambda_\mathbb{Z}$. Hence, for each $P \in \mathrm{SYT}(D)$, the element $g\mathbf{e}_{P,\mathbb{Z}}$ of $\mathcal{S}^\lambda_\mathbb{Z}$ can be written as a $\mathbb{Z}$-linear combination of this family:

$$g\mathbf{e}_{P,\mathbb{Z}} = \sum_{T \in \mathrm{SYT}(D)} a_{T,P}\mathbf{e}_{T,\mathbb{Z}} \tag{345}$$

for some scalars $a_{T,P} \in \mathbb{Z}$. Consider these scalars $a_{T,P}$. We now claim the following:

*Claim 1:* We have

$$\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g) = \sum_{T \in \mathrm{SYT}(D)} a_{T,T}.$$

*Proof of Claim 1.* Let $\tau_{g,\mathbb{Z}} : \mathcal{S}^\lambda_{\mathbb{Z}} \to \mathcal{S}^\lambda_{\mathbb{Z}}$ be the $\mathbb{Z}$-linear map $v \mapsto gv$. Then, $\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g) = \mathrm{Tr}\left(\tau_{g,\mathbb{Z}}\right)$ (by the definition of a character). However, the definition of $\tau_{g,\mathbb{Z}}$ shows that each $P \in \mathrm{SYT}(D)$ satisfies

$$\tau_{g,\mathbb{Z}}\left(\mathbf{e}_{P,\mathbb{Z}}\right) = g\mathbf{e}_{P,\mathbb{Z}} = \sum_{T \in \mathrm{SYT}(D)} a_{T,P}\mathbf{e}_{T,\mathbb{Z}} \qquad \text{(by (345))}.$$

Thus, the matrix that represents the $\mathbb{Z}$-linear map $\tau_{g,\mathbb{Z}} : \mathcal{S}^\lambda_{\mathbb{Z}} \to \mathcal{S}^\lambda_{\mathbb{Z}}$ with respect to the basis $\left(\mathbf{e}_{T,\mathbb{Z}}\right)_{T \in \mathrm{SYT}(D)}$ of the $\mathbb{Z}$-module $\mathcal{S}^\lambda_{\mathbb{Z}}$ is the matrix[119] $\left(a_{T,P}\right)_{T,P \in \mathrm{SYT}(D)}$. Hence,

$$\mathrm{Tr}\left(\tau_{g,\mathbb{Z}}\right) = \mathrm{Tr}\left(\left(a_{T,P}\right)_{T,P \in \mathrm{SYT}(D)}\right) = \sum_{T \in \mathrm{SYT}(D)} a_{T,T}$$

(by the definition of the trace of a matrix). In view of $\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g) = \mathrm{Tr}\left(\tau_{g,\mathbb{Z}}\right)$, we can rewrite this as $\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g) = \sum_{T \in \mathrm{SYT}(D)} a_{T,T}$. This proves Claim 1. □

Let us denote the Young module $\mathcal{M}^D$ (defined using the base ring $\mathbf{k}$) by $\mathcal{M}^D_{\mathbf{k}}$ in order to stress its dependence on $\mathbf{k}$. Note that $\mathcal{S}^\lambda_{\mathbf{k}} = \mathcal{S}^D \subseteq \mathcal{M}^D = \mathcal{M}^D_{\mathbf{k}}$ and similarly $\mathcal{S}^\lambda_{\mathbb{Z}} \subseteq \mathcal{M}^D_{\mathbb{Z}}$.

Consider the canonical ring morphism $f : \mathbb{Z} \to \mathbf{k}$, which sends each $m \in \mathbb{Z}$ to $m \cdot 1_{\mathbf{k}} \in \mathbf{k}$. This morphism canonically induces a $\mathbb{Z}$-linear map $f_D : \mathcal{M}^D_{\mathbb{Z}} \to \mathcal{M}^D_{\mathbf{k}}$ which sends each element $\sum_{\substack{\overline{T} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} c_{\overline{T}}\overline{T} \in \mathcal{M}^D_{\mathbb{Z}}$ (where the coefficients $c_{\overline{T}}$ belong to $\mathbb{Z}$) to $\sum_{\substack{\overline{T} \text{ is an } n\text{-tabloid} \\ \text{of shape } D}} f\left(c_{\overline{T}}\right)\overline{T} \in \mathcal{M}^D_{\mathbf{k}}$. It is clear that every $n$-tableau $T$ of shape $D$ satisfies

$$f_D\left(\mathbf{e}_{T,\mathbb{Z}}\right) = \mathbf{e}_{T,\mathbf{k}} \tag{346}$$

(since the definition of a polytabloid says that $\mathbf{e}_T = \sum_{w \in \mathcal{C}(T)} (-1)^w \overline{wT}$, independently of $\mathbf{k}$). It is also clear that the map $f_D$ is $S_n$-equivariant (since $S_n$ acts on $\mathcal{M}^D_{\mathbb{Z}}$ and on $\mathcal{M}^D_{\mathbf{k}}$ in the same way, namely by permuting the $n$-tabloids). Using these two observations, we can easily send the equality (345) "over into $\mathcal{S}^\lambda_{\mathbf{k}}$":

*Claim 2:* For each $P \in \mathrm{SYT}(D)$, we have

$$g\mathbf{e}_{P,\mathbf{k}} = \sum_{T \in \mathrm{SYT}(D)} f\left(a_{T,P}\right)\mathbf{e}_{T,\mathbf{k}}.$$

[119] This is a $\mathrm{SYT}(D) \times \mathrm{SYT}(D)$-matrix, i.e., a matrix whose rows and columns are indexed by the elements of $\mathrm{SYT}(D)$ rather than by integers. The trace of such a matrix can be defined in the same way as the trace of a "usual" $m \times m$-matrix: It is the sum of its diagonal entries, i.e., of all its entries whose row index and column index are equal.

*Proof of Claim 2.* Let $P \in \operatorname{SYT}(D)$. Since the map $f_D$ is $S_n$-equivariant, we have

$$f_D\left(g\mathbf{e}_{P,\mathbb{Z}}\right) = g \cdot \underbrace{f_D\left(\mathbf{e}_{P,\mathbb{Z}}\right)}_{=\mathbf{e}_{P,\mathbf{k}} \text{ (by (346), applied to } T=P)} = g\mathbf{e}_{P,\mathbf{k}}.$$

Hence,

$$\begin{aligned} g\mathbf{e}_{P,\mathbf{k}} &= f_D\left(g\mathbf{e}_{P,\mathbb{Z}}\right) = f_D\left(\sum_{T\in\operatorname{SYT}(D)} a_{T,P}\mathbf{e}_{T,\mathbb{Z}}\right) \qquad \text{(by (345))} \\ &= \sum_{T\in\operatorname{SYT}(D)} a_{T,P} \underbrace{f_D\left(\mathbf{e}_{T,\mathbb{Z}}\right)}_{=\mathbf{e}_{T,\mathbf{k}} \text{ (by (346))}} \qquad \text{(since the map } f_D \text{ is } \mathbb{Z}\text{-linear)} \\ &= \sum_{T\in\operatorname{SYT}(D)} \underbrace{a_{T,P}\mathbf{e}_{T,\mathbf{k}}}_{=(a_{T,P}\cdot 1_{\mathbf{k}})\mathbf{e}_{T,\mathbf{k}}} = \sum_{T\in\operatorname{SYT}(D)} \underbrace{\left(a_{T,P}\cdot 1_{\mathbf{k}}\right)}_{=f(a_{T,P})} \mathbf{e}_{T,\mathbf{k}} \\ &\qquad \text{(since the definition of } f \text{ yields } f(a_{T,P})=a_{T,P}\cdot 1_{\mathbf{k}}) \\ &= \sum_{T\in\operatorname{SYT}(D)} f\left(a_{T,P}\right)\mathbf{e}_{T,\mathbf{k}}. \end{aligned}$$

This proves Claim 2. □

As a consequence of Claim 2, we can see the following analogue of Claim 1:

*Claim 3:* We have $\chi_{\mathcal{S}_{\mathbf{k}}^{\lambda}}(g) = \sum_{T\in\operatorname{SYT}(D)} f\left(a_{T,T}\right)$.

*Proof of Claim 3.* Let $\tau_{g,\mathbf{k}} : \mathcal{S}_{\mathbf{k}}^{\lambda} \to \mathcal{S}_{\mathbf{k}}^{\lambda}$ be the $\mathbf{k}$-linear map $v \mapsto gv$. Then, $\chi_{\mathcal{S}_{\mathbf{k}}^{\lambda}}(g) = \operatorname{Tr}\left(\tau_{g,\mathbf{k}}\right)$ (by the definition of a character). However, the definition of $\tau_{g,\mathbf{k}}$ shows that each $P \in \operatorname{SYT}(D)$ satisfies

$$\tau_{g,\mathbf{k}}\left(\mathbf{e}_{P,\mathbf{k}}\right) = g\mathbf{e}_{P,\mathbf{k}} = \sum_{T\in\operatorname{SYT}(D)} f\left(a_{T,P}\right)\mathbf{e}_{T,\mathbf{k}} \qquad \text{(by Claim 2)}.$$

Thus, the matrix that represents the $\mathbf{k}$-linear map $\tau_{g,\mathbf{k}} : \mathcal{S}_{\mathbf{k}}^{\lambda} \to \mathcal{S}_{\mathbf{k}}^{\lambda}$ with respect to the basis $\left(\mathbf{e}_{T,\mathbf{k}}\right)_{T\in\operatorname{SYT}(D)}$ of the $\mathbf{k}$-module $\mathcal{S}_{\mathbf{k}}^{\lambda}$ is the matrix $\left(f\left(a_{T,P}\right)\right)_{T,P\in\operatorname{SYT}(D)}$. Hence,

$$\operatorname{Tr}\left(\tau_{g,\mathbf{k}}\right) = \operatorname{Tr}\left(\left(f\left(a_{T,P}\right)\right)_{T,P\in\operatorname{SYT}(D)}\right) = \sum_{T\in\operatorname{SYT}(D)} f\left(a_{T,T}\right)$$

(by the definition of the trace of a matrix). In view of $\chi_{\mathcal{S}_{\mathbf{k}}^{\lambda}}(g) = \operatorname{Tr}\left(\tau_{g,\mathbf{k}}\right)$, we can rewrite this as $\chi_{\mathcal{S}_{\mathbf{k}}^{\lambda}}(g) = \sum_{T\in\operatorname{SYT}(D)} f\left(a_{T,T}\right)$. This proves Claim 3. □

Now, Claim 3 yields

$$\begin{aligned} \chi_{\mathcal{S}_{\mathbf{k}}^{\lambda}}(g) &= \sum_{T\in\operatorname{SYT}(D)} \underbrace{f\left(a_{T,T}\right)}_{=a_{T,T}\cdot 1_{\mathbf{k}} \text{ (by the definition of } f)} = \sum_{T\in\operatorname{SYT}(D)} a_{T,T}\cdot 1_{\mathbf{k}} = \underbrace{\left(\sum_{T\in\operatorname{SYT}(D)} a_{T,T}\right)}_{=\chi_{\mathcal{S}_{\mathbb{Z}}^{\lambda}}(g) \text{ (by Claim 1)}} \cdot 1_{\mathbf{k}} \\ &= \chi_{\mathcal{S}_{\mathbb{Z}}^{\lambda}}(g)\cdot 1_{\mathbf{k}}. \end{aligned}$$

This proves Lemma 5.17.32 (rigorously this time). □

**Lemma 5.17.33.** Let $\lambda$ be a partition of $n$. We let $\mathbf{X}_{\lambda,\mathbf{k}}$ denote the element $\mathbf{X}_{\mathcal{S}^\lambda}$ corresponding to the Specht module $\mathcal{S}^\lambda$.

Consider the canonical ring morphism $f : \mathbb{Z} \to \mathbf{k}$ and the base change morphism $f_* : \mathbb{Z}[S_n] \to \mathbf{k}[S_n]$ corresponding to $f$. Then,

$$f_*(\mathbf{X}_{\lambda,\mathbb{Z}}) = \mathbf{X}_{\lambda,\mathbf{k}}.$$

*Proof.* Let us denote the Specht module $\mathcal{S}^\lambda$ (defined using the base ring $\mathbf{k}$) by $\mathcal{S}^\lambda_{\mathbf{k}}$ in order to stress its dependence on $\mathbf{k}$. Thus, $\mathcal{S}^\lambda = \mathcal{S}^\lambda_{\mathbf{k}}$.

Recall that $\mathbf{X}_{\lambda,\mathbf{k}}$ is a notation for $\mathbf{X}_{\mathcal{S}^\lambda}$. Thus,

$$\begin{aligned}\mathbf{X}_{\lambda,\mathbf{k}} = \mathbf{X}_{\mathcal{S}^\lambda} &= \sum_{g\in S_n} \chi_{\mathcal{S}^\lambda}(g)\, g \qquad \text{(by Definition 5.17.15)} \\ &= \sum_{g\in S_n} \chi_{\mathcal{S}^\lambda_{\mathbf{k}}}(g)\, g \qquad \left(\text{since } \mathcal{S}^\lambda = \mathcal{S}^\lambda_{\mathbf{k}}\right). \end{aligned} \tag{347}$$

The same argument (applied to $\mathbb{Z}$ instead of $\mathbf{k}$) shows that

$$\mathbf{X}_{\lambda,\mathbb{Z}} = \sum_{g\in S_n} \chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g)\, g.$$

Applying the map $f_*$ to both sides of this equality, we obtain

$$\begin{aligned} f_*(\mathbf{X}_{\lambda,\mathbb{Z}}) &= f_*\left(\sum_{g\in S_n} \chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g)\, g\right) \\ &= \sum_{g\in S_n} \underbrace{f\left(\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g)\right)}_{\substack{=\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g)\cdot 1_{\mathbf{k}} \\ \text{(by the definition of } f\text{)}}} g \qquad \text{(by the definition of } f_*\text{)} \\ &= \sum_{g\in S_n} \underbrace{\left(\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g)\cdot 1_{\mathbf{k}}\right)}_{\substack{=\chi_{\mathcal{S}^\lambda_{\mathbf{k}}}(g) \\ \text{(by Lemma 5.17.32)}}} g = \sum_{g\in S_n} \chi_{\mathcal{S}^\lambda_{\mathbf{k}}}(g)\, g = \mathbf{X}_{\lambda,\mathbf{k}} \end{aligned}$$

(by (347)). This proves Lemma 5.17.33. □

*Proof of Corollary 5.17.31.* We shall prove part **(a)** in detail, and leave the rest as an exercise (Exercise 5.17.3).

**(a)** Let $\lambda$ be a partition of $n$. We must prove that $\mathbf{X}_\lambda^2 = h^\lambda \mathbf{X}_\lambda$.

We shall first prove this for $\mathbf{k} = \mathbb{Q}$, and then extend it to the general case using a base change argument similar to the one we used back in the proof of Proposition 5.12.10 above.

*Proof in the case when* $\mathbf{k} = \mathbb{Q}$*:* Assume that $\mathbf{k} = \mathbb{Q}$. Theorem 5.12.14 says that $\mathbf{E}_\lambda^2 = \left(h^\lambda\right)^2 \mathbf{E}_\lambda$. But Theorem 5.17.28 yields that $\mathbf{E}_\lambda = h^\lambda \mathbf{X}_\lambda$. In view of this, we can rewrite $\mathbf{E}_\lambda^2 = \left(h^\lambda\right)^2 \mathbf{E}_\lambda$ as

$$\left(h^\lambda \mathbf{X}_\lambda\right)^2 = \left(h^\lambda\right)^2 h^\lambda \mathbf{X}_\lambda.$$

Thus,

$$\left(h^\lambda\right)^2 \mathbf{X}_\lambda^2 = \left(h^\lambda \mathbf{X}_\lambda\right)^2 = \left(h^\lambda\right)^2 h^\lambda \mathbf{X}_\lambda.$$

Dividing both sides of this equality by $\left(h^\lambda\right)^2$ (this is allowed, since $h^\lambda$ is a positive integer and since $\mathbf{k} = \mathbb{Q}$), we obtain

$$\mathbf{X}_\lambda^2 = h^\lambda \mathbf{X}_\lambda.$$

Thus, we have proved $\mathbf{X}_\lambda^2 = h^\lambda \mathbf{X}_\lambda$ in the case when $\mathbf{k} = \mathbb{Q}$.

*General discussion:* Now, forget our assumption that $\mathbf{k} = \mathbb{Q}$. Instead, let $\mathbf{k}$ be an arbitrary commutative ring again.

Let us denote the element $\mathbf{X}_\lambda$ of $\mathbf{k}[S_n]$ by $\mathbf{X}_{\lambda,\mathbf{k}}$ in order to stress its dependence on $\mathbf{k}$. Thus, $\mathbf{X}_{\lambda,\mathbf{k}}$ is the element $\mathbf{X}_{\mathcal{S}^\lambda}$ corresponding to the Specht module $\mathcal{S}^\lambda$.

Above, we have proved $\mathbf{X}_\lambda^2 = h^\lambda \mathbf{X}_\lambda$ in the case when $\mathbf{k} = \mathbb{Q}$. In other words, we have proved that

$$\mathbf{X}_{\lambda,\mathbb{Q}}^2 = h^\lambda \mathbf{X}_{\lambda,\mathbb{Q}}. \tag{348}$$

*Proof in the case when* $\mathbf{k} = \mathbb{Z}$*:* Next, let us prove that $\mathbf{X}_{\lambda,\mathbb{Z}}^2 = h^\lambda \mathbf{X}_{\lambda,\mathbb{Z}}$.

Indeed, consider the canonical ring morphism $f : \mathbb{Z} \to \mathbb{Q}$ and the base change morphism $f_* : \mathbb{Z}[S_n] \to \mathbb{Q}[S_n]$ corresponding to $f$. Then, Lemma 5.17.33 (applied to $\mathbb{Q}$ instead of $\mathbf{k}$) says that

$$f_*\left(\mathbf{X}_{\lambda,\mathbb{Z}}\right) = \mathbf{X}_{\lambda,\mathbb{Q}}.$$

The morphism $f$ is injective (indeed, it is simply the embedding of $\mathbb{Z}$ into $\mathbb{Q}$). Thus, the morphism $f_*$ is injective as well. But $f_*$ is furthermore a ring morphism; thus,

$$\begin{aligned} f_*\left(\mathbf{X}_{\lambda,\mathbb{Z}}^2\right) &= \left(\underbrace{f_*\left(\mathbf{X}_{\lambda,\mathbb{Z}}\right)}_{=\mathbf{X}_{\lambda,\mathbb{Q}}}\right)^2 = \mathbf{X}_{\lambda,\mathbb{Q}}^2 = h^\lambda \underbrace{\mathbf{X}_{\lambda,\mathbb{Q}}}_{=f_*\left(\mathbf{X}_{\lambda,\mathbb{Z}}\right)} \qquad \text{(by (348))} \\ &= h^\lambda f_*\left(\mathbf{X}_{\lambda,\mathbb{Z}}\right) = f_*\left(h^\lambda \mathbf{X}_{\lambda,\mathbb{Z}}\right) \qquad \text{(since } f_* \text{ is a ring morphism)}. \end{aligned}$$

Since $f_*$ is injective, we can "un-apply" $f_*$ to this equality, and obtain

$$\mathbf{X}_{\lambda,\mathbb{Z}}^2 = h^\lambda \mathbf{X}_{\lambda,\mathbb{Z}}. \tag{349}$$

Thus, we have proved $\mathbf{X}_\lambda^2 = h^\lambda \mathbf{X}_\lambda$ in the case when $\mathbf{k} = \mathbb{Z}$.

*Proof in the general case:* We are now ready to prove $\mathbf{X}_\lambda^2 = h^\lambda \mathbf{X}_\lambda$ for an arbitrary base ring $\mathbf{k}$.

Indeed, forget the morphisms $f$ and $f_*$ constructed above. Instead, let us consider the canonical ring morphism $f : \mathbb{Z} \to \mathbf{k}$ and the base change morphism $f_* : \mathbb{Z}[S_n] \to \mathbf{k}[S_n]$ corresponding to $f$. Then, Lemma 5.17.33 says that

$$f_*(\mathbf{X}_{\lambda,\mathbb{Z}}) = \mathbf{X}_{\lambda,\mathbf{k}}.$$

But $f_*$ is a ring morphism. Thus,

$$f_*\left(\mathbf{X}_{\lambda,\mathbb{Z}}^2\right) = \left(\underbrace{f_*(\mathbf{X}_{\lambda,\mathbb{Z}})}_{=\mathbf{X}_{\lambda,\mathbf{k}}}\right)^2 = \mathbf{X}_{\lambda,\mathbf{k}}^2.$$

Hence,

$$\begin{aligned}\mathbf{X}_{\lambda,\mathbf{k}}^2 &= f_*\left(\underbrace{\mathbf{X}_{\lambda,\mathbb{Z}}^2}_{\substack{=h^\lambda \mathbf{X}_{\lambda,\mathbb{Z}} \\ \text{(by (349))}}}\right) = f_*\left(h^\lambda \mathbf{X}_{\lambda,\mathbb{Z}}\right) = h^\lambda \underbrace{f_*(\mathbf{X}_{\lambda,\mathbb{Z}})}_{=\mathbf{X}_{\lambda,\mathbf{k}}} \qquad \left(\begin{array}{c}\text{since } f_* \text{ is a} \\ \text{ring morphism}\end{array}\right) \\ &= h^\lambda \mathbf{X}_{\lambda,\mathbf{k}}.\end{aligned}$$

Since $\mathbf{X}_{\lambda,\mathbf{k}} = \mathbf{X}_\lambda$, we can rewrite this as $\mathbf{X}_\lambda^2 = h^\lambda \mathbf{X}_\lambda$. Thus, we have proved $\mathbf{X}_\lambda^2 = h^\lambda \mathbf{X}_\lambda$ in the general case (i.e., for all commutative rings $\mathbf{k}$). This proves Corollary 5.17.31 **(a)**. □

**Exercise 5.17.3.** 1+2+1 Prove parts **(b)**, **(c)** and **(d)** of Corollary 5.17.31.

### 5.17.8. Characters of small symmetric groups

Theorem 5.17.28 provides an easy way to compute the elements $\mathbf{X}_\lambda$ when $\mathbf{k} = \mathbb{Q}$. Using Lemma 5.17.33, we can then extend these formulas for $\mathbf{X}_\lambda$ to $\mathbf{k} = \mathbb{Z}$ and thence to arbitrary rings $\mathbf{k}$. Then we can compute all character values $\chi_{\mathcal{S}^\lambda}(g)$ by taking appropriate coefficients of $\mathbf{X}_\lambda$.

We shall now give tables of these character values for the first five symmetric groups $S_0, S_1, S_2, S_3, S_4$. The tables should be read as follows:

- The rows of our tables correspond to the Specht modules $\mathcal{S}^\lambda$ indexed by the partitions $\lambda$ of $n$. We denote the character $\chi_{\mathcal{S}^\lambda}$ of a Specht module $\mathcal{S}^\lambda$ by $\chi_\lambda$. Note that $\chi_{(n)} = \chi_{\text{triv}}$ (see Example 5.17.7) and $\chi_{(1^n)} = \chi_{\text{sign}}$ (see Example 5.17.8).

- If two permutations $u, v \in S_n$ have the same cycle type, then they are conjugate (by Theorem 3.9.9) and thus satisfy $\chi_\lambda(u) = \chi_\lambda(v)$ for each partition $\lambda$ of $n$ (since Corollary 5.17.6 shows that a character of a representation sends conjugate elements to the same value). Hence, the value $\chi_\lambda(g)$ for a permutation $g \in S_n$ depends only on the cycle type $\operatorname{type} g$ of $g$, but not on $g$ itself. Thus, the columns of our tables correspond to the possible cycle types that a permutation $g \in S_n$ can have; we don't need a separate column for each permutation. Note that the possible cycle types are the partitions of $n$. Thus, our tables are of square shapes, with both rows and columns indexed by the partitions of $n$.

Here come the tables themselves:

| $\operatorname{type} g$ | $()$ |
|---|---|
| $\chi_{()}(g)$ | 1 |

$n=0$

| $\operatorname{type} g$ | $(1)$ |
|---|---|
| $\chi_{(1)}(g)$ | 1 |

$n=1$

| $\operatorname{type} g$ | $(2)$ | $(1,1)$ |
|---|---|---|
| $\chi_{(2)}(g)$ | 1 | 1 |
| $\chi_{(1,1)}(g)$ | $-1$ | 1 |

$n=2$

| $\operatorname{type} g$ | $(3)$ | $(2,1)$ | $(1,1,1)$ |
|---|---|---|---|
| $\chi_{(3)}(g)$ | 1 | 1 | 1 |
| $\chi_{(2,1)}(g)$ | $-1$ | 0 | 2 |
| $\chi_{(1,1,1)}(g)$ | 1 | $-1$ | 1 |

$n=3$

| $\operatorname{type} g$ | $(4)$ | $(3,1)$ | $(2,2)$ | $(2,1,1)$ | $(1,1,1,1)$ |
|---|---|---|---|---|---|
| $\chi_{(4)}(g)$ | 1 | 1 | 1 | 1 | 1 |
| $\chi_{(3,1)}(g)$ | $-1$ | 0 | $-1$ | 1 | 3 |
| $\chi_{(2,2)}(g)$ | 0 | $-1$ | 2 | 0 | 2 |
| $\chi_{(2,1,1)}(g)$ | 1 | 0 | $-1$ | $-1$ | 3 |
| $\chi_{(1,1,1,1)}(g)$ | $-1$ | 1 | 1 | $-1$ | 1 |

$n=4$

.

Character tables of $S_5, S_6, \ldots, S_{10}$ can be found in [JamKer81, §I.A]. Further character tables can be calculated on a computer, e.g., using SageMath[120].

[120] For instance, `SymmetricGroup(8).character_table()` returns the character table of $S_8$, and similarly for any $S_n$. Sadly, I don't know in which order the rows and the columns are

**Exercise 5.17.4.** Let $\lambda$ be a partition of $n$. Let us denote the character $\chi_{\mathcal{S}^\lambda}$ of the Specht module $\mathcal{S}^\lambda$ by $\chi_\lambda$.

**(a)** $\boxed{2}$ Prove that $\sum_{g \in S_n} \chi_\lambda(g) = \begin{cases} n!, & \text{if } \lambda = (n); \\ 0, & \text{if } \lambda \neq (n). \end{cases}$

**(b)** $\boxed{3}$ Prove that $\sum_{g \in S_n} (\chi_\lambda(g))^2 = n!$.

**(c)** $\boxed{?}$ Prove that $\sum_{g \in S_n} \chi_\lambda(g^2) = n!$.

[**Hint:** For part **(a)**, show that $\nabla \mathbf{X}_\lambda = 0$ for $\lambda \neq (n)$ first. What does this have to do with part **(a)**? Likewise, what does part **(b)** have to do with $\mathbf{X}_\lambda^2 = h^\lambda \mathbf{X}_\lambda$ ?

I don't know a proof of part **(c)** that does not use any deeper representation theory (at least multilinear algebra). I would welcome any suggestions!]

### 5.17.9. An explicit formula for the character of $\mathcal{S}^\lambda$

The following combinatorial formula gives a faster way to compute the character $\chi_{\mathcal{S}^\lambda}$ of a straight-shaped Specht module $\mathcal{S}^\lambda$ than Theorem 5.17.28.

**Theorem 5.17.34.** Let $\lambda$ be a partition of $n$. Consider the character $\chi_{\mathcal{S}^\lambda}$ of the corresponding Specht module $\mathcal{S}^\lambda$. Let $g \in S_n$ be any permutation. Let $C$ be the conjugacy class of $S_n$ that contains $g$. Let $T$ be any $n$-tableau of shape $\lambda$. Then,

$$\chi_{\mathcal{S}^\lambda}(g) = \underbrace{\frac{f^\lambda}{|C|} \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} (-1)^c \right)}_{\text{This is an integer.}} \cdot 1_{\mathbf{k}}.$$

(See Definition 5.12.11 **(a)** for the meaning of $f^\lambda$.)

*Proof.* Define the integer

$$k := \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} (-1)^c. \tag{350}$$

printed in this table (I think it is in reverse lexicographic order of the corresponding partitions; this would be the reverse of the order of my tables above).

The easiest way to obtain a given value $\chi_\lambda(g)$ in SageMath is probably using symmetric functions: The code

```
Sym = SymmetricFunctions(QQ); p = Sym.p(); s = Sym.s()
s[Partition([3,2])].scalar(p[Partition([2,1,1,1])])
```

returns the value $\chi_{(3,2)}(g)$ for a permutation $g \in S_5$ with cycle type type $g = (2,1,1,1)$. Similarly for other partitions and other permutations. The theory that underlies this method is that of the Frobenius characteristic map (see, e.g., [Sagan01, §4.7] or [Stanle23, §7.18]).

We start with the following observation:

*Claim 1:* We have $\chi_{\mathcal{S}^\lambda}(c) = \chi_{\mathcal{S}^\lambda}(g)$ for each $c \in C$.

*Proof of Claim 1.* Let $c \in C$. Then, $c$ and $g$ belong to the same conjugacy class of $S_n$ (namely, to the class $C$). In other words, the elements $c$ and $g$ are conjugate in $S_n$. In other words, $c = wgw^{-1}$ for some $w \in S_n$. Consider this $w$. From $c = wgw^{-1}$, we obtain $\chi_{\mathcal{S}^\lambda}(c) = \chi_{\mathcal{S}^\lambda}\left(wgw^{-1}\right) = \chi_{\mathcal{S}^\lambda}(g)$ (by Corollary 5.17.6, applied to $S_n$ and $\mathcal{S}^\lambda$ instead of $G$ and $V$). This proves Claim 1. □

Now, let $\mathbf{z}_C$ denote the conjugacy class sum $\sum_{c \in C} c \in \mathbf{k}[S_n]$ (as in Definition 3.9.4 **(c)**). Then, each $\mathbf{v} \in \mathcal{S}^\lambda$ satisfies

$$\mathbf{z}_C \mathbf{v} = \underbrace{\left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} (-1)^c \right)}_{\substack{=k \\ \text{(by (350))}}} \mathbf{v} \qquad \text{(by Theorem 5.16.1)}$$

$$= k\mathbf{v}. \tag{351}$$

Thus, we can easily see the following:

*Claim 2:* We have $\chi_{\mathcal{S}^\lambda}(\mathbf{z}_C) = f^\lambda k \cdot 1_{\mathbf{k}}$.

*Proof of Claim 2.* The $\mathbf{k}$-module $\mathcal{S}^\lambda$ is free of rank $f^\lambda$ (by Lemma 5.9.17).

For each $r \in \mathbf{k}[S_n]$, we let $\tau_r : \mathcal{S}^\lambda \to \mathcal{S}^\lambda$ be the $\mathbf{k}$-linear map $v \mapsto rv$. The definition of the character $\chi_{\mathcal{S}^\lambda}$ then shows that $\chi_{\mathcal{S}^\lambda}(r) = \operatorname{Tr}(\tau_r)$ for each $r \in \mathbf{k}[S_n]$. Thus, in particular, $\chi_{\mathcal{S}^\lambda}(\mathbf{z}_C) = \operatorname{Tr}(\tau_{\mathbf{z}_C})$.

However, from (351), we easily obtain $\tau_{\mathbf{z}_C} = (k \cdot 1_{\mathbf{k}}) \operatorname{id}_{\mathcal{S}^\lambda}$ [121]. Hence,

$$\begin{aligned}
\operatorname{Tr}(\tau_{\mathbf{z}_C}) &= \operatorname{Tr}\left((k \cdot 1_{\mathbf{k}}) \operatorname{id}_{\mathcal{S}^\lambda}\right) \\
&= f^\lambda (k \cdot 1_{\mathbf{k}}) \qquad \left( \begin{array}{c} \text{by Lemma 5.11.10,} \\ \text{applied to } W = \mathcal{S}^\lambda \text{ and } d = f^\lambda \text{ and } \kappa = k \cdot 1_{\mathbf{k}} \\ \text{(since } \mathcal{S}^\lambda \text{ is a free } \mathbf{k}\text{-module of rank } f^\lambda) \end{array} \right) \\
&= f^\lambda k \cdot 1_{\mathbf{k}}.
\end{aligned}$$

In view of $\chi_{\mathcal{S}^\lambda}(\mathbf{z}_C) = \operatorname{Tr}(\tau_{\mathbf{z}_C})$, we can rewrite this as $\chi_{\mathcal{S}^\lambda}(\mathbf{z}_C) = f^\lambda k \cdot 1_{\mathbf{k}}$. Thus, Claim 2 is proved. □

[121] *Proof.* For each $\mathbf{v} \in \mathcal{S}^\lambda$, we have

$$\begin{aligned}
\tau_{\mathbf{z}_C}(\mathbf{v}) &= \mathbf{z}_C \mathbf{v} \qquad \text{(by the definition of } \tau_{\mathbf{z}_C}) \\
&= k\mathbf{v} \qquad \text{(by (351))} \\
&= (k \cdot 1_{\mathbf{k}}) \mathbf{v} = \left((k \cdot 1_{\mathbf{k}}) \operatorname{id}_{\mathcal{S}^\lambda}\right)(\mathbf{v}).
\end{aligned}$$

In other words, $\tau_{\mathbf{z}_C} = (k \cdot 1_{\mathbf{k}}) \operatorname{id}_{\mathcal{S}^\lambda}$.

Combining Claim 1 with Claim 2, we obtain:

*Claim 3:* We have $|C| \cdot \chi_{\mathcal{S}^\lambda}(g) = f^\lambda k \cdot 1_{\mathbf{k}}$.

*Proof of Claim 3.* The character $\chi_{\mathcal{S}^\lambda} : \mathbf{k}[S_n] \to \mathbf{k}$ is a $\mathbf{k}$-linear map (by Proposition 5.17.2). Now, from $\mathbf{z}_C = \sum_{c \in C} c$, we obtain

$$\chi_{\mathcal{S}^\lambda}(\mathbf{z}_C) = \chi_{\mathcal{S}^\lambda}\left(\sum_{c \in C} c\right) = \sum_{c \in C} \underbrace{\chi_{\mathcal{S}^\lambda}(c)}_{\substack{=\chi_{\mathcal{S}^\lambda}(g) \\ \text{(by Claim 1)}}} \qquad \text{(since } \chi_{\mathcal{S}^\lambda} \text{ is a } \mathbf{k}\text{-linear map)}$$
$$= \sum_{c \in C} \chi_{\mathcal{S}^\lambda}(g) = |C| \cdot \chi_{\mathcal{S}^\lambda}(g).$$

Thus, $|C| \cdot \chi_{\mathcal{S}^\lambda}(g) = \chi_{\mathcal{S}^\lambda}(\mathbf{z}_C) = f^\lambda k \cdot 1_{\mathbf{k}}$ (by Claim 2). This proves Claim 3. □

From Claim 3, we aim to recover $\chi_{\mathcal{S}^\lambda}(g)$. Unfortunately, we cannot divide by $|C|$ when $\mathbf{k}$ is just an arbitrary commutative ring. Thus, we make a short detour via the ring $\mathbb{Z}$ and a base change argument.

Let us denote the Specht module $\mathcal{S}^\lambda$ (defined using the base ring $\mathbf{k}$) by $\mathcal{S}^\lambda_{\mathbf{k}}$ in order to stress its dependence on $\mathbf{k}$. Now, we claim the following:

*Claim 4:* We have $\dfrac{f^\lambda k}{|C|} \in \mathbb{Z}$ and $\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g) = \dfrac{f^\lambda k}{|C|}$.

*Proof of Claim 4.* Applying Claim 3 to $\mathbb{Z}$ instead of $\mathbf{k}$, we find $|C| \cdot \chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g) = f^\lambda k \cdot 1_{\mathbb{Z}} = f^\lambda k$. Since $|C|$ is a positive integer (because the conjugacy class $C$ is clearly a nonempty finite set), we can divide this equality by $|C|$, and find $\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g) = \dfrac{f^\lambda k}{|C|}$ in $\mathbb{Q}$. Thus, $\dfrac{f^\lambda k}{|C|} = \chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g) \in \mathbb{Z}$ (since $\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}$ is a map from $\mathbb{Z}[S_n]$ to $\mathbb{Z}$). Thus, Claim 4 is proved. □

We now return to $\mathbf{k}$:

*Claim 5:* We have $\chi_{\mathcal{S}^\lambda}(g) = \dfrac{f^\lambda k}{|C|} \cdot 1_{\mathbf{k}}$.

*Proof of Claim 5.* Lemma 5.17.32 yields

$$\chi_{\mathcal{S}^\lambda_{\mathbf{k}}}(g) = \underbrace{\chi_{\mathcal{S}^\lambda_{\mathbb{Z}}}(g)}_{\substack{=\dfrac{f^\lambda k}{|C|} \\ \text{(by Claim 4)}}} \cdot 1_{\mathbf{k}} = \frac{f^\lambda k}{|C|} \cdot 1_{\mathbf{k}}.$$

Since $\mathcal{S}^\lambda_{\mathbf{k}} = \mathcal{S}^\lambda$, we can rewrite this as $\chi_{\mathcal{S}^\lambda}(g) = \dfrac{f^\lambda k}{|C|} \cdot 1_{\mathbf{k}}$. Thus, Claim 5 is proved. □

Now, Claim 5 yields

$$\chi_{\mathcal{S}^\lambda}(g) = \frac{f^\lambda k}{|C|} \cdot 1_{\mathbf{k}} = \frac{f^\lambda}{|C|} k \cdot 1_{\mathbf{k}} = \frac{f^\lambda}{|C|} \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} (-1)^c \right) \cdot 1_{\mathbf{k}}$$

(by (350)). Moreover, $\underbrace{\frac{f^\lambda}{|C|} \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} (-1)^c \right)}_{=k} = \frac{f^\lambda}{|C|} k = \frac{f^\lambda k}{|C|} \in \mathbb{Z}$ (by Claim 4). In other words, $\frac{f^\lambda}{|C|} \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} (-1)^c \right)$ is an integer. Thus, Theorem 5.17.34 is proved. □

Specifically for transpositions, we obtain the following simple formula:

**Corollary 5.17.35.** Let $\lambda$ be a partition of $n$. Consider the character $\chi_{\mathcal{S}^\lambda}$ of the corresponding Specht module $\mathcal{S}^\lambda$. Let $g \in S_n$ be any transposition. Then,

$$\chi_{\mathcal{S}^\lambda}(g) = \underbrace{\frac{f^\lambda}{\binom{n}{2}} \left( \mathrm{n}\left(\lambda^t\right) - \mathrm{n}(\lambda) \right)}_{\text{This is an integer.}} \cdot 1_{\mathbf{k}}.$$

(See Definition 5.12.11 **(a)** for the meaning of $f^\lambda$, and Definition 5.16.2 for the meaning of $\mathrm{n}(\lambda)$.)

*Proof.* Pick any $n$-tableau $T$ of shape $Y(\lambda)$ (such a $T$ exists, since $|Y(\lambda)| = |\lambda| = n$). Thus, $T$ is an $n$-tableau of shape $\lambda$.

Let $C$ be the set of all transpositions in $S_n$. Then, $C$ is a conjugacy class of $S_n$ (since we know that all transpositions in $S_n$ form a single conjugacy class), and contains $g$ (since $g$ is a transposition). Thus, $C$ is the conjugacy class of $S_n$ that contains $g$. Theorem 5.17.34 thus yields

$$\chi_{\mathcal{S}^\lambda}(g) = \underbrace{\frac{f^\lambda}{|C|} \left( \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ cr \in C}} (-1)^c \right)}_{\text{This is an integer.}} \cdot 1_{\mathbf{k}}. \tag{352}$$

However, we have $|C| = \binom{n}{2}$ [122] and

$$\sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ cr\in C}} (-1)^c = \mathrm{n}\left(\lambda^t\right) - \mathrm{n}\left(\lambda\right) \qquad \text{(by Lemma 5.16.9)}.$$

In view of this, we can rewrite (352) as

$$\chi_{S^\lambda}(g) = \underbrace{\frac{f^\lambda}{\binom{n}{2}}\left(\mathrm{n}\left(\lambda^t\right) - \mathrm{n}\left(\lambda\right)\right)}_{\text{This is an integer.}} \cdot 1_{\mathbf{k}}.$$

Thus, Corollary 5.17.35 is proved. □

### 5.17.10. An outlook on the Murnaghan–Nakayama rule

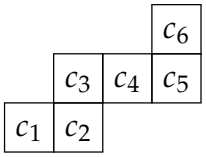
Let us now briefly note a different – combinatorial – approach to computing the characters of Specht modules. This is the famous *Murnaghan–Nakayama rule*. It relies on the concept of a *rim hook*:

**Definition 5.17.36.** A *rim hook* means a nonempty diagram $D$ that can be written in the form

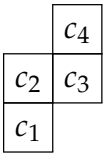
$$D = \{c_1, c_2, \ldots, c_k\},$$

where $c_1, c_2, \ldots, c_k \in \mathbb{Z}^2$ are cells such that for each $i \in \{2, 3, \ldots, k\}$, the cell $c_i$ is either the northern neighbor or the eastern neighbor of $c_{i-1}$.

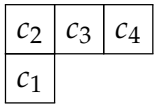
For example, the following diagrams are rim hooks:



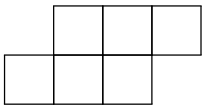

and



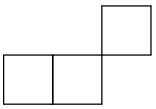

and



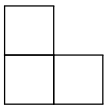

(where we have labelled the cells by $c_1, c_2, \ldots, c_k$ in a way that makes the requirement in Definition 5.17.36 true). In contrast, neither of the diagrams



and



and



[122] *Proof.* The transpositions in $S_n$ are the permutations of the form $t_{i,j}$, where $i$ and $j$ are two distinct elements of $[n]$. Moreover, the order of $i$ and $j$ does not matter (i.e., we have $t_{i,j} = t_{j,i}$). There are $\binom{n}{2}$ ways to choose two distinct elements $i$ and $j$ of $[n]$ if we disregard the order; and each of these $\binom{n}{2}$ choices produces a different transposition $t_{i,j}$. Thus, there are exactly $\binom{n}{2}$ transpositions in $S_n$. In other words, $|C| = \binom{n}{2}$ (since $C$ is the set of all transpositions in $S_n$).

is a rim hook. Rim hooks are also known as *ribbons* or *border strips* or *skew hooks*.

**Exercise 5.17.5.** [2] Prove that every rim hook is a skew Young diagram.

**Exercise 5.17.6.** [1] Let $\lambda$ be a partition. Prove that the diagram $Y(\lambda)$ is a rim hook if and only if $\lambda$ is a hook partition.

**Exercise 5.17.7.** [1] Recall the map $\mathbf{r} : \mathbb{Z}^2 \to \mathbb{Z}^2$ defined in Theorem 5.1.10. Let $D$ be a diagram. Prove that $D$ is a rim hook if and only if $\mathbf{r}(D)$ is a rim hook.

The following exercises give several equivalent characterizations of rim hooks:

**Exercise 5.17.8.** [2] Define a map $h : \mathbb{Z}^2 \to \mathbb{Z}$ by setting

$$h(i,j) = j - i \qquad \text{for each } (i,j) \in \mathbb{Z}^2.$$

Let $\lambda/\mu$ be a skew partition. Prove that the diagram $Y(\lambda/\mu)$ is a rim hook if and only if the restriction of the map $h$ to $Y(\lambda/\mu)$ is injective and its image $h(Y(\lambda/\mu))$ is the integer interval $\{u, u+1, \ldots, v\}$ for some integers $u \leq v$.

**Exercise 5.17.9.** [2] Let $\lambda/\mu$ be a skew partition. Let $G$ be the undirected graph whose vertices are the cells in $Y(\lambda/\mu)$, and whose edges are defined as follows: Two vertices $c$ and $d$ are adjacent if and only if the cells $c$ and $d$ are neighbors (i.e., the Euclidean distance between $c$ and $d$ is 1).

Prove that the diagram $Y(\lambda/\mu)$ is a rim hook if and only if the graph $G$ is connected and the diagram $Y(\lambda/\mu)$ contains no $2 \times 2$-square (i.e., there exists no $(i,j) \in Y(\lambda/\mu)$ such that $(i,j+1)$ and $(i+1,j)$ and $(i+1,j+1)$ all belong to $Y(\lambda/\mu)$ as well).

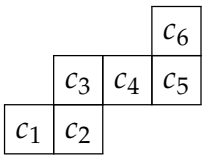
**Exercise 5.17.10.** [4] Let $\lambda/\mu$ be a skew partition. Prove that the diagram $Y(\lambda/\mu)$ is a rim hook if and only if there exist two positive integers $u < v$ such that

$$(\lambda_i = \mu_{i-1} + 1 \text{ for each } i \in \{u+1, u+2, \ldots, v\}) \qquad \text{and}$$
$$(\lambda_i = \mu_i \text{ for each } i \notin \{u, u+1, \ldots, v\}).$$

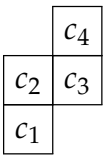
We next define the *leg length* of a rim hook:

**Definition 5.17.37.** Let $D$ be a rim hook. Then, the *leg length* of $D$ is defined to be $|\operatorname{Rows} D| - 1$, where $\operatorname{Rows} D := \{i \mid (i,j) \in D\}$.

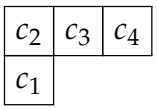
For instance, the rim hooks




and




and




have leg lengths 2 and 2 and 1, respectively.

Next, we define *rim hook tableaux*:

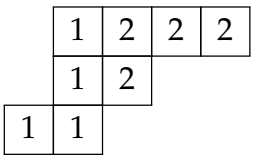
**Definition 5.17.38.** Let $D$ be a diagram. Let $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$ be a tuple of nonnegative integers.

**(a)** A *rim hook tableau* of shape $D$ and content $\mu$ means a tableau $T$ of shape $D$ with the following properties:

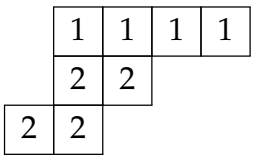
- All entries of $T$ belong to the set $[k]$.
- The tableau $T$ is row-semistandard and column-semistandard.
- For each $i \in [k]$, the set of all cells of $T$ that contain the entry $i$ is a rim hook of size $\mu_i$. We shall denote this rim hook by $R_i(T)$.

**(b)** If $T$ is such a rim hook tableau, then the *leg length* of $T$ means the sum of the leg lengths of the rim hooks $R_1(T)$, $R_2(T)$, $\ldots$, $R_k(T)$.

**Example 5.17.39.** Let $D$ be the skew Young diagram $Y((5,3,2)/(1,1))$. Then, the only rim hook tableaux of shape $D$ and content $(4,4)$ are




and




.

Their respective leg lengths are $2+1=3$ and $0+1=1$.

Note that a rim hook tableau of shape $D$ and content $(1,1,\ldots,1)$ is the same thing as a standard tableau of shape $D$ (assuming that the number of 1's in the content is the size of $D$).

We can finally state the *Murnaghan–Nakayama rule*:

**Theorem 5.17.40** (Murnaghan–Nakayama rule)**.** Let $D$ be a skew Young diagram of size $|D| = n$. Let $\chi_D$ be the character of the corresponding Specht module $\mathcal{S}^D$. Let $g \in S_n$ be a permutation, and let $\mu = \operatorname{type} g$ be its cycle type (or, more generally, any tuple of nonnegative integers such that the nonzero entries of $\mu$ are the lengths of the cycles of $g$, not necessarily in decreasing

order). Then,

$$\chi_D(g) = \sum_{\substack{T \text{ is a rim hook tableau} \\ \text{of shape } D \text{ and content } \mu}} (-1)^{(\text{leg length of } T)}.$$

We will not prove this theorem here, as it would require some further character theory that we have not introduced. A proof can be found in [CSScTo10, Remark 3.5.18] or in [Kleshc05, Corollary 2.3.7]. (In the case of a straight Young diagram, Theorem 5.17.40 also appears, e.g., in [Stanle23, Theorem 7.18.5], [Sagan01, Corollary 4.10.6], [Meliot17, Theorem 3.10]. Various other texts contain recursive rules that are easily seen to be equivalent.)

We observe a simple particular case of Theorem 5.17.40:

**Corollary 5.17.41.** Let $D$ be a skew Young diagram of size $|D| = n$. Let $\chi_D$ be the character of the corresponding Specht module $\mathcal{S}^D$. Let $g \in S_n$ be an $n$-cycle. Then,

$$\chi_D(g) = \begin{cases} (-1)^{(\text{leg length of } D)}, & \text{if } D \text{ is a rim hook;} \\ 0, & \text{otherwise.} \end{cases}$$

**Exercise 5.17.11. (a)** [1] Derive Corollary 5.17.41 from Theorem 5.17.40.

**(b)** [?] Prove Corollary 5.17.41 independently.

### 5.17.11. Characters uniquely determine $S_n$-representations

Let us next prove another useful fact, which is often applied to prove that two $S_n$-representations are isomorphic without constructing an explicit isomorphism:

**Proposition 5.17.42.** Assume that $\mathbf{k}$ is a field of characteristic 0. Let $V$ and $W$ be two finite-dimensional representations of $S_n$. Then, $V \cong W$ if and only if $\chi_V = \chi_W$.

This proposition is a particular case of Theorem 5.17.14, but let us outline an independent proof:

*Proof of Proposition 5.17.42 (sketched).* The "only if" part of Proposition 5.17.42 follows from Proposition 5.17.13. Thus, it remains to prove the "if" part. So let us assume that $\chi_V = \chi_W$. We must prove that $V \cong W$.

From $\chi_V = \chi_W$, we obtain $\mathbf{X}_V = \mathbf{X}_W$ (just as in the proof of Corollary 5.17.18).

For each partition $\lambda$ of $n$, we let $\mathbf{X}_\lambda$ denote the element $\mathbf{X}_{\mathcal{S}^\lambda}$ corresponding to the Specht module $\mathcal{S}^\lambda$.

Let $P$ be the set of all partitions of $n$. Then, the family $(\mathbf{X}_\lambda)_{\lambda\in P} = (\mathbf{X}_\lambda)_{\lambda \text{ is a partition of } n}$ is a basis of the $\mathbf{k}$-module $Z(\mathbf{k}[S_n])$ (by Corollary 5.17.31 **(d)**). Hence, this family is $\mathbf{k}$-linearly independent.

Proposition 5.17.42 shows that

$$V \cong \bigoplus_{\lambda\in P} \left(\mathcal{S}^\lambda\right)^{i_\lambda} \qquad \text{for some integers } i_\lambda \in \mathbb{N}.$$

Likewise, Proposition 5.17.42 (applied to $W$ instead of $V$) shows that

$$W \cong \bigoplus_{\lambda\in P} \left(\mathcal{S}^\lambda\right)^{j_\lambda} \qquad \text{for some integers } j_\lambda \in \mathbb{N}.$$

Consider these integers $i_\lambda$ and $j_\lambda$.

Corollary 5.17.23 shows that the character of a direct sum of several $S_n$-representations is the sum of the characters of these representations. Thus, from $V \cong \bigoplus_{\lambda\in P} \left(\mathcal{S}^\lambda\right)^{i_\lambda}$, we obtain $\chi_V = \sum_{\lambda\in P} i_\lambda \chi_{\mathcal{S}^\lambda}$ (since isomorphic $S_n$-representations have equal characters). Therefore, for each $g \in S_n$, we have

$$\chi_V(g) = \left(\sum_{\lambda\in P} i_\lambda \chi_{\mathcal{S}^\lambda}\right)(g) = \sum_{\lambda\in P} i_\lambda \chi_{\mathcal{S}^\lambda}(g). \tag{353}$$

Now, Definition 5.17.15 yields

$$\begin{aligned}
\mathbf{X}_V &= \sum_{g\in S_n} \underbrace{\chi_V(g)}_{=\sum_{\lambda\in P} i_\lambda \chi_{\mathcal{S}^\lambda}(g)} g = \sum_{g\in S_n} \left(\sum_{\lambda\in P} i_\lambda \chi_{\mathcal{S}^\lambda}(g)\right) g \\
&\qquad \text{(by (353))} \\
&= \sum_{\lambda\in P} i_\lambda \underbrace{\sum_{g\in S_n} \chi_{\mathcal{S}^\lambda}(g)\, g}_{=\mathbf{X}_{\mathcal{S}^\lambda}} = \sum_{\lambda\in P} i_\lambda \underbrace{\mathbf{X}_{\mathcal{S}^\lambda}}_{=\mathbf{X}_\lambda} = \sum_{\lambda\in P} i_\lambda \mathbf{X}_\lambda.
\end{aligned}$$

(by Definition 5.17.15)

Likewise,

$$\mathbf{X}_W = \sum_{\lambda\in P} j_\lambda \mathbf{X}_\lambda.$$

The left hand sides of these two equalities are equal (since $\mathbf{X}_V = \mathbf{X}_W$). Thus, so are the right hand sides. In other words,

$$\sum_{\lambda\in P} i_\lambda \mathbf{X}_\lambda = \sum_{\lambda\in P} j_\lambda \mathbf{X}_\lambda.$$

Since the family $(\mathbf{X}_\lambda)_{\lambda\in P}$ is $\mathbf{k}$-linearly independent, this equality entails that all $\lambda \in P$ satisfy $i_\lambda = j_\lambda$ in $\mathbf{k}$. Since $\mathbb{Z}$ is a subring of $\mathbf{k}$ (because $\mathbf{k}$ is a field of

characteristic 0), this entails that all $\lambda \in P$ satisfy $i_\lambda = j_\lambda$ in $\mathbb{Z}$ as well. Now, recall that

$$V \cong \bigoplus_{\lambda \in P} \left(\mathcal{S}^\lambda\right)^{i_\lambda} = \bigoplus_{\lambda \in P} \left(\mathcal{S}^\lambda\right)^{j_\lambda} \qquad \text{(since all } \lambda \in P \text{ satisfy } i_\lambda = j_\lambda\text{)}$$
$$\cong W \qquad \left(\text{since } W \cong \bigoplus_{\lambda \in P} \left(\mathcal{S}^\lambda\right)^{j_\lambda}\right).$$

This proves the "if" part of Proposition 5.17.42. The proof of this proposition is thus complete. □

**Remark 5.17.43.** Note that Proposition 5.17.42 really requires $\mathbf{k}$ to be a field of characteristic 0. Merely requiring $\mathbf{k}$ to be a field in which $n!$ is invertible would not suffice, since (e.g.) the representations $\mathbf{k}_{\text{triv}}^m = \underbrace{\mathbf{k}_{\text{triv}} \oplus \mathbf{k}_{\text{triv}} \oplus \cdots \oplus \mathbf{k}_{\text{triv}}}_{m \text{ times}}$ and 0 have the same character whenever $m = 0$ in $\mathbf{k}$.

## 5.18. The sign-twist of an $S_n$-representation

In this short section, we will explore a simple way to transform $S_n$-representations into new ones by "flipping a sign", and describe how it acts on Specht modules.

### 5.18.1. The sign-twist in general

In this section, we will oftentimes be dealing with two different $\mathbf{k}[S_n]$-module structures on one and the same set. Thus, it behooves us to have a notation that allows us to tell them apart. Here is the notation that we will use:

**Convention 5.18.1.** Let $R$ be a ring, and let $M$ be an $R$-module. Then, the action of $R$ on $M$ (that is, the map $R \times M \to M$ that sends each pair $(r, m)$ to $rm$) will be denoted $\overset{M}{\cdot}$. Thus, for any $r \in R$ and $m \in M$, we will write $r \overset{M}{\cdot} m$ for the image of $(r, m)$ under this action. Of course, this image has been hitherto denoted by $rm$, but the notation "$rm$" can be ambiguous when there are two different actions on the same set $M$, whereas the notation "$r \overset{M}{\cdot} m$" is unambiguous in this case, since it explicitly mentions the $R$-module $M$.

We could also introduce a similar notation for the addition on $M$, but we have no need for this, since we won't have to deal with differing additions on the same set $M$ here.

We shall now give two equivalent definitions of the sign-twist: one for representations of $S_n$ and one for left $\mathbf{k}[S_n]$-modules. As we know, a representation of $S_n$ (over $\mathbf{k}$) is "the same as" a left $\mathbf{k}[S_n]$-module (by Corollary 4.3.13, applied

to $G = S_n$), so that both definitions apply to the same objects; as we will soon see (Proposition 5.18.5 **(c)**), they also produce the same objects.

Here are the two definitions:

**Definition 5.18.2.** Let $V$ be a representation of $S_n$ over $\mathbf{k}$. Then, the *sign-twist* of $V$ is a new representation of $S_n$ over $\mathbf{k}$, which is called $V^{\mathrm{sign}}$ and is defined as follows:

- As a $\mathbf{k}$-module, $V^{\mathrm{sign}}$ is just $V$.
- The $S_n$-action on $V^{\mathrm{sign}}$ is the map $\overset{\mathrm{sign}}{\rightharpoonup} : S_n \times V \to V$ defined by the formula

$$g \overset{\mathrm{sign}}{\rightharpoonup} v := (-1)^g \, g \rightharpoonup v \qquad \text{for all } g \in S_n \text{ and } v \in V.$$

  Here, the "$\rightharpoonup$" symbol on the right hand side denotes the $S_n$-action on the original $S_n$-representation $V$.

This representation $V^{\mathrm{sign}}$ really exists (by Proposition 5.18.5 **(a)** below), so the sign-twist of $V$ is well-defined.

**Definition 5.18.3.** Let $V$ be a left $\mathbf{k}[S_n]$-module. Then, the *sign-twist* of $V$ is a new left $\mathbf{k}[S_n]$-module, which is called $V^{\mathrm{sign}}$ and is defined as follows:

- The addition and the zero of $V^{\mathrm{sign}}$ are the same as those of $V$. (In other words, $V^{\mathrm{sign}} = V$ as an additive group.)
- The action of $\mathbf{k}[S_n]$ on $V^{\mathrm{sign}}$ is given by[123]

$$\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v := T_{\mathrm{sign}}(\mathbf{a}) \overset{V}{\cdot} v \qquad \text{for all } \mathbf{a} \in \mathbf{k}[S_n] \text{ and } v \in V^{\mathrm{sign}}.$$

  In other words, the action of $\mathbf{k}[S_n]$ on $V^{\mathrm{sign}}$ is the map

$$\mathbf{k}[S_n] \times V \to V, \quad (\mathbf{a}, v) \mapsto T_{\mathrm{sign}}(\mathbf{a}) \overset{V}{\cdot} v.$$

This left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$ is well-defined (by Proposition 5.18.5 **(b)** below).

Soon (in Proposition 5.18.5) we will show that these two definitions are legitimate and equivalent. But first, let us give a simple example:

[123] We are using Convention 5.18.1 here.

**Example 5.18.4.** Consider the trivial representation $\mathbf{k}_{\mathrm{triv}}$ of $S_n$ constructed in Example 4.2.16. Let us find its sign-twist $(\mathbf{k}_{\mathrm{triv}})^{\mathrm{sign}}$ according to Definition 5.18.2. Indeed, as a $\mathbf{k}$-module, $(\mathbf{k}_{\mathrm{triv}})^{\mathrm{sign}} = \mathbf{k}_{\mathrm{triv}} = \mathbf{k}$, whereas the $S_n$-action on $(\mathbf{k}_{\mathrm{triv}})^{\mathrm{sign}}$ is the map $\overset{\mathrm{sign}}{\rightharpoonup} : S_n \times \mathbf{k} \to \mathbf{k}$ defined by the formula

$$g \overset{\mathrm{sign}}{\rightharpoonup} v := (-1)^g \, g \rightharpoonup v \qquad \text{for all } g \in S_n \text{ and } v \in \mathbf{k}_{\mathrm{triv}}.$$

This formula can be simplified to

$$g \overset{\mathrm{sign}}{\rightharpoonup} v := (-1)^g \, v \qquad \text{for all } g \in S_n \text{ and } v \in \mathbf{k}_{\mathrm{triv}}$$

(since $g \rightharpoonup v = v$ (because the $S_n$-action on $\mathbf{k}_{\mathrm{triv}}$ is trivial)). But this is precisely the left $S_n$-action on the sign representation $\mathbf{k}_{\mathrm{sign}}$ (as defined in Example 4.2.18). Hence, we conclude that $(\mathbf{k}_{\mathrm{triv}})^{\mathrm{sign}} = \mathbf{k}_{\mathrm{sign}}$.

More generally, the same arguments show that if $V$ is any trivial representation of $S_n$ (see Example 4.2.16), then its sign-twist $V^{\mathrm{sign}}$ is a sign representation of $S_n$ (see Example 4.2.18).

It is not much harder to see that $\left(\mathbf{k}_{\mathrm{sign}}\right)^{\mathrm{sign}} = \mathbf{k}_{\mathrm{triv}}$. (This also follows from $(\mathbf{k}_{\mathrm{triv}})^{\mathrm{sign}} = \mathbf{k}_{\mathrm{sign}}$ using Proposition 5.18.6.) We can easily find the sign-twist of the regular representation:

**Exercise 5.18.1.** $\boxed{1}$ Consider the left regular representation $\mathbf{k}[S_n]$ of $S_n$. Prove that its sign-twist $(\mathbf{k}[S_n])^{\mathrm{sign}}$ is isomorphic to $\mathbf{k}[S_n]$ itself, and that the map $T_{\mathrm{sign}} : \mathbf{k}[S_n] \to (\mathbf{k}[S_n])^{\mathrm{sign}}$ is an isomorphism of $S_n$-representations.

We will see other examples of sign-twists below.

We still owe a proof of the well-definedness of $V^{\mathrm{sign}}$, so let us give it now:

**Proposition 5.18.5.** Let $V$ be a left $\mathbf{k}[S_n]$-module, i.e., a representation of $S_n$ over $\mathbf{k}$. Then:

**(a)** The representation $V^{\mathrm{sign}}$ in Definition 5.18.2 is well-defined.

**(b)** The left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$ in Definition 5.18.3 is well-defined.

**(c)** The representation $V^{\mathrm{sign}}$ in Definition 5.18.2 and the left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$ in Definition 5.18.3 are the same (i.e., converting the former into a left $\mathbf{k}[S_n]$-module using Theorem 4.3.12 yields the latter).

*Proof.* We begin with part **(b)**, since this is the simplest part, and since establishing it will render the other two parts easier.

**(b)** Consider the map

$$\mathbf{k}[S_n] \times V \to V,$$
$$(\mathbf{a}, v) \mapsto T_{\mathrm{sign}}(\mathbf{a}) \overset{V}{\cdot} v.$$

This is the map that is intended to serve as the action of $\mathbf{k}[S_n]$ on $V^{\mathrm{sign}}$ (according to Definition 5.18.3). Hence, let us denote this map by $\overset{V^{\mathrm{sign}}}{\cdot}$ (that is, let us write $\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v$ for the image of a pair $(\mathbf{a}, v) \in \mathbf{k}[S_n] \times V$ under this map). Thus,

$$\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v = T_{\mathrm{sign}}(\mathbf{a}) \overset{V}{\cdot} v \tag{354}$$

for all $\mathbf{a} \in \mathbf{k}[S_n]$ and $v \in V$. The map $\overset{V^{\mathrm{sign}}}{\cdot}$ is clearly $\mathbf{k}$-bilinear (since the map $T_{\mathrm{sign}}$ is $\mathbf{k}$-linear, and since the action $\overset{V}{\cdot}$ of $\mathbf{k}[S_n]$ on $V$ is $\mathbf{k}$-bilinear).

In Definition 5.18.3, we defined the left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$ to be the additive group $V$, equipped with the map $\overset{V^{\mathrm{sign}}}{\cdot}$ as the action of $\mathbf{k}[S_n]$ on $V^{\mathrm{sign}}$. In order to prove that this left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$ is well-defined, we must show that this structure satisfies all the module axioms (see Definition 4.3.1). Most of these axioms are straightforward to check (e.g., both right and left distributivity laws follow from the $\mathbf{k}$-bilinearity of the map $\overset{V^{\mathrm{sign}}}{\cdot}$), so we only focus on two axioms: the associativity law and the "$1m = m$" axiom. Let us check them both:

- *The associativity law:* We must prove the associativity law for the left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$. In other words, we must prove that $(rs) \overset{V^{\mathrm{sign}}}{\cdot} m = r \overset{V^{\mathrm{sign}}}{\cdot} \left( s \overset{V^{\mathrm{sign}}}{\cdot} m \right)$ for all $r, s \in \mathbf{k}[S_n]$ and $m \in V$. So let $r, s \in \mathbf{k}[S_n]$ and $m \in V$ be arbitrary. Then, the definition of $\overset{V^{\mathrm{sign}}}{\cdot}$ yields $s \overset{V^{\mathrm{sign}}}{\cdot} m = T_{\mathrm{sign}}(s) \overset{V}{\cdot} m$ and $r \overset{V^{\mathrm{sign}}}{\cdot} \left( s \overset{V^{\mathrm{sign}}}{\cdot} m \right) = T_{\mathrm{sign}}(r) \overset{V}{\cdot} \left( s \overset{V^{\mathrm{sign}}}{\cdot} m \right)$ and $(rs) \overset{V^{\mathrm{sign}}}{\cdot} m = T_{\mathrm{sign}}(rs) \overset{V}{\cdot} m$. But $T_{\mathrm{sign}}$ is a $\mathbf{k}$-algebra morphism (by Theorem 3.11.5 **(a)**), and thus we have $T_{\mathrm{sign}}(rs) = T_{\mathrm{sign}}(r) \cdot T_{\mathrm{sign}}(s)$. Hence,

$$\begin{aligned}(rs) \overset{V^{\mathrm{sign}}}{\cdot} m &= \underbrace{T_{\mathrm{sign}}(rs)}_{= T_{\mathrm{sign}}(r) \cdot T_{\mathrm{sign}}(s)} \overset{V}{\cdot} m = \left( T_{\mathrm{sign}}(r) \cdot T_{\mathrm{sign}}(s) \right) \overset{V}{\cdot} m \\ &= T_{\mathrm{sign}}(r) \overset{V}{\cdot} \left( T_{\mathrm{sign}}(s) \overset{V}{\cdot} m \right)\end{aligned}$$

(since the original action $\overset{V}{\cdot}$ of $\mathbf{k}[S_n]$ on $V$ satisfies the associativity law). Comparing this with

$$r \overset{V^{\mathrm{sign}}}{\cdot} \left( s \overset{V^{\mathrm{sign}}}{\cdot} m \right) = T_{\mathrm{sign}}(r) \overset{V}{\cdot} \left( \underbrace{s \overset{V^{\mathrm{sign}}}{\cdot} m}_{= T_{\mathrm{sign}}(s) \overset{V}{\cdot} m} \right) = T_{\mathrm{sign}}(r) \overset{V}{\cdot} \left( T_{\mathrm{sign}}(s) \overset{V}{\cdot} m \right),$$

we obtain $(rs) \overset{V^{\mathrm{sign}}}{\cdot} m = r \overset{V^{\mathrm{sign}}}{\cdot} \left( s \overset{V^{\mathrm{sign}}}{\cdot} m \right)$. Thus, the associativity law for the $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$ is proved.

- *The $1m = m$ axiom:* We must prove the "$1m = m$" axiom for the left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$. In other words, we must prove that $1 \overset{V^{\mathrm{sign}}}{\cdot} m = m$ for all $m \in V$

(where 1 is the unity of $\mathbf{k}[S_n]$). So let $m \in V$. Then, the definition of $\overset{V^{\mathrm{sign}}}{\cdot}$ yields $1 \overset{V^{\mathrm{sign}}}{\cdot} m = T_{\mathrm{sign}}(1) \overset{V}{\cdot} m$. But $T_{\mathrm{sign}}$ is a $\mathbf{k}$-algebra morphism (by Theorem 3.11.5 **(a)**), and thus we have $T_{\mathrm{sign}}(1) = 1$. Hence, $1 \overset{V^{\mathrm{sign}}}{\cdot} m = \underbrace{T_{\mathrm{sign}}(1)}_{=1} \overset{V}{\cdot} m = 1 \overset{V}{\cdot} m = m$ (since the $\mathbf{k}[S_n]$-module $V$ satisfies the module axioms). Thus, we have proved the "$1m = m$" axiom for the left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$.

As we said, the other axioms are even easier to check. Thus, the left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$ is well-defined. This proves Proposition 5.18.5 **(b)**.

**(a)** Instead of proving the claim directly, we shall derive it from part **(b)** (which we have already proved).

First, we introduce some notations (exclusively for this proof):

- We shall write $V_1^{\mathrm{sign}}$ for the representation $V^{\mathrm{sign}}$ defined in Definition 5.18.2 (even though we have not yet shown that it is well-defined). Our goal is to prove that this representation $V_1^{\mathrm{sign}}$ is well-defined.
- We shall write $V_2^{\mathrm{sign}}$ for the left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$ defined in Definition 5.18.3. Proposition 5.18.5 **(b)** (which we have already proved) shows that the latter left $\mathbf{k}[S_n]$-module $V_2^{\mathrm{sign}}$ is well-defined. We reinterpret this left $\mathbf{k}[S_n]$-module $V_2^{\mathrm{sign}}$ as a representation of $S_n$ (using Corollary 4.3.13, as usual).

We shall now show that this representation $V_2^{\mathrm{sign}}$ is precisely the (purported) representation $V_1^{\mathrm{sign}}$. This will immediately entail that the latter representation is well-defined, and thus Proposition 5.18.5 **(a)** will follow (since $V_1^{\mathrm{sign}}$ is precisely the representation $V^{\mathrm{sign}}$ defined in Definition 5.18.2).

First, we recall that $T_{\mathrm{sign}}$ is a $\mathbf{k}$-algebra morphism (by Theorem 3.11.5 **(a)**), and thus we have $T_{\mathrm{sign}}\left(1_{\mathbf{k}[S_n]}\right) = 1_{\mathbf{k}[S_n]}$. But $T_{\mathrm{sign}}$ is $\mathbf{k}$-linear. Hence, for each $\lambda \in \mathbf{k}$, we have

$$T_{\mathrm{sign}}\left(\lambda 1_{\mathbf{k}[S_n]}\right) = \lambda \cdot \underbrace{T_{\mathrm{sign}}\left(1_{\mathbf{k}[S_n]}\right)}_{=1_{\mathbf{k}[S_n]}} = \lambda 1_{\mathbf{k}[S_n]}. \tag{355}$$

Next, we recall that $V_2^{\mathrm{sign}}$ is the left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$ from Definition 5.18.3. Hence, the action of $\mathbf{k}[S_n]$ on $V_2^{\mathrm{sign}}$ is given by

$$\mathbf{a} \overset{V_2^{\mathrm{sign}}}{\cdot} v = T_{\mathrm{sign}}(\mathbf{a}) \overset{V}{\cdot} v \tag{356}$$

for all $\mathbf{a} \in \mathbf{k}[S_n]$ and $v \in V$ (by Definition 5.18.3).

Now, we shall show that our $S_n$-representation $V_2^{\mathrm{sign}}$ (which we obtained from the left $\mathbf{k}[S_n]$-module $V_2^{\mathrm{sign}}$ defined in Definition 5.18.3) is the (purported) representation $V_1^{\mathrm{sign}}$ defined in Definition 5.18.2. Indeed:

- The underlying $\mathbf{k}$-module of our $S_n$-representation $V_2^{\mathrm{sign}}$ is the original $\mathbf{k}$-module $V$.

  [*Proof:* We must show that $V_2^{\mathrm{sign}} = V$ as $\mathbf{k}$-modules. By its definition, our $S_n$-representation $V_2^{\mathrm{sign}}$ has the same addition and zero as $V$. Thus, it remains to show that $V_2^{\mathrm{sign}}$ also has the same scaling (by elements of $\mathbf{k}$) as $V$. In other words, it remains to prove that every $\lambda \in \mathbf{k}$ and $v \in V$ satisfy $\lambda \overset{V_2^{\mathrm{sign}}}{\cdot} v = \lambda \overset{V}{\cdot} v$. So let us prove this. For any $\lambda \in \mathbf{k}$ and $v \in V$, we have

$$\begin{aligned}
\lambda \overset{V_2^{\mathrm{sign}}}{\cdot} v &= \left(\lambda 1_{\mathbf{k}[S_n]}\right) \overset{V_2^{\mathrm{sign}}}{\cdot} v && \left(\begin{array}{c}\text{since the } \mathbf{k}\text{-module structure on } V_2^{\mathrm{sign}} \text{ is} \\ \text{obtained from the } \mathbf{k}[S_n]\text{-module structure} \\ \text{by restriction of scalars}\end{array}\right) \\
&= \underbrace{T_{\mathrm{sign}}\left(\lambda 1_{\mathbf{k}[S_n]}\right)}_{\substack{=\lambda 1_{\mathbf{k}[S_n]} \\ \text{(by (355))}}} \overset{V}{\cdot} v && \left(\text{by (356), applied to } \mathbf{a} = \lambda 1_{\mathbf{k}[S_n]}\right) \\
&= \left(\lambda 1_{\mathbf{k}[S_n]}\right) \overset{V}{\cdot} v \\
&= \lambda \overset{V}{\cdot} \underbrace{\left(1_{\mathbf{k}[S_n]} \overset{V}{\cdot} v\right)}_{\substack{=v \\ \text{(by the module axioms)}}} && \left(\text{since the action } \overset{V}{\cdot} \text{ is } \mathbf{k}\text{-bilinear}\right) \\
&= \lambda \overset{V}{\cdot} v.
\end{aligned}$$

  Thus, the $\mathbf{k}$-module $V_2^{\mathrm{sign}}$ has the same scaling as $V$. Since $V_2^{\mathrm{sign}}$ also has the same addition and zero as $V$, we thus conclude that the $\mathbf{k}$-module $V_2^{\mathrm{sign}}$ is identical with the $\mathbf{k}$-module $V$.]

- The action of $S_n$ on our $S_n$-representation $V_2^{\mathrm{sign}}$ is exactly the map $\overset{\mathrm{sign}}{\rightharpoonup} : S_n \times V \to V$ defined in Definition 5.18.2.

  [*Proof:* Let $\underset{2}{\overset{\mathrm{sign}}{\rightharpoonup}} : S_n \times V_2^{\mathrm{sign}} \to V_2^{\mathrm{sign}}$ be the action of $S_n$ on the representation $V_2^{\mathrm{sign}}$ of $S_n$. Then, for any $g \in S_n$ and $v \in V_2^{\mathrm{sign}}$, we have

$$\begin{aligned}
g \underset{2}{\overset{\mathrm{sign}}{\rightharpoonup}} v = g \overset{V_2^{\mathrm{sign}}}{\cdot} v &= \underbrace{T_{\mathrm{sign}}(g)}_{\substack{=(-1)^g g \\ \text{(by the definition of } T_{\mathrm{sign}}\text{)}}} \overset{V}{\cdot} v && \text{(by (356), applied to } \mathbf{a} = g\text{)} \\
&= \left((-1)^g g\right) \overset{V}{\cdot} v = (-1)^g \underbrace{g \overset{V}{\cdot} v}_{\substack{=g \rightharpoonup v \\ \text{(where } \rightharpoonup \text{ denotes the action of } S_n \\ \text{on the original representation } V \text{ of } S_n\text{)}}} = (-1)^g\, g \rightharpoonup v.
\end{aligned}$$

  But this is precisely the formula for the map $\overset{\mathrm{sign}}{\rightharpoonup} : S_n \times V \to V$ defined in Definition 5.18.2. Thus, the map $\underset{2}{\overset{\mathrm{sign}}{\rightharpoonup}} : S_n \times V_2^{\mathrm{sign}} \to V_2^{\mathrm{sign}}$ is precisely the map $\overset{\mathrm{sign}}{\rightharpoonup} : S_n \times V \to V$ defined in Definition 5.18.2. In other words, the action of

$S_n$ on our $S_n$-representation $V_2^{\mathrm{sign}}$ is exactly the map $\overset{\mathrm{sign}}{\cdot} : S_n \times V \to V$ defined in Definition 5.18.2.]

Thus, we have shown that our representation $V_2^{\mathrm{sign}}$ of $S_n$ has the same underlying $\mathbf{k}$-module as $V$, and its action of $S_n$ is precisely the map $\overset{\mathrm{sign}}{\cdot} : S_n \times V \to V$ from Definition 5.18.2. Consequently, this representation $V_2^{\mathrm{sign}}$ is exactly the representation $V^{\mathrm{sign}}$ defined in Definition 5.18.2 (i.e., the representation that we now call $V_1^{\mathrm{sign}}$). Therefore, the latter representation $V^{\mathrm{sign}}$ is well-defined (since the former representation is well-defined). This proves Proposition 5.18.5 **(a)**.

**(c)** We must prove that the representation $V^{\mathrm{sign}}$ in Definition 5.18.2 and the left $\mathbf{k}[S_n]$-module $V^{\mathrm{sign}}$ in Definition 5.18.3 are the same. In our above proof of part **(a)**, we have denoted the former representation by $V_1^{\mathrm{sign}}$ and the latter $\mathbf{k}[S_n]$-module by $V_2^{\mathrm{sign}}$; thus, we must prove that $V_1^{\mathrm{sign}}$ and $V_2^{\mathrm{sign}}$ are the same.

But we have essentially done this already. Indeed, in our above proof of part **(a)**, we have shown that the representation $V_2^{\mathrm{sign}}$ is exactly the representation $V^{\mathrm{sign}}$ defined in Definition 5.18.2; but the latter representation is precisely the one that we called $V_1^{\mathrm{sign}}$. Hence, we have shown that $V_1^{\mathrm{sign}}$ and $V_2^{\mathrm{sign}}$ are the same. This proves Proposition 5.18.5 **(c)**. □

Proposition 5.18.5 **(a)** and **(b)** shows that our two definitions of $V^{\mathrm{sign}}$ are well-defined, and Proposition 5.18.5 **(c)** shows that they are equivalent. We now turn to some simple properties of sign-twists:

**Proposition 5.18.6.** Let $V$ be a representation of $S_n$. Then, $\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}} = V$. (This is saying that the $S_n$-representation $\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}$ is literally identical to $V$ – not just isomorphic.)

*Proof.* We shall use Definition 5.18.3 as a definition of $V^{\mathrm{sign}}$ and also of $\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}$. The three left $\mathbf{k}[S_n]$-modules $V$ and $V^{\mathrm{sign}}$ and $\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}$ all have the same underlying additive group (by Definition 5.18.3), but usually don't all have the same action of $\mathbf{k}[S_n]$. However, we claim that the first and the third of them (i.e., the left $\mathbf{k}[S_n]$-modules $V$ and $\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}$) do have the same action of $\mathbf{k}[S_n]$.

Indeed, the definition of the action of $\mathbf{k}[S_n]$ on $V^{\mathrm{sign}}$ (see Definition 5.18.3) shows that

$$\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v = T_{\mathrm{sign}}(\mathbf{a}) \overset{V}{\cdot} v \tag{357}$$

for all $\mathbf{a} \in \mathbf{k}[S_n]$ and $v \in V^{\mathrm{sign}}$. Applying the same reasoning to $V^{\mathrm{sign}}$ instead of $V$, we obtain

$$\mathbf{a} \overset{\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}}{\cdot} v = T_{\mathrm{sign}}(\mathbf{a}) \overset{V^{\mathrm{sign}}}{\cdot} v \tag{358}$$

for all $\mathbf{a} \in \mathbf{k}[S_n]$ and $v \in \left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}$. Now, for any $\mathbf{a} \in \mathbf{k}[S_n]$ and $v \in V$, we have

$v \in V = V^{\mathrm{sign}} = \left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}$ and therefore

$$\begin{aligned}
\mathbf{a} \overset{\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}}{\cdot} v &= T_{\mathrm{sign}}\left(\mathbf{a}\right) \overset{V^{\mathrm{sign}}}{\cdot} v \qquad \text{(by (358))} \\
&= \underbrace{T_{\mathrm{sign}}\left(T_{\mathrm{sign}}\left(\mathbf{a}\right)\right)}_{\substack{=\left(T_{\mathrm{sign}} \circ T_{\mathrm{sign}}\right)(\mathbf{a}) \\ =\mathbf{a} \\ \text{(since Theorem 3.11.5 \textbf{(b)}} \\ \text{says that } T_{\mathrm{sign}} \circ T_{\mathrm{sign}} = \mathrm{id}\text{)}}} \overset{V}{\cdot} v \qquad \left( \begin{array}{c} \text{by (357), applied to } T_{\mathrm{sign}}\left(\mathbf{a}\right) \\ \text{instead of } \mathbf{a} \end{array} \right) \\
&= \mathbf{a} \overset{V}{\cdot} v.
\end{aligned}$$

This shows that the maps $\overset{\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}}{\cdot}$ and $\overset{V}{\cdot}$ are equal. In other words, the actions of $\mathbf{k}\left[S_n\right]$ on the left $\mathbf{k}\left[S_n\right]$-modules $\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}$ and $V$ are equal (since these actions are $\overset{\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}}{\cdot}$ and $\overset{V}{\cdot}$, respectively).

Now, the left $\mathbf{k}\left[S_n\right]$-modules $\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}$ and $V$ have the same underlying additive group (as we saw above), but also have the same action of $\mathbf{k}\left[S_n\right]$ (since we just showed that the actions of $\mathbf{k}\left[S_n\right]$ on the left $\mathbf{k}\left[S_n\right]$-modules $\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}}$ and $V$ are equal). Hence, these two left $\mathbf{k}\left[S_n\right]$-modules are completely identical. In other words, $\left(V^{\mathrm{sign}}\right)^{\mathrm{sign}} = V$. This proves Proposition 5.18.6. □

We give two more basic properties of sign-twists without proof:

**Proposition 5.18.7.** Let $V$ and $W$ be two representations of $S_n$. Then, $\left(V \oplus W\right)^{\mathrm{sign}} = V^{\mathrm{sign}} \oplus W^{\mathrm{sign}}$.

**Proposition 5.18.8.** Let $V$ and $W$ be two representations of $S_n$. Let $f : V \to W$ be a map. Then:

**(a)** The map $f$ is a morphism of representations from $V$ to $W$ if and only if $f$ is a morphism of representations from $V^{\mathrm{sign}}$ to $W^{\mathrm{sign}}$.

**(b)** The map $f$ is an isomorphism of representations from $V$ to $W$ if and only if $f$ is an isomorphism of representations from $V^{\mathrm{sign}}$ to $W^{\mathrm{sign}}$.

**Corollary 5.18.9.** Let $V$ and $W$ be two representations of $S_n$. If $V \cong W$, then $V^{\mathrm{sign}} \cong W^{\mathrm{sign}}$.

**Exercise 5.18.2. (a)** $\boxed{1}$ Prove Proposition 5.18.7.

**(b)** $\boxed{2}$ Prove Proposition 5.18.8 and Corollary 5.18.9.

**Remark 5.18.10.** Proposition 5.18.8 **(a)** shows that we can define a "sign-twist functor" on the category of $S_n$-representations. This is a (covariant) functor from the category of $S_n$-representations to itself. On objects, it acts by sending each $S_n$-representation $V$ to $V^{\mathrm{sign}}$. On morphisms, it does nothing (i.e., it sends any morphism $f$ to $f$ itself).

**Remark 5.18.11.** The definition of the sign-twist (Definition 5.18.3) can be generalized: Let $R$ and $R'$ be two rings, and let $\varphi : R \to R'$ be any ring morphism. Then, any left $R'$-module $V$ can be turned into a left $R$-module by setting

$$r \cdot v := \varphi(r) \cdot v \qquad \text{for all } r \in R \text{ and } v \in V$$

(where the "$\cdot$" on the left hand side means the action of $R$ on $V$ that we are defining, whereas the "$\cdot$" on the right hand side means the existing action of $R'$ on $V$). The addition on this new $R$-module $V$ is the same as on the original $R'$-module $V$. We denote this new $R$-module $V$ by $V^{\varphi}$.

This way of transforming $R'$-modules $V$ into $R$-modules $V^{\varphi}$ (with the same ground set and the same addition) is known as *restriction of scalars*, and has already been introduced in Definition 4.3.5 in a slightly less general context (with the letters $\mathbf{k}$ and $R$ used instead of $R$ and $R'$). Nothing prevents us from taking $R = \mathbf{k}[S_n]$ and $R' = \mathbf{k}[S_n]$ here, and letting $\varphi$ be the ring morphism $T_{\mathrm{sign}} : \mathbf{k}[S_n] \to \mathbf{k}[S_n]$. If we do so, then $V^{\varphi} = V^{T_{\mathrm{sign}}}$ is precisely the sign-twist $V^{\mathrm{sign}}$ defined in Definition 5.18.3. Many properties of sign-twists are particular cases of general facts about restriction of scalars. For example, if $\varphi : R \to R'$ and $\psi : R' \to R''$ are two ring morphisms, and if $V$ is a left $R''$-module, then $\left(V^{\psi}\right)^{\varphi} = V^{\psi \circ \varphi}$. Applying this to $R = \mathbf{k}[S_n]$ and $R' = \mathbf{k}[S_n]$ and $R'' = \mathbf{k}[S_n]$ and $\varphi = T_{\mathrm{sign}}$ and $\psi = T_{\mathrm{sign}}$, we obtain $\left(V^{T_{\mathrm{sign}}}\right)^{T_{\mathrm{sign}}} = V^{T_{\mathrm{sign}} \circ T_{\mathrm{sign}}} = V^{\mathrm{id}}$ (since $T_{\mathrm{sign}} \circ T_{\mathrm{sign}} = \mathrm{id}$). Since $V^{\mathrm{id}}$ is easily seen to be just $V$, whereas $V^{T_{\mathrm{sign}}}$ is the sign-twist $V^{\mathrm{sign}}$, this is precisely the claim of Proposition 5.18.6.

**Remark 5.18.12.** There is yet another way of defining the sign-twist $V^{\mathrm{sign}}$ of an $S_n$-representation $V$: Namely, it can be viewed as an instance of the *internal tensor product*, namely the tensor product $\mathbf{k}_{\mathrm{sign}} \otimes V$. Personally, I find this definition unnecessarily convoluted (internal tensor products in general are rather complicated), but it is worth knowing about its existence, so here is an outline:

The *internal tensor product* of two $S_n$-representations $U$ and $V$ is defined to be the $S_n$-representation which (as a $\mathbf{k}$-module) is the tensor product $U \otimes V$ and whose $S_n$-action is given by the rule

$$g \rightharpoonup (u \otimes v) = (g \rightharpoonup u) \otimes (g \rightharpoonup v) \qquad \text{for all } g \in S_n \text{ and } u \in U \text{ and } v \in V.$$

(This rule determines the $S_n$-action on all pure tensors, and thus – by linearity – on all tensors.) It is not hard to see that this really is a well-defined $S_n$-action (indeed, this works just as well for any group $G$ instead of $S_n$), and that each $S_n$-representation $V$ satisfies $\mathbf{k}_{\mathrm{triv}} \otimes V \cong V$ (by the isomorphism $\lambda \otimes v \mapsto \lambda v$) and $\mathbf{k}_{\mathrm{sign}} \otimes V \cong V^{\mathrm{sign}}$ (by the isomorphism $\lambda \otimes v \mapsto \lambda v$ again). The latter isomorphism allows us to work with $\mathbf{k}_{\mathrm{sign}} \otimes V$ instead of $V^{\mathrm{sign}}$, thus sidestepping the need for defining $V^{\mathrm{sign}}$.

### 5.18.2. The sign-twist of a Specht module

We have seen some examples of sign-twists in Example 5.18.4 and shortly thereafter. Let us now state the main result of this section:

**Theorem 5.18.13.** Let $\lambda$ be any partition of $n$. Set $\mathcal{S}^\lambda := \mathcal{S}^{Y(\lambda)}$ (the Specht module for shape $Y(\lambda)$). Let $\lambda^t$ be the conjugate partition of $\lambda$ (defined in Theorem 5.1.10). Let $h^\lambda := \dfrac{n!}{f^\lambda}$.

Assume that $h^\lambda$ is invertible in $\mathbf{k}$. Then, the sign-twist $\left(\mathcal{S}^\lambda\right)^{\mathrm{sign}}$ of the Specht module $\mathcal{S}^\lambda$ is isomorphic to the Specht module $\mathcal{S}^{\lambda^t}$.

For example, for $n = 4$ and $\lambda = (3,1)$, this is saying that $\left(\mathcal{S}^{(3,1)}\right)^{\mathrm{sign}} \cong \mathcal{S}^{(3,1)^t} = \mathcal{S}^{(2,1,1)}$ as long as $h^{(3,1)} = \dfrac{4!}{f^{(3,1)}} = 8$ is invertible in $\mathbf{k}$.

To prove Theorem 5.18.13, we recall that we denote the $\mathbf{k}$-algebra $\mathbf{k}[S_n]$ by $\mathcal{A}$. Now, any left ideal $\mathcal{A}\mathbf{a}$ of $\mathcal{A}$ (with $\mathbf{a} \in \mathcal{A}$) is a left $\mathcal{A}$-module, that is, a left $\mathbf{k}[S_n]$-module, i.e., a representation of $S_n$. Hence, it has a well-defined sign-twist $(\mathcal{A}\mathbf{a})^{\mathrm{sign}}$. We claim the following general formula for this sign-twist:

**Lemma 5.18.14.** Let $\mathbf{a} \in \mathcal{A}$. Then,

$$(\mathcal{A}\mathbf{a})^{\mathrm{sign}} \cong \mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a}) \qquad \text{as } S_n\text{-representations.}$$

*Proof.* We will use an easy general fact about $\mathbf{k}$-algebras:

> *Claim 1:* Let $R$ and $R'$ be two $\mathbf{k}$-algebras. Let $\varphi : R \to R'$ be a $\mathbf{k}$-algebra isomorphism. Let $r \in R$. Then, the restriction of $\varphi$ to the left ideal $Rr$ is a $\mathbf{k}$-module isomorphism from $Rr$ to $R' \cdot \varphi(r)$.

The map $T_{\mathrm{sign}} : \mathcal{A} \to \mathcal{A}$ is a $\mathbf{k}$-algebra isomorphism (by Theorem 3.11.5 **(a)**, since $\mathcal{A} = \mathbf{k}[S_n]$). Hence, Claim 1 (applied to $R = \mathcal{A}$ and $R' = \mathcal{A}$ and $\varphi = T_{\mathrm{sign}}$ and $r = \mathbf{a}$) shows that the restriction of $T_{\mathrm{sign}}$ to the left ideal $\mathcal{A}\mathbf{a}$ is a $\mathbf{k}$-module isomorphism from $\mathcal{A}\mathbf{a}$ to $\mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a})$. Let us denote this restriction by $f$. Thus, $f : \mathcal{A}\mathbf{a} \to \mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a})$ is a $\mathbf{k}$-module isomorphism from $\mathcal{A}\mathbf{a}$ to $\mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a})$, and hence is invertible. Moreover,

$$f(\mathbf{u}) = T_{\mathrm{sign}}(\mathbf{u}) \qquad \text{for each } \mathbf{u} \in \mathcal{A}\mathbf{a} \tag{359}$$

(since $f$ is a restriction of $T_{\mathrm{sign}}$).

The map $f$ is $\mathbf{k}$-linear, but not left $\mathcal{A}$-linear – at least not when regarded as a map from $\mathcal{A}\mathbf{a}$ to $\mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a})$. However, we can also regard $f$ as a map from the sign-twist $(\mathcal{A}\mathbf{a})^{\mathrm{sign}}$ to $\mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a})$ (since $(\mathcal{A}\mathbf{a})^{\mathrm{sign}} = \mathcal{A}\mathbf{a}$ as sets), and then the situation improves:

*Claim 2:* The map $f$ is a left $\mathcal{A}$-linear map from the sign-twist $(\mathcal{A}\mathbf{a})^{\mathrm{sign}}$ to $\mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a})$.

*Proof of Claim 2.* Let us denote the left $\mathcal{A}$-module $\mathcal{A}\mathbf{a}$ by $V$. Thus, $\mathcal{A}\mathbf{a} = V$. Definition 5.18.3 yields that

$$\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v = T_{\mathrm{sign}}(\mathbf{a}) \overset{V}{\cdot} v \qquad \text{for all } \mathbf{a} \in \mathbf{k}[S_n] \text{ and } v \in V^{\mathrm{sign}} \tag{360}$$

(using Convention 5.18.1).

Let us denote the left $\mathcal{A}$-module $\mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a})$ by $W$. Thus, $\mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a}) = W$.

We must show that $f$ is a left $\mathcal{A}$-linear map from $(\mathcal{A}\mathbf{a})^{\mathrm{sign}}$ to $\mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a})$. In other words, we must show that $f$ is a left $\mathcal{A}$-linear map from $V^{\mathrm{sign}}$ to $W$ (since $\mathcal{A}\mathbf{a} = V$ and $\mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a}) = W$). Since we already know that $f$ respects addition and zero (because $f$ is $\mathbf{k}$-linear), we only need to show that

$$f\left(\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v\right) = \mathbf{a} \overset{W}{\cdot} f(v) \qquad \text{for all } \mathbf{a} \in \mathcal{A} \text{ and } v \in V^{\mathrm{sign}}.$$

So let us show this. Let $\mathbf{a} \in \mathcal{A}$ and $v \in V^{\mathrm{sign}}$. Thus, $\mathbf{a} \in \mathcal{A} = \mathbf{k}[S_n]$, so that (360) yields $\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v = T_{\mathrm{sign}}(\mathbf{a}) \overset{V}{\cdot} v$. Furthermore, $v \in V^{\mathrm{sign}} = V$ (we are talking about equality of sets here, not of $\mathcal{A}$-modules), so that $v \in V = \mathcal{A}\mathbf{a}$.

But $V$ is a left ideal of $\mathcal{A}$ (since $V = \mathcal{A}\mathbf{a}$); thus, the action of $\mathcal{A}$ on $V$ is just multiplication inside $\mathcal{A}$. In other words, $\mathbf{b} \overset{V}{\cdot} w = \mathbf{b}w$ for all $\mathbf{b} \in \mathcal{A}$ and $w \in V$. Hence, in particular, $T_{\mathrm{sign}}(\mathbf{a}) \overset{V}{\cdot} v = T_{\mathrm{sign}}(\mathbf{a}) \cdot v$. Similarly, we find $\mathbf{a} \overset{W}{\cdot} f(v) = \mathbf{a} \cdot f(v)$ (since $W$ is also a left ideal of $\mathcal{A}$). Now,

$$\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v = T_{\mathrm{sign}}(\mathbf{a}) \overset{V}{\cdot} v = T_{\mathrm{sign}}(\mathbf{a}) \cdot v.$$

But $\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v \in V^{\mathrm{sign}} = V = \mathcal{A}\mathbf{a}$. Hence, (359) (applied to $\mathbf{u} = \mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v$) yields

$$\begin{aligned} f\left(\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v\right) &= T_{\mathrm{sign}}\left(\underbrace{\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v}_{= T_{\mathrm{sign}}(\mathbf{a}) \cdot v}\right) = T_{\mathrm{sign}}\left(T_{\mathrm{sign}}(\mathbf{a}) \cdot v\right) \\ &= \underbrace{T_{\mathrm{sign}}\left(T_{\mathrm{sign}}(\mathbf{a})\right)}_{\substack{= \left(T_{\mathrm{sign}} \circ T_{\mathrm{sign}}\right)(\mathbf{a}) \\ = \mathbf{a} \\ \text{(since Theorem 3.11.5 \textbf{(b)}} \\ \text{says that } T_{\mathrm{sign}} \circ T_{\mathrm{sign}} = \mathrm{id}\text{)}}} \cdot T_{\mathrm{sign}}(v) \qquad \left(\begin{array}{c}\text{since } T_{\mathrm{sign}} \text{ is a } \mathbf{k}\text{-algebra} \\ \text{morphism}\end{array}\right) \\ &= \mathbf{a} \cdot T_{\mathrm{sign}}(v). \end{aligned}$$

Comparing this with

$$\mathbf{a} \overset{W}{\cdot} f(v) = \mathbf{a} \cdot \underbrace{f(v)}_{\substack{=T_{\mathrm{sign}}(v) \\ \text{(by (359),} \\ \text{applied to } \mathbf{u}=v)}} = \mathbf{a} \cdot T_{\mathrm{sign}}(v),$$

we obtain $f\left(\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v\right) = \mathbf{a} \overset{W}{\cdot} f(v)$.

Forget that we fixed $\mathbf{a}$ and $v$. We thus have proved that

$$f\left(\mathbf{a} \overset{V^{\mathrm{sign}}}{\cdot} v\right) = \mathbf{a} \overset{W}{\cdot} f(v) \qquad \text{for all } \mathbf{a} \in \mathcal{A} \text{ and } v \in V^{\mathrm{sign}}.$$

As explained above, this completes the proof of Claim 2. □

Now, Claim 2 shows that the map $f$ is a left $\mathcal{A}$-linear map from the sign-twist $(\mathcal{A}\mathbf{a})^{\mathrm{sign}}$ to $\mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a})$. Since $f$ is invertible, we thus conclude that $f$ is a left $\mathcal{A}$-module isomorphism (since any invertible left $\mathcal{A}$-linear map is a left $\mathcal{A}$-module isomorphism). Hence, $(\mathcal{A}\mathbf{a})^{\mathrm{sign}} \cong \mathcal{A} \cdot T_{\mathrm{sign}}(\mathbf{a})$ as left $\mathcal{A}$-modules, i.e., as left $\mathbf{k}[S_n]$-modules, i.e., as $S_n$-representations. This proves Lemma 5.18.14. □

Our next lemma is a general "left ideal avatar" for the sign-twist of a Specht module:

**Lemma 5.18.15.** Let $D$ be a diagram with $|D| = n$. Let $T$ be an $n$-tableau of shape $D$. Let $\mathbf{F}_T := \nabla_{\mathrm{Row}\,T} \nabla^{-}_{\mathrm{Col}\,T} \in \mathcal{A}$. Then,

$$\mathcal{S}^D \cong \mathcal{A}\mathbf{E}_T \qquad \text{and} \qquad \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}} \cong \mathcal{A}\mathbf{F}_T$$

as $S_n$-representations. (See Theorem 5.1.10 for the meaning of $\mathbf{r}$.)

*Proof.* From (228), we know that $\mathcal{S}^D \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules, i.e., as $S_n$-representations (since $\mathcal{A} = \mathbf{k}[S_n]$). It remains to prove that $\left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}} \cong \mathcal{A}\mathbf{F}_T$.

A straightforward generalization of Proposition 5.11.19 **(a)** (with the same proof) says that the antipode $S$ of $\mathbf{k}[S_n]$ satisfies $S(\mathbf{E}_T) = \mathbf{F}_T$.

We know that $T$ is an $n$-tableau of shape $D$. By Definition 5.15.10, we conclude that $T\mathbf{r}$ is an $n$-tableau of shape $\mathbf{r}(D)$. Hence, (228) (applied to $\mathbf{r}(D)$ and $T\mathbf{r}$ instead of $D$ and $T$) shows that $\mathcal{S}^{\mathbf{r}(D)} \cong \mathcal{A}\mathbf{E}_{T\mathbf{r}}$ as left $\mathcal{A}$-modules.

Furthermore, (277) (applied to $P = T$) yields $\mathbf{E}_{T\mathbf{r}} = T_{\mathrm{sign}}(S(\mathbf{E}_T))$. In view of $S(\mathbf{E}_T) = \mathbf{F}_T$, we can rewrite this as $\mathbf{E}_{T\mathbf{r}} = T_{\mathrm{sign}}(\mathbf{F}_T)$. Applying the map $T_{\mathrm{sign}}$ to this equality, we obtain

$$T_{\mathrm{sign}}(\mathbf{E}_{T\mathbf{r}}) = T_{\mathrm{sign}}\left(T_{\mathrm{sign}}(\mathbf{F}_T)\right) = \mathbf{F}_T$$

(since Theorem 3.11.5 **(b)** says that $T_{\mathrm{sign}} \circ T_{\mathrm{sign}} = \mathrm{id}$).

But we have $\mathcal{S}^{\mathbf{r}(D)} \cong \mathcal{A}\mathbf{E}_{T\mathbf{r}}$ as left $\mathcal{A}$-modules, i.e., as $S_n$-representations (since $\mathcal{A} = \mathbf{k}[S_n]$). Hence, by Corollary 5.18.9 (applied to $V = \mathcal{S}^{\mathbf{r}(D)}$ and $W = \mathcal{A}\mathbf{E}_{T\mathbf{r}}$), we obtain $\left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}} \cong \left(\mathcal{A}\mathbf{E}_{T\mathbf{r}}\right)^{\mathrm{sign}}$. Furthermore, Lemma 5.18.14 (applied to $\mathbf{a} = \mathbf{E}_{T\mathbf{r}}$) yields $\left(\mathcal{A}\mathbf{E}_{T\mathbf{r}}\right)^{\mathrm{sign}} \cong \mathcal{A} \cdot T_{\mathrm{sign}}\left(\mathbf{E}_{T\mathbf{r}}\right)$ as $S_n$-representations. In view of $T_{\mathrm{sign}}\left(\mathbf{E}_{T\mathbf{r}}\right) = \mathbf{F}_T$, we can rewrite this as $\left(\mathcal{A}\mathbf{E}_{T\mathbf{r}}\right)^{\mathrm{sign}} \cong \mathcal{A}\mathbf{F}_T$. Thus, $\left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}} \cong \left(\mathcal{A}\mathbf{E}_{T\mathbf{r}}\right)^{\mathrm{sign}} \cong \mathcal{A}\mathbf{F}_T$. This completes the proof of Lemma 5.18.15. □

We can now prove Theorem 5.18.13:

*Proof of Theorem 5.18.13.* Pick any $n$-tableau $T$ of shape $Y(\lambda)$. (Such a $T$ exists, as we saw in the proof of Corollary 5.7.10.) Thus, $T$ is an $n$-tableau of shape $\lambda$. Note that $|Y(\lambda)| = |\lambda| = n$ (since $\lambda$ is a partition of $n$).

Set $\mathbf{F}_T := \nabla_{\mathrm{Row}\,T} \nabla^{-}_{\mathrm{Col}\,T} \in \mathcal{A}$. Thus, Lemma 5.18.15 (applied to $D = Y(\lambda)$) yields that

$$\mathcal{S}^D \cong \mathcal{A}\mathbf{E}_T \qquad \text{and} \qquad \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}} \cong \mathcal{A}\mathbf{F}_T$$

as $S_n$-representations. But Proposition 5.11.19 **(c)** says that $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{E}_T \cong \mathcal{A}\mathbf{F}_T$ as left $\mathcal{A}$-modules, i.e., as $S_n$-representations.

However, $D = Y(\lambda)$, so that $\mathbf{r}(D) = \mathbf{r}(Y(\lambda)) = Y\left(\lambda^t\right)$ (by the definition of $\lambda^t$). Thus, $\mathcal{S}^{\mathbf{r}(D)} = \mathcal{S}^{Y\left(\lambda^t\right)} = \mathcal{S}^{\lambda^t}$ (since $\mathcal{S}^{\lambda^t}$ is a shorthand for $\mathcal{S}^{Y\left(\lambda^t\right)}$). Hence, $\mathcal{S}^{\lambda^t} = \mathcal{S}^{\mathbf{r}(D)}$, so that $\left(\mathcal{S}^{\lambda^t}\right)^{\mathrm{sign}} = \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}} \cong \mathcal{A}\mathbf{F}_T \cong \mathcal{S}^\lambda$ as $S_n$-representations (since $\mathcal{S}^\lambda \cong \mathcal{A}\mathbf{F}_T$ as $S_n$-representations). Hence, by Corollary 5.18.9 (applied to $V = \left(\mathcal{S}^{\lambda^t}\right)^{\mathrm{sign}}$ and $W = \mathcal{S}^\lambda$), we obtain

$$\left(\left(\mathcal{S}^{\lambda^t}\right)^{\mathrm{sign}}\right)^{\mathrm{sign}} \cong \left(\mathcal{S}^\lambda\right)^{\mathrm{sign}} \tag{361}$$

as $S_n$-representations.

Finally, Proposition 5.18.6 (applied to $V = \mathcal{S}^{\lambda^t}$) yields $\left(\left(\mathcal{S}^{\lambda^t}\right)^{\mathrm{sign}}\right)^{\mathrm{sign}} = \mathcal{S}^{\lambda^t}$. Hence, we can rewrite (361) as $\mathcal{S}^{\lambda^t} \cong \left(\mathcal{S}^\lambda\right)^{\mathrm{sign}}$. In other words, $\left(\mathcal{S}^\lambda\right)^{\mathrm{sign}} \cong \mathcal{S}^{\lambda^t}$. This proves Theorem 5.18.13. □

Note that Theorem 5.18.13 is not generally true without the "$h^\lambda$ is invertible" condition, as the following exercise shows:

**Exercise 5.18.3.** [2] As we know from Example 5.4.2, the Specht module $\mathcal{S}^{(2,1)}$ is isomorphic to the zero-sum subrepresentation $R\left(\mathbf{k}^3\right)$ of the natural $S_3$-representation $\mathbf{k}^3$.

Assume that $\mathbf{k}$ is a field of characteristic 3. Prove that $\left(\mathcal{S}^{(2,1)}\right)^{\mathrm{sign}}$ is not isomorphic to $\mathcal{S}^{(2,1)^t}$. (Note that $(2,1)^t = (2,1)$.)

In many situations, Theorem 5.18.13 can be generalized from straight-shaped Young diagrams to arbitrary diagrams:

**Theorem 5.18.16.** Assume that $\mathbf{k}$ is a field of characteristic 0. Let $D$ be any diagram with $|D| = n$. Then, the sign-twist $\left(\mathcal{S}^D\right)^{\mathrm{sign}}$ of the Specht module $\mathcal{S}^D$ is isomorphic to the Specht module $\mathcal{S}^{\mathbf{r}(D)}$ (where $\mathbf{r} : \mathbb{Z}^2 \to \mathbb{Z}^2$ is the map from Theorem 5.1.10).

However, we will need the tools of the next section to prove this theorem.

**Question 5.18.17.** Does Theorem 5.18.16 still hold if we merely assume $\mathbf{k}$ to be a commutative ring in which $n!$ is invertible (as opposed to a field of characteristic 0)?

**Question 5.18.18.** Is there an explicit construction of an isomorphism in Theorem 5.18.16?

## 5.19. The dual of an $S_n$-representation

We now move to a different operation that transforms $S_n$-representations (and, more generally, $G$-representations for any group $G$): dualization.

### 5.19.1. Reminders on dual k-modules

This operation is based on the linear-algebraic notion of a dual $\mathbf{k}$-module, which, in turn, generalizes the classical concept of a dual $\mathbf{k}$-vector space. We begin by recalling how these notions are defined. While duals are generally better-behaved for $\mathbf{k}$-vector spaces (i.e., when $\mathbf{k}$ is a field) than for $\mathbf{k}$-modules (i.e., when $\mathbf{k}$ is just a commutative ring), their definition is the same in both settings; thus we shall waste no time talking about the case of $\mathbf{k}$-vector spaces except in those situations where we can genuinely say more about it than about the general case.

We recall the definition of duals ([Axler24, 3.110]):

**Definition 5.19.1. (a)** If $V$ and $W$ are two $\mathbf{k}$-modules, then $\mathrm{Hom}(V, W)$ means the $\mathbf{k}$-module of all $\mathbf{k}$-linear maps from $V$ to $W$. (Its addition and scaling are entrywise: e.g., we have $(f+g)(v) = f(v) + g(v)$ for all $f, g \in \mathrm{Hom}(V, W)$ and $v \in V$.)

**(b)** If $V$ is any $\mathbf{k}$-module, then the $\mathbf{k}$-module $\mathrm{Hom}(V, \mathbf{k})$ is also denoted by $V^*$ and is called the *dual* (or *dual* $\mathbf{k}$*-module*, or *dual space*) of $V$.

Thus, the dual $V^*$ of a $\mathbf{k}$-module $V$ consists of all $\mathbf{k}$-linear maps from $V$ to $\mathbf{k}$; these are often called "linear functions" or "linear functionals" on $V$. Some authors also prefer the notation $V^\vee$ for $V^*$.

**Warning 5.19.2.** The dual $V^*$ of a $\mathbf{k}$-module $V$ depends not just on $V$ but also on the base ring $\mathbf{k}$. Thus, for example, a $\mathbb{Q}$-vector space $V$ is always a $\mathbb{Z}$-module (by restriction of scalars), but the dual $V^*$ of the $\mathbb{Q}$-vector space $V$ is not the same as the dual $V^*$ of the $\mathbb{Z}$-module $V$, not even as a set. (Indeed, it can be shown that the latter dual is always trivial, since the only $\mathbb{Z}$-linear map $f : V \to \mathbb{Z}$ from a $\mathbb{Q}$-vector space $V$ to $\mathbb{Z}$ is the constant-zero map.)

We can avoid this pitfall by using the unambiguous notation $V^*_{\mathbf{k}}$ (instead of $V^*$) for the dual of a $\mathbf{k}$-module $V$. However, we shall keep such precautions restricted to situations in which they are genuinely necessary (i.e., when we are considering one and the same module as a module over different base rings).

The following properties of duals are known from linear algebra (see [Conrad, "Dual modules", §5][124]):

**Proposition 5.19.3.** Let $V$ be a free $\mathbf{k}$-module with a finite basis $(e_i)_{i\in I}$. Then, its dual $V^*$ has a finite basis $\left(e_i^*\right)_{i\in I}$, which is defined as follows: For each $i \in I$, we let $e_i^* : V \to \mathbf{k}$ be the $\mathbf{k}$-linear map that sends each basis vector $e_j$ of $V$ to the Kronecker delta $\delta_{i,j} \in \mathbf{k}$.

This basis $\left(e_i^*\right)_{i\in I}$ is called the *dual basis* of the original basis $(e_i)_{i\in I}$. Its definition works even if the basis $(e_i)_{i\in I}$ is not finite, but in that case it is not a basis of $V^*$ but merely a $\mathbf{k}$-linearly independent family.

**Proposition 5.19.4.** Let $V$ and $W$ be two $\mathbf{k}$-modules.

**(a)** For each $\mathbf{k}$-linear map $f : V \to W$, there is a $\mathbf{k}$-linear map $f^* : W^* \to V^*$ (called the *dual* or the *adjoint* of $f$) defined by

$$f^*(g) := g \circ f \qquad \text{for each } g \in W^*.$$

**(b)** If $U$ is a further $\mathbf{k}$-module, and if $f : V \to W$ and $g : W \to U$ are two $\mathbf{k}$-linear maps, then $(g \circ f)^* = f^* \circ g^*$.

**(c)** If $f : V \to W$ is a $\mathbf{k}$-module isomorphism, then its dual $f^* : W^* \to V^*$ is also a $\mathbf{k}$-module isomorphism.

**(d)** Assume that the $\mathbf{k}$-modules $V$ and $W$ are free, with finite bases $(v_i)_{i\in I}$ and $\left(w_j\right)_{j\in J}$. Let $f : V \to W$ be a $\mathbf{k}$-linear map, and let $A \in \mathbf{k}^{J\times I}$ be the $J \times I$-matrix that represents $f$ with respect to the bases $(v_i)_{i\in I}$ and $\left(w_j\right)_{j\in J}$.

[124] In the case when $\mathbf{k}$ is a field, you can find these facts and more in any good text on linear algebra, such as [StoLui19, §6].

Let $\left(v_i^*\right)_{i\in I}$ and $\left(w_j^*\right)_{j\in J}$ be the dual bases of these bases $\left(v_i\right)_{i\in I}$ and $\left(w_j\right)_{j\in J}$. Then, the matrix that represents the map $f^* : W^* \to V^*$ with respect to these dual bases $\left(w_j^*\right)_{j\in J}$ and $\left(v_i^*\right)_{i\in I}$ is the transpose $A^T$ of the matrix $A$.

Proposition 5.19.4 **(d)** is the conceptual meaning of transpose matrices (if you view matrices as models for linear maps). Another cluster of well-known facts concerns double duals (i.e., duals of duals):

**Proposition 5.19.5.** Let $V$ be a $\mathbf{k}$-module.

**(a)** There is a canonical $\mathbf{k}$-linear map $\iota : V \to V^{**}$ (where $V^{**} = \left(V^*\right)^*$) that sends each vector $v \in V$ to the $\mathbf{k}$-linear map

$$\begin{aligned} \mathrm{ev}_v : V^* &\to \mathbf{k}, \\ g &\mapsto g\left(v\right). \end{aligned}$$

**(b)** If the $\mathbf{k}$-module $V$ has a finite basis, then this map $\iota$ is an isomorphism.

**(c)** In particular, if $\mathbf{k}$ is a field and $V$ is a finite-dimensional $\mathbf{k}$-vector space, then $\dim\left(V^*\right) = \dim V$.

A proof of Proposition 5.19.5 can be found in [Conrad, "Dual modules", §4].[125] Note that Proposition 5.19.5 **(b)** requires a **finite basis**; neither a finite generating set nor an infinite basis would be sufficient.

See [Conrad, "Dual modules"] for more about duals of $\mathbf{k}$-modules.

### 5.19.2. Bilinear forms

Bilinear forms are an important tool for understanding duals. To remind:

- A $\mathbf{k}$-bilinear map is a map $f : U \times V \to W$ (where $U$, $V$ and $W$ are three $\mathbf{k}$-modules) such that if we fix either argument, the map $f$ becomes a linear map in the other (see below for a detailed definition).
- A $\mathbf{k}$-bilinear form is just a $\mathbf{k}$-bilinear map to the base ring $\mathbf{k}$ itself.

Let us state these definitions in more detail, as the relevant notations will be useful later on. We begin with a notation for turning a two-argument map into a one-argument map by fixing one of the arguments to a constant:

[125] Those who believe in the axiom of choice (and classical logic) can say a bit more: The map $\iota$ in Proposition 5.19.5 **(a)** is injective whenever $\mathbf{k}$ is a field. See, e.g., `https://math.stackexchange.com/questions/4894044` for the proof of this fact (which we will not need in these notes). Keep in mind that it is false when $\mathbf{k}$ is not a field (not even the axiom of choice can make the map $\iota : \mathbb{Q} \to \mathbb{Q}^{**}$ injective for $\mathbf{k} = \mathbb{Z}$, since $\mathbb{Q}^* = 0$ entails $\mathbb{Q}^{**} = 0^* = 0$ here).

**Definition 5.19.6.** Let $U$, $V$ and $W$ be three sets, and $f : U \times V \to W$ be any map. Then:

**(a)** For any $u \in U$, we let $f(u, ?)$ be the map from $V$ to $W$ that sends each $v \in V$ to $f(u, v)$.

**(b)** For any $v \in V$, we let $f(?, v)$ be the map from $U$ to $W$ that sends each $u \in U$ to $f(u, v)$.

For instance, if $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ is the addition map (i.e., the map $(a, b) \mapsto a + b$), then $f(3, ?)$ is the map that sends each integer $b$ to $3 + b$. (This is the same as the map $f(?, 3)$, since addition is commutative.)

Next, we define bilinearity in detail:

**Definition 5.19.7.** Let $U$, $V$ and $W$ be three $\mathbf{k}$-modules, and $f : U \times V \to W$ be any map.

**(a)** We say that the map $f$ is *linear in the second argument* if for each $u \in U$, the map $f(u, ?) : V \to W$ is $\mathbf{k}$-linear. In this case, we define the map

$$\begin{aligned} f_L : U &\to \operatorname{Hom}(V, W), \\ u &\mapsto f(u, ?). \end{aligned}$$

We call $f_L$ the *left-curried form of* $f$. Explicitly, $f_L$ is thus given by the equality

$$(f_L(u))(v) = f(u, v) \qquad \text{for all } u \in U \text{ and } v \in V.$$

**(b)** We say that the map $f$ is *linear in the first argument* if for each $v \in V$, the map $f(?, v) : U \to W$ is $\mathbf{k}$-linear. In this case, we define the map

$$\begin{aligned} f_R : V &\to \operatorname{Hom}(U, W), \\ v &\mapsto f(?, v). \end{aligned}$$

We call $f_R$ the *right-curried form of* $f$. Explicitly, $f_R$ is thus given by the equality

$$(f_R(v))(u) = f(u, v) \qquad \text{for all } u \in U \text{ and } v \in V.$$

**(c)** We say that the map $f$ is *bilinear* (or $\mathbf{k}$-*bilinear*, to be more precise) if $f$ is both linear in the first argument and linear in the second argument. In this case, both maps $f_L : U \to \operatorname{Hom}(V, W)$ and $f_R : V \to \operatorname{Hom}(U, W)$ are easily seen to be $\mathbf{k}$-linear.

We can also rewrite this definition of "bilinear" explicitly as a list of axioms:

**Remark 5.19.8.** Let $U$, $V$ and $W$ be three $\mathbf{k}$-modules. Let $f : U \times V \to W$ be

a map. Then, $f$ is $\mathbf{k}$-bilinear if and only if it satisfies the six axioms

$$\begin{aligned}
f\left(u,\ v_1+v_2\right) &= f\left(u,v_1\right)+f\left(u,v_2\right) && \text{for all } u\in U \text{ and } v_1,v_2\in V;\\
f\left(u,\ \lambda v\right) &= \lambda f\left(u,v\right) && \text{for all } u\in U \text{ and } v\in V \text{ and } \lambda\in\mathbf{k};\\
f\left(u,0\right) &= 0 && \text{for all } u\in U;\\
f\left(u_1+u_2,\ v\right) &= f\left(u_1,v\right)+f\left(u_2,v\right) && \text{for all } u_1,u_2\in U \text{ and } v\in V;\\
f\left(\lambda u,\ v\right) &= \lambda f\left(u,v\right) && \text{for all } u\in U \text{ and } v\in V \text{ and } \lambda\in\mathbf{k};\\
f\left(0,v\right) &= 0 && \text{for all } v\in V.
\end{aligned}$$

**Definition 5.19.9.** Let $U$ and $V$ be two $\mathbf{k}$-modules. A *bilinear form* on $U$ and $V$ means a bilinear map $f : U\times V\to\mathbf{k}$.

Thus, if $f : U\times V\to\mathbf{k}$ is a bilinear form, then its left-curried form $f_L$ is a $\mathbf{k}$-linear map from $U$ to $V^*$, whereas its right-curried form $f_R$ is a $\mathbf{k}$-linear map from $V$ to $U^*$.

Bilinear forms on free $\mathbf{k}$-modules can be described by the values they take on the basis vectors:

**Proposition 5.19.10.** Let $U$ and $V$ be two free $\mathbf{k}$-modules with bases $\left(u_i\right)_{i\in I}$ and $\left(v_j\right)_{j\in J}$. Then, any bilinear form $f : U\times V\to\mathbf{k}$ is uniquely determined by its values $f\left(u_i,v_j\right)$ for all $i\in I$ and $j\in J$. Moreover, for any family $\left(a_{i,j}\right)_{(i,j)\in I\times J}\in\mathbf{k}^{I\times J}$ of scalars, there is a unique bilinear form $f : U\times V\to\mathbf{k}$ that satisfies

$$\left(f\left(u_i,v_j\right)=a_{i,j} \qquad \text{for all } i\in I \text{ and } j\in J\right).$$

Thus, a bilinear form $f : U\times V\to\mathbf{k}$ can be defined by providing its values $f\left(u_i,v_j\right)$ for all $i\in I$ and $j\in J$.

*Proof.* See [23wa, Theorem 3.9.2] (which proves the same claim in the more general setting where bilinear forms $f : U\times V\to\mathbf{k}$ are replaced by bilinear maps $f : U\times V\to W$). □

**Proposition 5.19.11.** Let $U$ and $V$ be two free $\mathbf{k}$-modules with finite bases $\left(u_i\right)_{i\in I}$ and $\left(v_j\right)_{j\in J}$. Let $f : U\times V\to\mathbf{k}$ be a bilinear form. Then:

**(a)** The $\mathbf{k}$-linear map $f_L : U\to V^*$ is invertible if and only if the matrix $\left(f\left(u_i,v_j\right)\right)_{(i,j)\in I\times J}\in\mathbf{k}^{I\times J}$ is invertible.

**(b)** The $\mathbf{k}$-linear map $f_R : V\to U^*$ is invertible if and only if the matrix $\left(f\left(u_i,v_j\right)\right)_{(i,j)\in I\times J}\in\mathbf{k}^{I\times J}$ is invertible.

*Proof.* **(b)** Let $\left(u_i^*\right)_{i\in I}$ and $\left(v_j^*\right)_{j\in J}$ be the dual bases of the bases $\left(u_i\right)_{i\in I}$ and $\left(v_j\right)_{j\in J}$ of $U$ and $V$. For each $j \in J$, we have

$$f_R\left(v_j\right) = \sum_{i\in I} f\left(u_i, v_j\right) u_i^*. \tag{362}$$

*Proof of (362).* Let $j \in J$. Then, both $f_R\left(v_j\right)$ and $\sum_{i\in I} f\left(u_i, v_j\right) u_i^*$ are $\mathbf{k}$-linear maps from $U$ to $\mathbf{k}$. Hence, in order to prove that these two maps are equal, it suffices to show that they agree on each of the basis vectors $u_k$ of $U$. But this is easy: For each $k \in I$, we have

$$\begin{aligned}
\left(\sum_{i\in I} f\left(u_i, v_j\right) u_i^*\right)\left(u_k\right) &= \sum_{i\in I} f\left(u_i, v_j\right) \underbrace{u_i^*\left(u_k\right)}_{\substack{=\delta_{i,k} \\ \text{(by the definition of the} \\ \text{dual basis } \left(u_i^*\right)_{i\in I}\text{)}}} = \sum_{i\in I} f\left(u_i, v_j\right) \delta_{i,k} \\
&= f\left(u_k, v_j\right) \underbrace{\delta_{k,k}}_{=1} + \sum_{\substack{i\in I; \\ i\neq k}} f\left(u_i, v_j\right) \underbrace{\delta_{i,k}}_{\substack{=0 \\ \text{(since } i\neq k\text{)}}} \\
&= f\left(u_k, v_j\right) + \underbrace{\sum_{\substack{i\in I; \\ i\neq k}} f\left(u_i, v_j\right) 0}_{=0} \\
&= f\left(u_k, v_j\right). \qquad (363)
\end{aligned}$$

On the other hand, for each $k \in I$, we have

$$\begin{aligned}
\left(f_R\left(v_j\right)\right)\left(u_k\right) &= f\left(u_k, v_j\right) \qquad \text{(by the definition of } f_R\text{)} \\
&= \left(\sum_{i\in I} f\left(u_i, v_j\right) u_i^*\right)\left(u_k\right) \qquad \text{(by (363))}.
\end{aligned}$$

In other words, the two $\mathbf{k}$-linear maps $f_R\left(v_j\right)$ and $\sum_{i\in I} f\left(u_i, v_j\right) u_i^*$ agree on each of the basis vectors $u_k$ of $U$. Because of their $\mathbf{k}$-linearity, this entails that these two maps are equal. Thus, (362) is proved. □

Now, the equality (362) shows that the matrix $\left(f\left(u_i, v_j\right)\right)_{(i,j)\in I\times J}$ represents the $\mathbf{k}$-linear map $f_R$ with respect to the bases $\left(v_j\right)_{j\in J}$ and $\left(u_i^*\right)_{i\in I}$ of $V$ and $U^*$. Hence, the latter map is invertible if and only if the former matrix is invertible (since invertible matrices represent invertible linear maps). This proves Proposition 5.19.11 **(b)**.

**(a)** In our proof of part **(b)**, we have shown that the $\mathbf{k}$-linear map $f_R : V \to U^*$ is invertible if and only if the matrix $\left(f\left(u_i, v_j\right)\right)_{(i,j)\in I\times J} \in \mathbf{k}^{I\times J}$ is invertible.

Similarly, we can show that the $\mathbf{k}$-linear map $f_L : U \to V^*$ is invertible if and only if the matrix $\left(f\left(u_i, v_j\right)\right)_{(j,i)\in J\times I} \in \mathbf{k}^{J\times I}$ (note that the two indices now appear in reverse order!) is invertible.[126] Hence, we have the following chain of logical equivalences:

$$\begin{aligned}
&\left(\text{the } \mathbf{k}\text{-linear map } f_L : U \to V^* \text{ is invertible}\right) \\
&\iff \left(\text{the matrix } \left(f\left(u_i, v_j\right)\right)_{(j,i)\in J\times I} \in \mathbf{k}^{J\times I} \text{ is invertible}\right) \\
&\iff \left(\text{the transpose of the matrix } \left(f\left(u_i, v_j\right)\right)_{(i,j)\in I\times J} \in \mathbf{k}^{I\times J} \text{ is invertible}\right) \\
&\qquad \left(\begin{array}{c} \text{since the matrix } \left(f\left(u_i, v_j\right)\right)_{(j,i)\in J\times I} \in \mathbf{k}^{J\times I} \text{ is the} \\ \text{transpose of the matrix } \left(f\left(u_i, v_j\right)\right)_{(i,j)\in I\times J} \in \mathbf{k}^{I\times J} \end{array}\right) \\
&\iff \left(\text{the matrix } \left(f\left(u_i, v_j\right)\right)_{(i,j)\in I\times J} \in \mathbf{k}^{I\times J} \text{ is invertible}\right)
\end{aligned}$$

(since the transpose of a matrix $A$ is invertible if and only if the matrix $A$ is invertible). This proves Proposition 5.19.11 **(a)**. □

**Definition 5.19.12.** The matrix $\left(f\left(u_i, v_j\right)\right)_{(i,j)\in I\times J} \in \mathbf{k}^{I\times J}$ in Proposition 5.19.11 is called the *Gram matrix* of the bilinear form $f$ with respect to the bases $\left(u_i\right)_{i\in I}$ and $\left(v_j\right)_{j\in J}$.

Another useful fact from linear algebra is the following:

**Lemma 5.19.13.** Assume that $\mathbf{k}$ is a field. Let $U$ and $V$ be two finite-dimensional $\mathbf{k}$-vector spaces. Let $\alpha : U \to V^*$ and $\beta : V \to U^*$ be two injective $\mathbf{k}$-linear maps. Then, these two maps $\alpha$ and $\beta$ are bijective.

*Proof.* Since $V$ is finite-dimensional, we have $\dim\left(V^*\right) = \dim V$ (by Proposition 5.19.5 **(c)**). Likewise, $\dim\left(U^*\right) = \dim U$.

Now, recall the following two classical facts from linear algebra (which are easy consequences of the rank-nullity theorem):

> *Fact 1:* If $\varphi : X \to Y$ is an injective $\mathbf{k}$-linear map between two finite-dimensional $\mathbf{k}$-vector spaces $X$ and $Y$, then $\dim X \leq \dim Y$.
>
> *Fact 2:* If $\varphi : X \to Y$ is an injective $\mathbf{k}$-linear map between two finite-dimensional $\mathbf{k}$-vector spaces $X$ and $Y$ that satisfy $\dim X \geq \dim Y$, then $\varphi$ is bijective.

[126] The proof relies on the identity

$$f_L\left(u_i\right) = \sum_{j\in J} f\left(u_i, v_j\right) v_j^* \qquad \text{for all } i \in I;$$

this is an analogue of (362).

We have assumed that the $\mathbf{k}$-linear map $\alpha : U \to V^*$ is injective. Hence, Fact 1 (applied to $X = U$ and $Y = V^*$ and $\varphi = \alpha$) yields $\dim U \leq \dim (V^*) = \dim V$. Hence, $\dim V \geq \dim U = \dim (U^*)$ (since $\dim (U^*) = \dim U$).

But we have also assumed that the $\mathbf{k}$-linear map $\beta : V \to U^*$ is injective. Hence, Fact 2 (applied to $X = V$ and $Y = U^*$ and $\varphi = \beta$) yields that $\beta$ is bijective (since $\dim V \geq \dim (U^*)$). Similarly, we can show that $\alpha$ is bijective. This proves Lemma 5.19.13. □

### 5.19.3. The dual of a group representation

We will now define the *dual* of a representation $V$ of a group $G$. Just as for sign-twists (in Section 5.18), we will give two definitions: one for representations of $G$ and one for left $\mathbf{k}[G]$-modules. As we know, a representation of $G$ (over $\mathbf{k}$) is "the same as" a left $\mathbf{k}[G]$-module (by Corollary 4.3.13), so that both definitions apply to the same objects; as we will soon see (Proposition 5.19.20 **(c)**), they also produce the same objects.

We begin with the first definition:

**Definition 5.19.14.** Let $G$ be a group. Let $V$ be a representation of $G$ over $\mathbf{k}$. For each $g \in G$, we consider the map

$$\begin{aligned} \tau_g : V &\to V, \\ x &\mapsto g \rightharpoonup x. \end{aligned}$$

This map $\tau_g$ is $\mathbf{k}$-linear (by Definition 4.2.2).

Now, the *dual* of $V$ is a new representation of $G$ over $\mathbf{k}$, which is called $V^*$ and is defined as follows:

- As a $\mathbf{k}$-module, $V^*$ is just the dual $V^*$ of the $\mathbf{k}$-module $V$ (as defined in Definition 5.19.1 **(b)**).
- The $G$-action on $V^*$ is the map $\overset{*}{\rightharpoonup} : G \times V^* \to V^*$ defined by the formula

$$g \overset{*}{\rightharpoonup} f := f \circ \tau_{g^{-1}} \qquad \text{for all } g \in G \text{ and } f \in V^*.$$

This representation $V^*$ really exists (by Proposition 5.19.20 **(a)** below), so the dual of $V$ is well-defined.

**Remark 5.19.15.** The definition of the $G$-action on $V^*$ in Definition 5.19.14 can be restated as follows: The $G$-action on $V^*$ is the map $\overset{*}{\rightharpoonup} : G \times V^* \to V^*$ that satisfies

$$\left(g \overset{*}{\rightharpoonup} f\right)(v) = f\left(g^{-1} \rightharpoonup v\right) \qquad \text{for all } g \in G \text{ and } f \in V^* \text{ and } v \in V \tag{364}$$

(where the "$\rightharpoonup$" on the right hand side means the action of $G$ on $V$). Indeed, the formula

$$g \overset{*}{\rightharpoonup} f := f \circ \tau_{g^{-1}} \qquad \text{for all } g \in G \text{ and } f \in V^*$$

(which defines $\overset{*}{\rightharpoonup}$ in Definition 5.19.14) is clearly equivalent to the formula

$$\left(g \overset{*}{\rightharpoonup} f\right)(v) = \left(f \circ \tau_{g^{-1}}\right)(v)$$
$$\text{for all } g \in G \text{ and } f \in V^* \text{ and } v \in V$$

(since two maps from $V$ to $\mathbf{k}$ are equal if and only if they give equal values on every $v \in V$). But the latter formula is, in turn, equivalent to (364), since all $g \in G$ and $f \in V^*$ and $v \in V$ satisfy

$$\left(f \circ \tau_{g^{-1}}\right)(v) = f\left(\underbrace{\tau_{g^{-1}}(v)}_{\substack{=g^{-1} \rightharpoonup v \\ \text{(by the definition of } \tau_{g^{-1}}\text{)}}}\right) = f\left(g^{-1} \rightharpoonup v\right).$$

For the second definition, we need a bit of preparation:

**Definition 5.19.16.** Let $G$ be a group. Consider the $\mathbf{k}$-linear map

$$S : \mathbf{k}[G] \to \mathbf{k}[G],$$
$$w \mapsto w^{-1} \qquad \text{for all } w \in G.$$

This map $S$ is called the *antipode* of $\mathbf{k}[G]$; for $G = S_n$, it was already defined in Definition 3.11.12. Its main properties are the following:

**(a)** The map $S : \mathbf{k}[G] \to \mathbf{k}[G]$ is a $\mathbf{k}$-algebra anti-automorphism.

**(b)** It is furthermore an involution (i.e., it satisfies $S \circ S = \mathrm{id}$).

Indeed, both of these properties were proved for $G = S_n$ in Theorem 3.11.14; the proofs in the general case are analogous.

We are now ready to define duals for left $\mathbf{k}[G]$-modules:

**Definition 5.19.17.** Let $G$ be a group. Let $V$ be a left $\mathbf{k}[G]$-module. Then, the *dual* of $V$ is a new left $\mathbf{k}[G]$-module, which is called $V^*$ and is defined as follows:

- As an additive group, $V^*$ is just the dual $V^*$ of the $\mathbf{k}$-module $V$ (as defined in Definition 5.19.1 **(b)**).

- The action of $\mathbf{k}[G]$ on $V^*$ is given by

$$\mathbf{a} \cdot f := f \circ \tau_{S(\mathbf{a})} \qquad \text{for all } \mathbf{a} \in \mathbf{k}[G] \text{ and } f \in V^*,$$

where $\tau_{S(\mathbf{a})}$ denotes the $\mathbf{k}$-linear map

$$\begin{aligned} V &\to V, \\ v &\mapsto S(\mathbf{a}) \cdot v. \end{aligned}$$

In other words, the action of $\mathbf{k}[G]$ on $V^*$ is the map

$$\begin{aligned} \mathbf{k}[G] \times V^* &\to V^*, \\ (\mathbf{a}, f) &\mapsto f \circ \tau_{S(\mathbf{a})}. \end{aligned}$$

This left $\mathbf{k}[G]$-module $V^*$ is well-defined (by Proposition 5.19.20 **(b)** below).

**Remark 5.19.18.** The definition of the action of $\mathbf{k}[G]$ on $V^*$ in Definition 5.19.17 can be restated as follows: The action of $\mathbf{k}[G]$ on $V^*$ is given by the rule

$$(\mathbf{a} \cdot f)(v) = f(S(\mathbf{a}) \cdot v) \qquad \text{for all } \mathbf{a} \in \mathbf{k}[G] \text{ and } f \in V^* \text{ and } v \in V \tag{365}$$

(where the "$\cdot$" on the right hand side means the action of $\mathbf{k}[G]$ on $V$). Indeed, the formula

$$\mathbf{a} \cdot f := f \circ \tau_{S(\mathbf{a})} \qquad \text{for all } \mathbf{a} \in \mathbf{k}[G] \text{ and } f \in V^*$$

(which defines the action of $\mathbf{k}[G]$ on $V^*$ in Definition 5.19.17) is clearly equivalent to the formula

$$(\mathbf{a} \cdot f)(v) = \left(f \circ \tau_{S(\mathbf{a})}\right)(v) \qquad \text{for all } \mathbf{a} \in \mathbf{k}[G] \text{ and } f \in V^* \text{ and } v \in V$$

(since two maps from $V$ to $\mathbf{k}$ are equal if and only if they give equal values on every $v \in V$). But the latter formula is, in turn, equivalent to (365), since all $\mathbf{a} \in \mathbf{k}[G]$ and $f \in V^*$ and $v \in V$ satisfy

$$\left(f \circ \tau_{S(\mathbf{a})}\right)(v) = f\left(\underbrace{\tau_{S(\mathbf{a})}(v)}_{\substack{= S(\mathbf{a}) \cdot v \\ \text{(by the definition of } \tau_{S(\mathbf{a})}\text{)}}}\right) = f(S(\mathbf{a}) \cdot v).$$

Soon (in Proposition 5.19.20) we will show that the above two definitions of duals are legitimate and equivalent. But first, let us give a simple example:

**Example 5.19.19.** Let $G$ be a group. Consider the trivial representation $\mathbf{k}_{\mathrm{triv}}$ of $G$, which is the $\mathbf{k}$-module $\mathbf{k}$ equipped with the trivial $G$-action given by

$$g \rightharpoonup v = v \qquad \text{for all } g \in G \text{ and } v \in \mathbf{k}.$$

(This generalizes the $\mathbf{k}_{\mathrm{triv}}$ from Example 4.2.16 to arbitrary groups.)

Let us find its dual $\left(\mathbf{k}_{\mathrm{triv}}\right)^*$ according to Definition 5.19.14. Indeed, as a $\mathbf{k}$-module, $\left(\mathbf{k}_{\mathrm{triv}}\right)^* = \mathbf{k}^* \cong \mathbf{k}$, whereas the $G$-action on $\left(\mathbf{k}_{\mathrm{triv}}\right)^*$ is the map $\overset{*}{\rightharpoonup} : G \times \mathbf{k}^* \to \mathbf{k}^*$ defined by the formula

$$g \overset{*}{\rightharpoonup} f := f \circ \tau_{g^{-1}} \qquad \text{for all } g \in G \text{ and } f \in \mathbf{k}^*.$$

This formula can be simplified to

$$g \overset{*}{\rightharpoonup} f := f \qquad \text{for all } g \in G \text{ and } f \in \mathbf{k}^*$$

(since the triviality of the $G$-action on $\mathbf{k}_{\mathrm{triv}}$ yields that $\tau_g = \mathrm{id}_V$ for each $g \in G$, and thus $f \circ \underbrace{\tau_{g^{-1}}}_{=\mathrm{id}_V} = f \circ \mathrm{id}_V = f$ for each $g \in G$ and $f \in \mathbf{k}^*$). This means that the $G$-action on $\left(\mathbf{k}_{\mathrm{triv}}\right)^*$ is again trivial. Hence, we conclude that $\left(\mathbf{k}_{\mathrm{triv}}\right)^* \cong \mathbf{k}_{\mathrm{triv}}$ (since $\left(\mathbf{k}_{\mathrm{triv}}\right)^* \cong \mathbf{k}$ as a $\mathbf{k}$-module).

More generally, the same arguments show that if $V$ is any trivial representation of $G$ (that is, any representation of $G$ that satisfies $g \rightharpoonup v = v$ for all $g \in G$ and $v \in V$), then its dual $V^*$ is again a trivial representation of $G$.

It is not much harder to see that the sign representation $\mathbf{k}_{\mathrm{sign}}$ of $S_n$ satisfies $\left(\mathbf{k}_{\mathrm{sign}}\right)^* \cong \mathbf{k}_{\mathrm{sign}}$.

We can easily find the dual of the regular representation of a finite group $G$ and of the natural representation of a symmetric group $S_n$:

**Exercise 5.19.1.** [2] Let $G$ be a finite group. Consider the left regular representation $\mathbf{k}[G]$ of $G$. Prove that its dual $(\mathbf{k}[G])^*$ is isomorphic to $\mathbf{k}[G]$ itself, and that the $\mathbf{k}$-linear map

$$\begin{aligned} \mathbf{k}[G] &\to (\mathbf{k}[G])^*, \\ g &\mapsto g^* \end{aligned}$$

(where $(h^*)_{h \in G}$ denotes the dual basis of the standard basis $(h)_{h \in G}$ of $\mathbf{k}[G]$) is an isomorphism of $G$-representations.

**Exercise 5.19.2.** [1] Consider the natural representation $\mathbf{k}^n$ of the symmetric group $S_n$ (defined in Example 4.2.17). Prove that its dual $(\mathbf{k}^n)^*$ is isomorphic to $\mathbf{k}^n$ itself, and that the $\mathbf{k}$-linear map

$$\begin{aligned} \mathbf{k}^n &\to (\mathbf{k}^n)^*, \\ e_i &\mapsto e_i^* \end{aligned}$$

(where $(e_1^*, e_2^*, \ldots, e_n^*)$ denotes the dual basis of the standard basis $(e_1, e_2, \ldots, e_n)$ of $\mathbf{k}^n$) is an isomorphism of $S_n$-representations.

These examples might suggest that each representation $V$ of a finite group $G$ satisfies $V^* \cong V$; but this is not always the case (for example, a nontrivial irrep of the cyclic group $C_3$ over $\mathbb{C}$ is not isomorphic to its dual[127]). Nevertheless, this is the case when $G$ is a symmetric group and $\mathbf{k}$ is a field of characteristic 0 and $V$ is finite-dimensional; this is a nontrivial and rather useful result, which we will prove below. We will also describe the duals of Specht modules in most cases. Note that Specht modules are not always isomorphic to their own duals in positive characteristic:

**Exercise 5.19.3.** [2] Let $\mathbf{k}$ be a field of characteristic 3. Prove that $\left(\mathcal{S}^{(2,1)}\right)^* \not\cong \mathcal{S}^{(2,1)}$ as $S_3$-representations.

We still owe a proof of the well-definedness of $V^*$, so let us give it now:

**Proposition 5.19.20.** Let $G$ be a group. Let $V$ be a left $\mathbf{k}[G]$-module, i.e., a representation of $G$ over $\mathbf{k}$. Then:

**(a)** The representation $V^*$ in Definition 5.19.14 is well-defined.

**(b)** The left $\mathbf{k}[G]$-module $V^*$ in Definition 5.19.17 is well-defined.

**(c)** The representation $V^*$ in Definition 5.19.14 and the left $\mathbf{k}[G]$-module $V^*$ in Definition 5.19.17 are the same (i.e., converting the former into a left $\mathbf{k}[G]$-module using Theorem 4.3.12 yields the latter).

*Proof.* Structurally, this is similar to the proof of Proposition 5.18.5, although the details of the computations differ. We again shall first prove part **(b)**, then use it to prove the other two parts.

However, before we start with the proof of part **(b)**, let us state some features of $V$.

[127] In more detail: Let $C_3 = \{1, g, g^2\}$ be the cyclic group of size 3. Let $\zeta$ be one of the two nontrivial 3-rd roots of unity in $\mathbb{C}$ (these are $e^{2\pi i/3}$ and $e^{-2\pi i/3}$). Let $\mathbb{C}_\zeta$ be the $C_3$-representation on the vector space $\mathbb{C}$ in which $g$ acts as multiplication by $\zeta$, and let $\mathbb{C}_{\zeta^2}$ be the $C_3$-representation on the vector space $\mathbb{C}$ in which $g$ acts as multiplication by $\zeta^2 = \zeta^{-1}$. Then, $(\mathbb{C}_\zeta)^*$ is isomorphic to $\mathbb{C}_{\zeta^2}$, not to $\mathbb{C}_\zeta$.

For each $\mathbf{b} \in \mathbf{k}[G]$, we let $\tau_{\mathbf{b}}$ denote the map

$$\begin{aligned} V &\to V, \\ v &\mapsto \mathbf{b} \cdot v. \end{aligned}$$

This generalizes the map $\tau_{S(\mathbf{a})}$ in Definition 5.19.17.

Recall that the action of $\mathbf{k}[G]$ on $V$ (viewed as the map $\mathbf{k}[G] \times V \to V$, $(\mathbf{b}, v) \mapsto \mathbf{b} \cdot v$) is $\mathbf{k}$-bilinear. Thus, for each $\mathbf{b} \in \mathbf{k}[G]$, the map $\tau_{\mathbf{b}}$ is $\mathbf{k}$-linear, and depends $\mathbf{k}$-linearly on $\mathbf{b}$ (meaning that $\tau_{\lambda\mathbf{b}+\mu\mathbf{c}} = \lambda\tau_{\mathbf{b}} + \mu\tau_{\mathbf{c}}$ for any $\lambda, \mu \in \mathbf{k}$ and any $\mathbf{b}, \mathbf{c} \in \mathbf{k}[G]$).

Furthermore, for any $\mathbf{b}, \mathbf{c} \in \mathbf{k}[G]$, we have

$$\tau_{\mathbf{bc}} = \tau_{\mathbf{b}} \circ \tau_{\mathbf{c}} \tag{366}$$

[128]. Moreover, the unity 1 of $\mathbf{k}[G]$ satisfies $\tau_1 = \mathrm{id}$ (since each $v \in V$ satisfies $\tau_1(v) = 1 \cdot v = v = \mathrm{id}(v)$). Hence, for each $\lambda \in \mathbf{k}$, we have

$$\begin{aligned} \tau_{\lambda \cdot 1} &= \lambda \underbrace{\tau_1}_{=\mathrm{id}} \qquad \text{(since } \tau_{\mathbf{b}} \text{ depends } \mathbf{k}\text{-linearly on } \mathbf{b}\text{)} \\ &= \lambda\,\mathrm{id}\,. \end{aligned} \tag{367}$$

Now, we step to the proofs of parts **(b)**, **(a)** and **(c)**.

**(b)** Consider the map

$$\begin{aligned} \mathbf{k}[G] \times V^* &\to V^*, \\ (\mathbf{a}, f) &\mapsto f \circ \tau_{S(\mathbf{a})}. \end{aligned}$$

This is the map that is intended to serve as the action of $\mathbf{k}[G]$ on $V^*$ (according to Definition 5.19.17). Hence, let us denote this map by $\overset{V^*}{\cdot}$ (that is, let us write $\mathbf{a} \overset{V^*}{\cdot} f$ for the image of a pair $(\mathbf{a}, f) \in \mathbf{k}[G] \times V^*$ under this map). Thus,

$$\mathbf{a} \overset{V^*}{\cdot} f = f \circ \tau_{S(\mathbf{a})} \tag{368}$$

for all $\mathbf{a} \in \mathbf{k}[G]$ and $f \in V^*$.

The map $\overset{V^*}{\cdot}$ is easily seen to be $\mathbf{k}$-bilinear (since composition of $\mathbf{k}$-linear maps is $\mathbf{k}$-bilinear; since the map $S$ is $\mathbf{k}$-linear; and since the linear map $\tau_{\mathbf{b}} : V \to V$ depends $\mathbf{k}$-linearly on $\mathbf{b}$).

In Definition 5.19.17, we defined the left $\mathbf{k}[G]$-module $V^*$ to be the additive group $V^*$ (the dual of the $\mathbf{k}$-module $V$), equipped with the map $\overset{V^*}{\cdot}$ as the action of $\mathbf{k}[G]$ on

[128] *Proof.* Let $\mathbf{b}, \mathbf{c} \in \mathbf{k}[G]$. Let $v \in V$. Then, $\tau_{\mathbf{bc}}(v) = \mathbf{bc} \cdot v$ (by the definition of $\tau_{\mathbf{bc}}$) and $\tau_{\mathbf{c}}(v) = \mathbf{c} \cdot v$ (similarly). Now,

$$\begin{aligned} (\tau_{\mathbf{b}} \circ \tau_{\mathbf{c}})(v) = \tau_{\mathbf{b}}(\tau_{\mathbf{c}}(v)) &= \mathbf{b} \cdot \underbrace{\tau_{\mathbf{c}}(v)}_{=\mathbf{c}\cdot v} \qquad \text{(by the definition of } \tau_{\mathbf{b}}\text{)} \\ &= \mathbf{b} \cdot (\mathbf{c} \cdot v) = \mathbf{bc} \cdot v \qquad \text{(since } V \text{ is a left } \mathbf{k}[G]\text{-module)}\,. \end{aligned}$$

Comparing this with $\tau_{\mathbf{bc}}(v) = \mathbf{bc} \cdot v$, we obtain $\tau_{\mathbf{bc}}(v) = (\tau_{\mathbf{b}} \circ \tau_{\mathbf{c}})(v)$.

Forget that we fixed $v$. We thus have proved that $\tau_{\mathbf{bc}}(v) = (\tau_{\mathbf{b}} \circ \tau_{\mathbf{c}})(v)$ for each $v \in V$. In other words, $\tau_{\mathbf{bc}} = \tau_{\mathbf{b}} \circ \tau_{\mathbf{c}}$. This proves (366).

$V^*$. In order to prove that this left $\mathbf{k}[G]$-module $V^*$ is well-defined, we must show that this structure satisfies all the module axioms (see Definition 4.3.1). Most of these axioms are straightforward to check (e.g., both right and left distributivity laws follow from the $\mathbf{k}$-bilinearity of the map $\overset{V^*}{\cdot}$), so we only focus on two axioms: the associativity law and the "$1m = m$" axiom. Let us check them both:

- *The associativity law:* We must prove the associativity law for the left $\mathbf{k}[G]$-module $V^*$. In other words, we must prove that $(rs) \overset{V^*}{\cdot} m = r \overset{V^*}{\cdot} \left(s \overset{V^*}{\cdot} m\right)$ for all $r, s \in \mathbf{k}[G]$ and $m \in V^*$. So let $r, s \in \mathbf{k}[G]$ and $m \in V^*$ be arbitrary. Then, the definition of $\overset{V^*}{\cdot}$ yields $s \overset{V^*}{\cdot} m = m \circ \tau_{S(s)}$ and $r \overset{V^*}{\cdot} \left(s \overset{V^*}{\cdot} m\right) = \left(s \overset{V^*}{\cdot} m\right) \circ \tau_{S(r)}$ and $(rs) \overset{V^*}{\cdot} m = m \circ \tau_{S(rs)}$. But $S$ is a $\mathbf{k}$-algebra anti-morphism (by Definition 5.19.16 **(a)**), and thus we have $S(rs) = S(s) \cdot S(r)$. Hence, $\tau_{S(rs)} = \tau_{S(s) \cdot S(r)} = \tau_{S(s)} \circ \tau_{S(r)}$ (by (366)). Thus,

$$(rs) \overset{V^*}{\cdot} m = m \circ \underbrace{\tau_{S(rs)}}_{= \tau_{S(s)} \circ \tau_{S(r)}} = m \circ \tau_{S(s)} \circ \tau_{S(r)}.$$

  Comparing this with

$$r \overset{V^*}{\cdot} \left(s \overset{V^*}{\cdot} m\right) = \underbrace{\left(s \overset{V^*}{\cdot} m\right)}_{= m \circ \tau_{S(s)}} \circ \tau_{S(r)} = m \circ \tau_{S(s)} \circ \tau_{S(r)},$$

  we obtain $(rs) \overset{V^*}{\cdot} m = r \overset{V^*}{\cdot} \left(s \overset{V^*}{\cdot} m\right)$. Thus, the associativity law for the $\mathbf{k}[G]$-module $V^*$ is proved.

- *The $1m = m$ axiom:* We must prove the "$1m = m$" axiom for the left $\mathbf{k}[G]$-module $V^*$. In other words, we must prove that $1 \overset{V^*}{\cdot} m = m$ for all $m \in V^*$ (where 1 is the unity of $\mathbf{k}[G]$). So let $m \in V^*$. Then, the definition of $\overset{V^*}{\cdot}$ yields $1 \overset{V^*}{\cdot} m = m \circ \tau_{S(1)}$. But $S$ is a $\mathbf{k}$-algebra anti-morphism (by Definition 5.19.16 **(a)**), and thus we have $S(1) = 1$. Hence, $\tau_{S(1)} = \tau_1 = \mathrm{id}$. Thus, $1 \overset{V^*}{\cdot} m = m \circ \underbrace{\tau_{S(1)}}_{= \mathrm{id}} = m$. Thus, we have proved the "$1m = m$" axiom for the left $\mathbf{k}[G]$-module $V^*$.

As we said, the other axioms are even easier to check. Thus, the left $\mathbf{k}[G]$-module $V^*$ is well-defined. This proves Proposition 5.19.20 **(b)**.

**(a)** Instead of proving the claim directly, we shall derive it from part **(b)** (which we have already proved).

First, we introduce some notations (exclusively for this proof):

- We shall write $V_1^*$ for the representation $V^*$ defined in Definition 5.19.14 (even though we have not yet shown that it is well-defined). Our goal is to prove that this representation $V_1^*$ is well-defined.

- We shall write $V_2^*$ for the left $\mathbf{k}[G]$-module $V^*$ defined in Definition 5.19.17. Proposition 5.19.20 **(b)** (which we have already proved) shows that the latter left $\mathbf{k}[G]$-module $V_2^*$ is well-defined. We reinterpret this left $\mathbf{k}[G]$-module $V_2^*$ as a representation of $G$ (using Corollary 4.3.13, as usual).

We shall now show that this representation $V_2^*$ is precisely the (purported) representation $V_1^*$. This will immediately entail that the latter representation is well-defined, and thus Proposition 5.19.20 **(a)** will follow (since $V_1^*$ is precisely the representation $V^*$ defined in Definition 5.19.14).

First, we recall that $S$ is a $\mathbf{k}$-algebra anti-morphism (by Definition 5.19.16 **(a)**), and thus we have $S\left(1_{\mathbf{k}[G]}\right) = 1_{\mathbf{k}[G]}$. But $S$ is also $\mathbf{k}$-linear. Hence, for each $\lambda \in \mathbf{k}$, we have

$$S\left(\lambda 1_{\mathbf{k}[G]}\right) = \lambda \cdot \underbrace{S\left(1_{\mathbf{k}[G]}\right)}_{=1_{\mathbf{k}[G]}} = \lambda 1_{\mathbf{k}[G]}. \tag{369}$$

Next, we recall that $V_2^*$ is the left $\mathbf{k}[G]$-module $V^*$ from Definition 5.19.17. Hence, the action of $\mathbf{k}[G]$ on $V_2^*$ is given by

$$\mathbf{a} \overset{V_2^*}{\cdot} f := f \circ \tau_{S(\mathbf{a})} \tag{370}$$

for all $\mathbf{a} \in \mathbf{k}[G]$ and $f \in V^*$ (by Definition 5.19.17).

Now, we shall show that our $G$-representation $V_2^*$ (which we obtained from the left $\mathbf{k}[G]$-module $V_2^*$ defined in Definition 5.19.17) is the (purported) representation $V_1^*$ defined in Definition 5.19.14. Indeed:

- The underlying $\mathbf{k}$-module of our $G$-representation $V_2^*$ is the dual $V^*$ of the $\mathbf{k}$-module $V$.

  [*Proof:* We must show that $V_2^* = V^*$ as $\mathbf{k}$-modules. As an additive group, $V_2^*$ is just the dual $V^*$ of the $\mathbf{k}$-module $V$ (by its definition). In other words, our $G$-representation $V_2^*$ has the same addition and zero as $V^*$. Thus, it remains to show that $V_2^*$ also has the same scaling (by elements of $\mathbf{k}$) as $V^*$. In other words, it remains to prove that every $\lambda \in \mathbf{k}$ and $f \in V^*$ satisfy $\lambda \overset{V_2^*}{\cdot} f = \lambda f$ (where the right hand side is computed in the dual $V^*$ of the $\mathbf{k}$-module $V$). So let us prove this. For any $\lambda \in \mathbf{k}$ and $f \in V^*$, we have

$$\begin{aligned}
\lambda \overset{V_2^*}{\cdot} f &= \left(\lambda 1_{\mathbf{k}[G]}\right) \overset{V_2^*}{\cdot} f && \left(\begin{array}{c}\text{since the } \mathbf{k}\text{-module structure on } V_2^* \text{ is} \\ \text{obtained from the } \mathbf{k}[G]\text{-module structure} \\ \text{by restriction of scalars}\end{array}\right) \\
&= f \circ \tau_{S\left(\lambda 1_{\mathbf{k}[G]}\right)} && \left(\text{by (370), applied to } \mathbf{a} = \lambda 1_{\mathbf{k}[G]}\right) \\
&= f \circ \underbrace{\tau_{\lambda 1_{\mathbf{k}[G]}}}_{\substack{=\tau_{\lambda 1} \\ =\lambda \operatorname{id} \\ \text{(by (367))}}} && \left(\text{since (369) yields } S\left(\lambda 1_{\mathbf{k}[G]}\right) = \lambda 1_{\mathbf{k}[G]}\right) \\
&= f \circ (\lambda \operatorname{id}) = \lambda f
\end{aligned}$$

(since each $v \in V$ satisfies $(f \circ (\lambda \operatorname{id}))(v) = f\left(\underbrace{(\lambda \operatorname{id})(v)}_{=\lambda v}\right) = f(\lambda v) = \lambda f(v)$ due to the $\mathbf{k}$-linearity of $f$). Thus, we have shown that that every $\lambda \in \mathbf{k}$ and $f \in V^*$ satisfy $\lambda \overset{V_2^*}{\cdot} f = \lambda f$. In other words, the $\mathbf{k}$-module $V_2^*$ has the same scaling as $V^*$. Since $V_2^*$ also has the same addition and zero as $V^*$, we thus conclude that the $\mathbf{k}$-module $V_2^*$ is identical with the $\mathbf{k}$-module $V^*$.]

- The action of $G$ on our $G$-representation $V_2^*$ is exactly the map $\overset{*}{\rightharpoonup} : G \times V^* \to V^*$ defined in Definition 5.19.14.

  [*Proof:* Let $\underset{2}{\overset{*}{\rightharpoonup}} : G \times V_2^* \to V_2^*$ be the action of $G$ on the representation $V_2^*$ of $G$. Then, for any $g \in G$ and $f \in V_2^*$, we have

$$\begin{aligned} g \underset{2}{\overset{*}{\rightharpoonup}} f &= g \overset{V_2^*}{\cdot} f = f \circ \tau_{S(g)} && \text{(by (370), applied to } \mathbf{a} = g) \\ &= f \circ \tau_{g^{-1}} && \left( \begin{array}{c} \text{since the definition of } S \text{ yields } S(g) = g^{-1} \\ \text{(because } g \in G) \end{array} \right). \end{aligned}$$

  But this is precisely the formula for the map $\overset{*}{\rightharpoonup} : G \times V^* \to V^*$ defined in Definition 5.19.14. Thus, the map $\underset{2}{\overset{*}{\rightharpoonup}} : G \times V_2^* \to V_2^*$ is precisely the map $\overset{*}{\rightharpoonup} : G \times V^* \to V^*$ defined in Definition 5.19.14. In other words, the action of $G$ on our $G$-representation $V_2^*$ is exactly the map $\overset{*}{\rightharpoonup} : G \times V^* \to V^*$ defined in Definition 5.19.14.]

Thus, we have shown that our representation $V_2^*$ of $G$ has the same underlying $\mathbf{k}$-module as $V^*$, and its action of $G$ is precisely the map $\overset{*}{\rightharpoonup} : G \times V^* \to V^*$ from Definition 5.19.14. Consequently, this representation $V_2^*$ is exactly the representation $V^*$ defined in Definition 5.19.14 (i.e., the representation that we now call $V_1^*$). Therefore, the latter representation $V^*$ is well-defined (since the former representation is well-defined). This proves Proposition 5.19.20 **(a)**.

**(c)** We must prove that the representation $V^*$ in Definition 5.19.14 and the left $\mathbf{k}[G]$-module $V^*$ in Definition 5.19.17 are the same. In our above proof of part **(a)**, we have denoted the former representation by $V_1^*$ and the latter $\mathbf{k}[G]$-module by $V_2^*$; thus, we must prove that $V_1^*$ and $V_2^*$ are the same.

But we have essentially done this already. Indeed, in our above proof of part **(a)**, we have shown that the representation $V_2^*$ is exactly the representation $V^*$ defined in Definition 5.19.14; but the latter representation is precisely the one that we called $V_1^*$. Hence, we have shown that $V_1^*$ and $V_2^*$ are the same. This proves Proposition 5.19.20 **(c)**. □

Proposition 5.19.20 **(a)** and **(b)** shows that our two definitions of $V^*$ are well-defined, and Proposition 5.19.20 **(c)** shows that they are equivalent. We now turn to some simple properties of duals. For the following proposition, we recall the notion of the dual $f^*$ of a $\mathbf{k}$-linear map $f$, as defined in Proposition 5.19.4 **(a)**:

**Proposition 5.19.21.** Let $G$ be a group. Let $V$ be a representation of $G$. Then:

**(a)** For each $g \in G$, we have

$$\left(\tau_g \text{ on } V^*\right) = \left(\tau_{g^{-1}} \text{ on } V\right)^*.$$

Here, "$\tau_g$ on $V^*$" means the $\mathbf{k}$-linear map $V^* \to V^*$, $f \mapsto g \rightharpoonup f$, whereas "$\tau_{g^{-1}}$ on $V$" means the $\mathbf{k}$-linear map $V \to V$, $v \mapsto g^{-1} \rightharpoonup v$.

**(b)** More generally, for each $\mathbf{a} \in \mathbf{k}[G]$, we have

$$\left(\tau_{\mathbf{a}} \text{ on } V^*\right) = \left(\tau_{S(\mathbf{a})} \text{ on } V\right)^*.$$

Here, "$\tau_{\mathbf{a}}$ on $V^*$" means the $\mathbf{k}$-linear map $V^* \to V^*$, $f \mapsto \mathbf{a} \cdot f$, whereas "$\tau_{S(\mathbf{a})}$ on $V$" means the $\mathbf{k}$-linear map $V \to V$, $v \mapsto S(\mathbf{a}) \cdot v$.

*Proof.* These are essentially just the definitions of the action of $G$ (resp. $\mathbf{k}[G]$) on $V^*$, restated in terms of the dual of a map. In more detail:

**(b)** Let $\mathbf{a} \in \mathbf{k}[G]$. Then, for each $f \in V^*$, we have

$$\begin{aligned}
\left(\tau_{\mathbf{a}} \text{ on } V^*\right)(f) &= \mathbf{a} \cdot f \qquad \left(\text{by the definition of } \left(\tau_{\mathbf{a}} \text{ on } V^*\right)\right) \\
&= f \circ \tau_{S(\mathbf{a})} \qquad \left(\begin{array}{c}\text{by the definition of the } \mathbf{k}[G]\text{-action on } V^* \\ \text{(Definition 5.19.17)}\end{array}\right) \\
&= f \circ \left(\tau_{S(\mathbf{a})} \text{ on } V\right) \qquad \left(\begin{array}{c}\text{since the } \tau_{S(\mathbf{a})} \text{ in Definition 5.19.17} \\ \text{is what we call } \left(\tau_{S(\mathbf{a})} \text{ on } V\right) \text{ here}\end{array}\right) \\
&= \left(\tau_{S(\mathbf{a})} \text{ on } V\right)^* (f)
\end{aligned}$$

(since the definition of the dual $\left(\tau_{S(\mathbf{a})} \text{ on } V\right)^*$ yields $\left(\tau_{S(\mathbf{a})} \text{ on } V\right)^* (f) = f \circ \left(\tau_{S(\mathbf{a})} \text{ on } V\right)$). In other words, $\left(\tau_{\mathbf{a}} \text{ on } V^*\right) = \left(\tau_{S(\mathbf{a})} \text{ on } V\right)^*$. This proves Proposition 5.19.21 **(b)**.

**(a)** This is analogous to part **(b)**, except that we are using Definition 5.19.14 instead of Definition 5.19.17 as the definition of $V^*$ here. Alternatively, because Definition 5.19.14 and Definition 5.19.17 are equivalent, we can derive this from Proposition 5.19.21 **(b)** (applied to $\mathbf{a} = g$). □

**Proposition 5.19.22.** Let $G$ be a group. Let $V$ be a representation of $G$. Then:

**(a)** The map $\iota : V \to V^{**}$ from Proposition 5.19.5 **(a)** is a morphism of $G$-representations.

**(b)** Thus, if $V$ has a finite basis, then $V^{**} \cong V$ as $G$-representations.

*Proof.* **(a)** For any $G$-representation $W$ and any $\mathbf{b} \in \mathbf{k}[G]$, we let $\tau_{\mathbf{b}}^{W}$ denote the $\mathbf{k}$-linear map $W \to W,\ v \mapsto \mathbf{b}v$. (This map would have been previously denoted by $\tau_{\mathbf{b}}$, but we want to make the role of $W$ explicit.) Thus, the action of $\mathbf{k}[G]$ on $V^*$ is given by

$$\mathbf{a} \cdot f = f \circ \tau_{S(\mathbf{a})}^{V} \qquad \text{for all } \mathbf{a} \in \mathbf{k}[G] \text{ and } f \in V^* \tag{371}$$

(indeed, this is precisely the formula $\mathbf{a} \cdot f = f \circ \tau_{S(\mathbf{a})}$ from Definition 5.19.17, since the map that was called $\tau_{S(\mathbf{a})}$ in Definition 5.19.17 is now called $\tau_{S(\mathbf{a})}^{V}$). Applying the same reasoning to $V^*$ instead of $V$, we see that the action of $\mathbf{k}[G]$ on $V^{**}$ is given by

$$\mathbf{a} \cdot f = f \circ \tau_{S(\mathbf{a})}^{V^*} \qquad \text{for all } \mathbf{a} \in \mathbf{k}[G] \text{ and } f \in V^{**}. \tag{372}$$

The definition of $\iota$ yields that

$$(\iota(v))(g) = g(v) \qquad \text{for each } g \in V^* \text{ and each } v \in V. \tag{373}$$

We must prove that the map $\iota : V \to V^{**}$ is a morphism of $G$-representations. Equivalently, we must prove that this map $\iota$ is left $\mathbf{k}[G]$-linear. But we already know (from Proposition 5.19.5 **(a)**) that it is $\mathbf{k}$-linear. Thus, all we need to show is that it satisfies $\iota(\mathbf{a} \cdot v) = \mathbf{a} \cdot \iota(v)$ for all $\mathbf{a} \in \mathbf{k}[G]$ and all $v \in V$.

Let us show this.

Let $\mathbf{a} \in \mathbf{k}[G]$ and $v \in V$. We must show that $\iota(\mathbf{a} \cdot v) = \mathbf{a} \cdot \iota(v)$.

Definition 5.19.16 **(b)** says that $S \circ S = \mathrm{id}$. Thus, $S(S(\mathbf{a})) = \mathbf{a}$.

Applying (372) to $f = \iota(v)$, we find $\mathbf{a} \cdot \iota(v) = \iota(v) \circ \tau_{S(\mathbf{a})}^{V^*}$. Hence, for each $f \in V^*$, we have

$$\underbrace{(\mathbf{a} \cdot \iota(v))}_{=\iota(v)\circ\tau_{S(\mathbf{a})}^{V^*}}(f) = \left(\iota(v) \circ \tau_{S(\mathbf{a})}^{V^*}\right)(f) = (\iota(v))\ \underbrace{\left(\tau_{S(\mathbf{a})}^{V^*}(f)\right)}_{\substack{=S(\mathbf{a})\cdot f \\ \text{(by the definition of } \tau_{S(\mathbf{a})}^{V^*}\text{)}}} = (\iota(v))(S(\mathbf{a}) \cdot f)$$

$$= \underbrace{(S(\mathbf{a}) \cdot f)}_{\substack{=f\circ\tau_{S(S(\mathbf{a}))}^{V} \\ \text{(by (371), applied to } S(\mathbf{a}) \\ \text{instead of } \mathbf{a}\text{)}}}(v) \qquad \text{(by (373), applied to } g = S(\mathbf{a}) \cdot f)$$

$$= \left(f \circ \tau_{S(S(\mathbf{a}))}^{V}\right)(v) = \left(f \circ \tau_{\mathbf{a}}^{V}\right)(v) \qquad (\text{since } S(S(\mathbf{a})) = \mathbf{a})$$

$$= f\left(\underbrace{\tau_{\mathbf{a}}^{V}(v)}_{\substack{=\mathbf{a}\cdot v \\ \text{(by the definition of } \tau_{\mathbf{a}}^{V}\text{)}}}\right) = f(\mathbf{a} \cdot v) = (\iota(\mathbf{a} \cdot v))(f)$$

(since (373) (applied to $\mathbf{a} \cdot v$ and $f$ instead of $v$ and $g$) yields $(\iota(\mathbf{a} \cdot v))(f) = f(\mathbf{a} \cdot v)$). In other words, $\mathbf{a} \cdot \iota(v) = \iota(\mathbf{a} \cdot v)$. In other words, $\iota(\mathbf{a} \cdot v) = \mathbf{a} \cdot \iota(v)$. As we explained above, this completes the proof of Proposition 5.19.22 **(a)**.

**(b)** Assume that $V$ has a finite basis. Then, Proposition 5.19.5 **(b)** shows that the map $\iota : V \to V^{**}$ is an isomorphism of $\mathbf{k}$-modules, hence is invertible. But Proposition 5.19.22 **(a)** shows that this map $\iota$ is a morphism of $G$-representations. Thus, $\iota$ is an invertible morphism of $G$-representations, hence an isomorphism of $G$-representations. Therefore, $V^{**} \cong V$. This proves Proposition 5.19.22 **(b)**. $\square$

We give two more basic properties of duals without proof:

**Proposition 5.19.23.** Let $G$ be a group. Let $V$ and $W$ be two representations of $G$. Then, $(V \oplus W)^* \cong V^* \oplus W^*$ canonically. (To be more specific: The standard $\mathbf{k}$-module isomorphism $(V \oplus W)^* \to V^* \oplus W^*$ is an isomorphism of $G$-representations.)

**Proposition 5.19.24.** Let $G$ be a group. Let $V$ and $W$ be two representations of $G$. Let $f : V \to W$ be a map. Then:

**(a)** If $f$ is a morphism of representations from $V$ to $W$, then its dual $f^*$ (as defined in Proposition 5.19.4 **(a)**) is a morphism of representations from $W^*$ to $V^*$.

**(b)** If $f$ is an isomorphism of representations from $V$ to $W$, then its dual $f^*$ is an isomorphism of representations from $W^*$ to $V^*$.

**Corollary 5.19.25.** Let $G$ be a group. Let $V$ and $W$ be two representations of $G$. If $V \cong W$, then $V^* \cong W^*$.

**Exercise 5.19.4. (a)** $\boxed{1}$ Prove Proposition 5.19.23.

**(b)** $\boxed{2}$ Prove Proposition 5.19.24 and Corollary 5.19.25.

**Remark 5.19.26.** Proposition 5.19.24 **(a)** (combined with the facts that $(f \circ g)^* = g^* \circ f^*$ and $\mathrm{id}^* = \mathrm{id}$) shows that we can define a "dualization functor" on the category of $G$-representations for any group $G$. This is a contravariant functor from the category of $G$-representations to itself. On objects, it acts by sending each $G$-representation $V$ to $V^*$. On morphisms, it acts by sending each morphism $f$ to $f^*$.

### 5.19.4. $G$-invariant bilinear forms and duals

Bilinear forms are often the tool of choice to understand duals of $\mathbf{k}$-modules. To understand duals of $G$-representations, we need a slightly restrictive kind of bilinear forms, namely the *$G$-invariant bilinear forms*:

**Definition 5.19.27.** Let $G$ be a group. Let $U$ and $V$ be two representations of $G$ over $\mathbf{k}$. Let $f : U \times V \to \mathbf{k}$ be a bilinear form. Then, $f$ is said to be *$G$-invariant* if and only if every $g \in G$ and $u \in U$ and $v \in V$ satisfy the equality

$$f(gu, gv) = f(u, v).$$

(You can rewrite this equality as $f(g \rightharpoonup u, g \rightharpoonup v) = f(u, v)$ if you wish.)

**Example 5.19.28.** Consider the natural representation $\mathbf{k}^n$ of the symmetric group $S_n$ (defined in Example 4.2.17). Then:

**(a)** Define the *dot product* $\langle u, v \rangle$ of two vectors $u = (u_1, u_2, \ldots, u_n) \in \mathbf{k}^n$ and $v = (v_1, v_2, \ldots, v_n) \in \mathbf{k}^n$ to be the scalar $\sum_{i=1}^{n} u_i v_i$. (This can also be rewritten as $u^T v$, if you identify the $1 \times 1$-matrix $u^T v$ with its only entry.) The map

$$\begin{aligned} \mathbf{k}^n \times \mathbf{k}^n &\to \mathbf{k}, \\ (u, v) &\mapsto \langle u, v \rangle \end{aligned}$$

(which sends a pair of vectors to their dot product) is a bilinear form, known as the *dot product* or the *standard scalar product* on $\mathbf{k}^n$. This form is $S_n$-invariant, since each $g \in S_n$ and any two vectors $u \in \mathbf{k}^n$ and $v \in \mathbf{k}^n$ satisfy the equality $\langle gu, gv \rangle = \langle u, v \rangle$. (Indeed, if we write $u$ and $v$ as $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$, then this equality rewrites as $\sum_{i=1}^{n} u_{g^{-1}(i)} v_{g^{-1}(i)} = \sum_{i=1}^{n} u_i v_i$, which is true because $g^{-1}$ is a permutation of $[n]$.)

**(b)** Another $S_n$-invariant bilinear form on the natural representation $\mathbf{k}^n$ is the map

$$\begin{aligned} \mathbf{k}^n \times \mathbf{k}^n &\to \mathbf{k}, \\ (u, v) &\mapsto (\operatorname{sum} u)\,(\operatorname{sum} v), \end{aligned}$$

where $\operatorname{sum} w$ denotes the sum of all entries of any vector $w \in \mathbf{k}^n$ (that is, $\operatorname{sum}(w_1, w_2, \ldots, w_n) := w_1 + w_2 + \cdots + w_n$). This follows easily from the fact that $\operatorname{sum}(gw) = \operatorname{sum} w$ for any $w \in \mathbf{k}^n$ and any $g \in S_n$.

**(c)** Of course, the zero form $\mathbf{k}^n \times \mathbf{k}^n \to \mathbf{k}$ (which sends every $(u, v)$ to 0) is also $S_n$-invariant.

**(d)** An example of a bilinear form that is not $S_n$-invariant is

$$\begin{aligned} &\mathbf{k}^n \times \mathbf{k}^n \to \mathbf{k}, \\ &\quad (u, v) \mapsto u_1 v_1 \qquad \left(\text{where } u = (u_1, u_2, \ldots, u_n) \text{ and } v = (v_1, v_2, \ldots, v_n)\right) \end{aligned}$$

(unless $n = 1$, in which case this form is the dot product form and thus is $S_n$-invariant).

**Exercise 5.19.5. (a)** $\boxed{1}$ Let $G$ be a group. Let $U$ and $V$ be two trivial $G$-representations. Prove that any bilinear form $f : U \times V \to \mathbf{k}$ is $G$-invariant.

**(b)** $\boxed{1}$ Consider the trivial representation $\mathbf{k}_{\text{triv}}$ and the sign representation $\mathbf{k}_{\text{sign}}$ of $S_n$ (see Example 4.2.16 and Example 4.2.18). Assume that 2 is invertible in $\mathbf{k}$, and that $n \geq 2$. Prove that the only $S_n$-invariant bilinear form

- $f : \mathbf{k}_{\mathrm{triv}} \times \mathbf{k}_{\mathrm{sign}} \to \mathbf{k}$ is the zero form (sending everything to 0).

**Exercise 5.19.6.** [1] Let $U$ and $V$ be two $\mathbf{k}$-modules.

**(a)** Prove that the set of all bilinear forms $f : U \times V \to \mathbf{k}$ is a $\mathbf{k}$-module (with respect to pointwise addition and scaling).

**(b)** Now assume that $U$ and $V$ are representations of a group $G$. Prove that the set of all $G$-invariant bilinear forms $f : U \times V \to \mathbf{k}$ is a $\mathbf{k}$-submodule of the $\mathbf{k}$-module of all bilinear forms $f : U \times V \to \mathbf{k}$.

**Exercise 5.19.7.** [1] Let $G$ be a group. Define a map $S : \mathbf{k}[G] \to \mathbf{k}[G]$ as in Corollary 5.17.26.

Let $U$ and $V$ be two $G$-representations. Let $f : U \times V \to \mathbf{k}$ be a bilinear form. Prove that $f$ is $G$-invariant if and only if every $\mathbf{a} \in \mathbf{k}[G]$ and every $u \in U$ and $v \in V$ satisfy

$$f(\mathbf{a} \cdot u,\ v) = f(u,\ S(\mathbf{a}) \cdot v).$$

As we said, $G$-invariant forms provide information about dual representations. This is due to the following easy fact:

**Proposition 5.19.29.** Let $G$ be a group. Let $U$ and $V$ be two $G$-representations. Let $f : U \times V \to \mathbf{k}$ be a $G$-invariant bilinear form. Then:

**(a)** The maps $f_L : U \to V^*$ and $f_R : V \to U^*$ (defined in Definition 5.19.7 for $W = \mathbf{k}$) are morphisms of $G$-representations.

**(b)** If $f_L$ is bijective, then $U \cong V^*$ as $G$-representations.

*Proof.* The form $f$ is $G$-invariant. In other words, every $g \in G$ and $u \in U$ and $v \in V$ satisfy the equality

$$f(gu,\ gv) = f(u,\ v). \tag{374}$$

**(a)** Let us prove the claim about $f_L$ only (the claim about $f_R$ is analogous).

So we must show that $f_L$ is a morphism of $G$-representations. Since we already know that $f_L$ is $\mathbf{k}$-linear, we must only show that $f_L$ is $G$-equivariant. In other words, we must show that $f_L(g \rightharpoonup u) = g \overset{*}{\rightharpoonup} f_L(u)$ for all $g \in G$ and $u \in U$, where $\overset{*}{\rightharpoonup}$ denotes the $G$-action on $V^*$ (as defined in Definition 5.19.14).

So let us fix $g \in G$ and $u \in U$. Then, for any $v \in V$, we have

$$\begin{aligned}
\left(f_L\left(g \rightharpoonup u\right)\right)(v) &= f\left(\underbrace{g \rightharpoonup u}_{=gu}, \underbrace{v}_{=gg^{-1}v}\right) && \text{(by the definition of } f_L\text{)} \\
&= f\left(gu,\ gg^{-1}v\right) \\
&= f\left(u,\ g^{-1}v\right) && \left(\text{by (374), applied to } g^{-1}v \text{ instead of } v\right) \\
&= \left(f_L\left(u\right)\right)\left(g^{-1}v\right) && \left(\begin{array}{c}\text{since the definition of } f_L \\ \text{yields } \left(f_L\left(u\right)\right)\left(g^{-1}v\right) = f\left(u,\ g^{-1}v\right)\end{array}\right)
\end{aligned}$$

and

$$\begin{aligned}
\left(g \overset{*}{\rightharpoonup} f_L\left(u\right)\right)(v) &= \left(f_L\left(u\right)\right)\left(g^{-1} \rightharpoonup v\right) && \text{(by (364), applied to } f_L\left(u\right) \text{ instead of } f\text{)} \\
&= \left(f_L\left(u\right)\right)\left(g^{-1}v\right) && \left(\text{since } g^{-1} \rightharpoonup v = g^{-1}v\right).
\end{aligned}$$

Comparing these two equalities, we obtain

$$\left(f_L\left(g \rightharpoonup u\right)\right)(v) = \left(g \overset{*}{\rightharpoonup} f_L\left(u\right)\right)(v) \qquad \text{for each } v \in V.$$

In other words, $f_L\left(g \rightharpoonup u\right) = g \overset{*}{\rightharpoonup} f_L\left(u\right)$.

Forget that we fixed $g$ and $u$. We thus have shown that $f_L\left(g \rightharpoonup u\right) = g \overset{*}{\rightharpoonup} f_L\left(u\right)$ for all $g \in G$ and $u \in U$. In other words, the map $f_L$ is $G$-equivariant. As we said, this completes our proof of Proposition 5.19.29 **(a)**.

**(b)** Assume that $f_L$ is bijective. Thus, $f_L$ is invertible. But Proposition 5.19.29 **(a)** shows that $f_L : U \to V^*$ is a morphism of $G$-representations. Hence, $f_L : U \to V^*$ is an invertible morphism of $G$-representations, thus an isomorphism (by Proposition 4.2.24). Hence, $U \cong V^*$. This proves Proposition 5.19.29 **(b)**. □

As an example of how Proposition 5.19.29 can be used, let us show that any finite permutation module of any group $G$ is isomorphic to its own dual:

**Theorem 5.19.30.** Let $G$ be a group, and let $X$ be a finite left $G$-set. Recall the permutation module $\mathbf{k}^{(X)}$ defined in Example 4.3.9; this is the left $\mathbf{k}\left[G\right]$-module on which $G$ acts by permuting the standard basis vectors (see the equality (106) for details).

Then, $\left(\mathbf{k}^{(X)}\right)^* \cong \mathbf{k}^{(X)}$ as $G$-representations.

*Proof.* As in Example 4.3.9, we denote the standard basis vector $e_x$ of $\mathbf{k}^{(X)}$ as $x$ for each $x \in X$. Then, $\left(x\right)_{x \in X}$ is the standard basis of $\mathbf{k}^{(X)}$. Thus, by Proposition 5.19.10, we can define a bilinear form $f : \mathbf{k}^{(X)} \times \mathbf{k}^{(X)} \to \mathbf{k}$ by setting

$$f\left(x, y\right) = \delta_{x,y} \qquad \text{for all } x \in X \text{ and } y \in X$$

(where $\delta_{x,y}$ denotes the Kronecker delta). Consider this bilinear form $f$.

We shall now show that the bilinear form $f$ is $G$-invariant. Indeed, by Definition 5.19.27, this amounts to showing that every $g \in G$ and $u \in \mathbf{k}^{(X)}$ and $v \in \mathbf{k}^{(X)}$ satisfy the equality $f(gu, gv) = f(u, v)$. So let us show this. Fix $g \in G$ and $u \in \mathbf{k}^{(X)}$ and $v \in \mathbf{k}^{(X)}$. We must prove the equality $f(gu, gv) = f(u, v)$. Both sides of this equality depend $\mathbf{k}$-linearly on each of $u$ and $v$ (since $f$ is $\mathbf{k}$-bilinear). Hence, in proving it, we can WLOG assume that both $u$ and $v$ belong to the standard basis $(x)_{x \in X}$ of $\mathbf{k}^{(X)}$. Assume this. Thus, $u = x$ and $v = y$ for some $x, y \in X$. Consider these $x, y$.

Now, the definition of $f$ yields

$$\begin{aligned}
f(gx, gy) &= \delta_{gx,gy} \\
&= \begin{cases} 1, & \text{if } gx = gy; \\ 0, & \text{if } gx \neq gy \end{cases} \qquad \left(\text{by the definition of the Kronecker delta}\right) \\
&= \begin{cases} 1, & \text{if } x = y; \\ 0, & \text{if } x \neq y \end{cases} \qquad \left( \begin{array}{c} \text{since } gx = gy \text{ if and only if } x = y \\ \text{(because } g \text{ is invertible)} \end{array} \right) \\
&= \delta_{x,y} \qquad \left(\text{by the definition of the Kronecker delta}\right) \\
&= f(x, y) \qquad \left(\text{by the definition of } f\right).
\end{aligned}$$

In other words, $f(gu, gv) = f(u, v)$ (since $u = x$ and $v = y$).

Forget that we fixed $u$ and $v$. We thus have shown that every $g \in G$ and $u \in \mathbf{k}^{(X)}$ and $v \in \mathbf{k}^{(X)}$ satisfy the equality $f(gu, gv) = f(u, v)$. In other words, the bilinear form $f$ is $G$-invariant.

Now, recall that $\mathbf{k}^{(X)}$ is a free $\mathbf{k}$-module with a finite basis $(x)_{x \in X}$ (it is finite since $X$ is finite). Let $(x^*)_{x \in X}$ be the dual basis of this basis $(x)_{x \in X}$. Then, $(x^*)_{x \in X}$ is a basis of the dual $\mathbf{k}$-module $\left(\mathbf{k}^{(X)}\right)^*$ (since the set $X$ is finite).

However, the map $f_L : \mathbf{k}^{(X)} \to \left(\mathbf{k}^{(X)}\right)^*$ (defined as in Definition 5.19.7) is easily seen to satisfy

$$f_L(x) = x^* \qquad \text{for each } x \in X$$

[129]. In other words, the map $f_L$ sends the entries $x$ of the basis $(x)_{x \in X}$ of $\mathbf{k}^{(X)}$ to the corresponding entries $x^*$ of the basis $(x^*)_{x \in X}$ of $\left(\mathbf{k}^{(X)}\right)^*$. Hence, this map $f_L$ is a $\mathbf{k}$-module isomorphism (since it is $\mathbf{k}$-linear and sends a basis to a

[129] *Proof.* Let $x \in X$. Then, for each $y \in X$, we have

$$\begin{aligned}
\left(f_L(x)\right)(y) &= f(x, y) \qquad \left(\text{by the definition of } f_L\right) \\
&= \delta_{x,y} \qquad \left(\text{by the definition of } f\right) \\
&= x^*(y)
\end{aligned}$$

(since the definition of a dual basis yields $x^*(y) = \delta_{x,y}$). In other words, $f_L(x) = x^*$, qed.

basis), thus bijective. Hence, Proposition 5.19.29 **(b)** (applied to $U = \mathbf{k}^{(X)}$ and $V = \mathbf{k}^{(X)}$) shows that $\mathbf{k}^{(X)} \cong \left(\mathbf{k}^{(X)}\right)^*$ as $G$-representations. Thus, Theorem 5.19.30 is proved. □

Two particular cases of Theorem 5.19.30 are worth mentioning. The first has already been stated in Exercise 5.19.2, but we shall now prove it anew by interpreting it as a special case of Theorem 5.19.30:

**Corollary 5.19.31.** Consider the natural representation $\mathbf{k}^n$ of the symmetric group $S_n$ (defined in Example 4.2.17). Then, $(\mathbf{k}^n)^* \cong \mathbf{k}^n$ as $S_n$-representations.

*Proof sketch.* I claim that the natural representation $\mathbf{k}^n$ of $S_n$ is a permutation module. Indeed, consider the natural action of $S_n$ on the set $[n] = \{1, 2, \ldots, n\}$ given by the rule

$$g \rightharpoonup i := g(i) \qquad \text{for all } g \in S_n \text{ and } i \in [n]$$

(this is the action from Example 4.1.12, applied to $A = [n]$). Thus, $[n]$ is a left $S_n$-set. The corresponding permutation module $\mathbf{k}^{([n])}$ (as defined in Example 4.3.9) is the free $\mathbf{k}$-module with basis $(e_i)_{i \in [n]} = (e_1, e_2, \ldots, e_n)$ and with a linear $S_n$-action given by

$$g \rightharpoonup e_i = e_{g \rightharpoonup i} = e_{g(i)} \qquad \text{for all } g \in S_n \text{ and } i \in [n].$$

But this is precisely the natural representation $\mathbf{k}^n$ of $S_n$ with its standard basis.

Thus, this natural representation $\mathbf{k}^n$ is the permutation module $\mathbf{k}^{([n])}$ corresponding to the left $S_n$-set $[n]$. But Proposition 5.19.30 (applied to $G = S_n$ and $X = [n]$) shows that $\left(\mathbf{k}^{([n])}\right)^* \cong \mathbf{k}^{([n])}$ as $S_n$-representations. Since the natural representation $\mathbf{k}^n$ is the permutation module $\mathbf{k}^{([n])}$, we can rewrite this as $(\mathbf{k}^n)^* \cong \mathbf{k}^n$ (as $S_n$-representations). This proves Corollary 5.19.31. □

The next particular case of Theorem 5.19.30 is the case of Young modules:

**Corollary 5.19.32.** Let $D$ be a diagram with $|D| = n$. Then, the Young module $\mathcal{M}^D$ satisfies $\left(\mathcal{M}^D\right)^* \cong \mathcal{M}^D$ as $S_n$-representations.

*Proof.* The Young module $\mathcal{M}^D$ is a permutation module. Indeed, recall (from Definition 5.3.10) that the set $\{n\text{-tabloids of shape } D\}$ is a left $S_n$-set. The permutation module corresponding to this $S_n$-set (as defined in Example 4.3.9) is precisely $\mathcal{M}^D$ (by Definition 5.3.15). In other words,

$$\mathcal{M}^D = \mathbf{k}^{(\{n\text{-tabloids of shape } D\})} \tag{375}$$

as $S_n$-representations. But Proposition 5.19.30 (applied to $G = S_n$ and $X = \{n\text{-tabloids of shape } D\}$) shows that

$$\left(\mathbf{k}^{(\{n\text{-tabloids of shape } D\})}\right)^* \cong \mathbf{k}^{(\{n\text{-tabloids of shape } D\})} \text{ as } S_n\text{-representations.}$$

In view of (375), we can rewrite this as follows:

$$\left(\mathcal{M}^D\right)^* \cong \mathcal{M}^D \text{ as } S_n\text{-representations.}$$

This proves Corollary 5.19.32. □

### 5.19.5. The dual of a Specht module: The case of a field

So the dual of a Young module is the Young module itself (up to a rather obvious isomorphism). We shall now study the dual of a Specht module, which is far less obvious. We begin with the case when $\mathbf{k}$ is a field.

Recall the map $\mathbf{r} : \mathbb{Z}^2 \to \mathbb{Z}^2$ (defined in Theorem 5.1.10), which sends each cell $(i, j) \in \mathbb{Z}^2$ to $(j, i)$. Recall also the sign-twist $V^{\mathrm{sign}}$ of an $S_n$-representation $V$, defined in Definition 5.18.2 or in Definition 5.18.3. Now we claim (generalizing [James78, Theorems 6.7 and 8.15]):

**Theorem 5.19.33.** Let $D$ be a diagram with $|D| = n$. Assume that $\mathbf{k}$ is a field. Then, the Specht module $\mathcal{S}^D$ satisfies

$$\left(\mathcal{S}^D\right)^* \cong \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}} \text{ as } S_n\text{-representations.}$$

I don't know if the "$\mathbf{k}$ is a field" assumption is actually necessary here; the proof below makes use of it, but I don't know of any counterexample when $\mathbf{k}$ is not a field. I will say more about this below, after proving Theorem 5.19.33 and Theorem 5.19.34.

*Proof of Theorem 5.19.33.* Let $\mathcal{A} = \mathbf{k}[S_n]$. Recall the antipode $S$ of $\mathbf{k}[S_n]$.

Furthermore, recall Definition 5.11.6. Thus, if an element $\mathbf{a} \in \mathcal{A}$ is written as $\mathbf{a} = \sum_{g \in S_n} \alpha_g g$ for some scalars $\alpha_g \in \mathbf{k}$, then $[w]\,\mathbf{a} := \alpha_w$ for any $w \in S_n$. We begin by showing a few properties of coefficients of elements of $\mathcal{A}$:

*Claim 1:* Let $\mathbf{a} \in \mathbf{k}[S_n]$ and $g \in S_n$. Then,

$$[1]\left(g\mathbf{a}g^{-1}\right) = [1]\,\mathbf{a}; \tag{376}$$

$$[1]\left(\mathbf{a}g^{-1}\right) = [g]\,\mathbf{a}; \tag{377}$$

$$[g]\,\mathbf{a} = [1]\left(gS\left(\mathbf{a}\right)\right). \tag{378}$$

*Proof of Claim 1.* All three equalities (376), (377) and (378) depend $\mathbf{k}$-linearly on $\mathbf{a}$ (since $S$ is a $\mathbf{k}$-linear map). Thus, by linearity, we can WLOG assume that $\mathbf{a}$ is a standard basis vector of $\mathbf{k}[S_n]$. Assume this. Thus, $\mathbf{a} = w$ for some $w \in S_n$. Consider this $w$. Hence, $S(\mathbf{a}) = S(w) = w^{-1}$ (by the definition of the antipode $S$).

Definition 5.11.6 yields that any two elements $u, v \in S_n$ satisfy $[u]\, v = \delta_{u,v}$ (Kronecker delta). Hence, $[g]\, w = \delta_{g,w}$ and $[1]\left(gw^{-1}\right) = \delta_{1,gw^{-1}} = \delta_{g,w}$ (since the equation $1 = gw^{-1}$ is equivalent to $g = w$). Comparing these two equalities, we obtain $[g]\, w = [1]\left(gw^{-1}\right)$. In other words, $[g]\, \mathbf{a} = [1]\left(gS(\mathbf{a})\right)$ (since $\mathbf{a} = w$ and $S(\mathbf{a}) = w^{-1}$). Thus, (378) is proved.

Recall again the fact that any two elements $u, v \in S_n$ satisfy $[u]\, v = \delta_{u,v}$. Hence, $[1]\left(wg^{-1}\right) = \delta_{1,wg^{-1}} = \delta_{g,w}$ (since the equation $1 = wg^{-1}$ holds if and only if $g = w$) and $[g]\, w = \delta_{g,w}$. Comparing these two equalities, we obtain $[1]\left(wg^{-1}\right) = [g]\, w$. In other words, $[1]\left(\mathbf{a}g^{-1}\right) = [g]\, \mathbf{a}$ (since $\mathbf{a} = w$). Thus, (377) is proved.

It remains to prove (376). For this, we observe that the equation $1 = gwg^{-1}$ is equivalent to $1 = w$ (because multiplying the equation $1 = gwg^{-1}$ by $g$ on the right transforms it into $g = gw$, which is clearly equivalent to $1 = w$). Now, recall once again the fact that any two elements $u, v \in S_n$ satisfy $[u]\, v = \delta_{u,v}$. Hence, $[1]\left(gwg^{-1}\right) = \delta_{1,gwg^{-1}} = \delta_{1,w}$ (since the equation $1 = gwg^{-1}$ is equivalent to $1 = w$) and $[1]\, w = \delta_{1,w}$. Comparing these two equalities, we obtain $[1]\left(gwg^{-1}\right) = [1]\, w$. In other words, $[1]\left(g\mathbf{a}g^{-1}\right) = [1]\, \mathbf{a}$ (since $\mathbf{a} = w$). Thus, (376) is proved.

Now, we have proved all three equalities (376), (377) and (378). Thus, Claim 1 is proved. □

Fix an $n$-tableau $T$ of shape $D$ (such an $n$-tableau clearly exists, since $|D| = n$). Consider the elements

$$\mathbf{E}_T := \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \qquad \text{and} \qquad \mathbf{F}_T := \nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T}$$

of $\mathcal{A}$. Proposition 5.11.19 **(a)** says that the antipode $S$ of $\mathbf{k}[S_n]$ satisfies $S(\mathbf{E}_T) = \mathbf{F}_T$. Applying the map $S$ to this equality, we obtain $S(S(\mathbf{E}_T)) = S(\mathbf{F}_T)$. But $S(S(\mathbf{E}_T)) = \mathbf{E}_T$ (since Definition 5.19.16 **(b)** says that $S \circ S = \mathrm{id}$). Comparing these two equalities, we find $S(\mathbf{F}_T) = \mathbf{E}_T$. Moreover, recall (from Definition 5.19.16 **(a)**) that $S$ is an algebra anti-morphism. Lemma 5.18.15 yields that

$$\mathcal{S}^D \cong \mathcal{A}\mathbf{E}_T \qquad \text{and} \qquad \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}} \cong \mathcal{A}\mathbf{F}_T \tag{379}$$

as $S_n$-representations. From $\mathcal{S}^D \cong \mathcal{A}\mathbf{E}_T$, we obtain $\left(\mathcal{S}^D\right)^* \cong \left(\mathcal{A}\mathbf{E}_T\right)^*$ (by Corollary 5.19.25).

We must prove that $\left(\mathcal{S}^D\right)^* \cong \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}}$ as $S_n$-representations. In view of (379), it suffices to show that $\left(\mathcal{A}\mathbf{E}_T\right)^* \cong \mathcal{A}\mathbf{F}_T$ as $S_n$-representations (because

once this is proved, then we will immediately obtain $\left(\mathcal{S}^D\right)^* \cong \left(\mathcal{A}\mathbf{E}_T\right)^* \cong \mathcal{A}\mathbf{F}_T \cong \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}}$ by (379)). In order to show this, we will construct an $S_n$-invariant bilinear form $f : \mathcal{A}\mathbf{E}_T \times \mathcal{A}\mathbf{F}_T \to \mathbf{k}$ such that the map $f_R : \mathcal{A}\mathbf{F}_T \to \left(\mathcal{A}\mathbf{E}_T\right)^*$ is an isomorphism of $S_n$-representations.

We define this bilinear form $f : \mathcal{A}\mathbf{E}_T \times \mathcal{A}\mathbf{F}_T \to \mathbf{k}$ by the equality

$$f\left(u\mathbf{E}_T,\ v\mathbf{F}_T\right) := [1]\left(u\mathbf{E}_T S\left(v\right)\right) \qquad \text{for all } u, v \in \mathcal{A}. \tag{380}$$

But first, we must show that this form $f$ is well-defined, i.e., that the equality (380) really uniquely determines each value of $f$ and that the resulting map $f : \mathcal{A}\mathbf{E}_T \times \mathcal{A}\mathbf{F}_T \to \mathbf{k}$ is a bilinear form.

In other words, we must prove the following:

*Claim 2:* The bilinear form $f : \mathcal{A}\mathbf{E}_T \times \mathcal{A}\mathbf{F}_T \to \mathbf{k}$ is well-defined.

*Proof of Claim 2.* First, we shall show that the equality (380) defines a valid **map** $f : \mathcal{A}\mathbf{E}_T \times \mathcal{A}\mathbf{F}_T \to \mathbf{k}$. (We will prove its bilinearity later.)

Each element of $\mathcal{A}\mathbf{E}_T$ can be written as $u\mathbf{E}_T$ for some $u \in \mathcal{A}$. Each element of $\mathcal{A}\mathbf{F}_T$ can be written as $v\mathbf{F}_T$ for some $v \in \mathcal{A}$. Thus, if we want to compute $f\left(a, b\right)$ for some $a \in \mathcal{A}\mathbf{E}_T$ and some $b \in \mathcal{A}\mathbf{F}_T$, then we can write $a$ as $u\mathbf{E}_T$ and write $b$ as $v\mathbf{F}_T$, and then compute the value $f\left(a, b\right)$ as follows:

$$f\left(\underbrace{a}_{=u\mathbf{E}_T}, \underbrace{b}_{=v\mathbf{F}_T}\right) = f\left(u\mathbf{E}_T,\ v\mathbf{F}_T\right) = [1]\left(u\mathbf{E}_T S\left(v\right)\right) \qquad \left(\text{by (380)}\right).$$

Thus, the equality (380) allows us to compute each value $f\left(a, b\right)$ of $f$. However, this does not yet mean that $f$ is well-defined: Indeed, when we choose the elements $u$ and $v$ satisfying $a = u\mathbf{E}_T$ and $b = v\mathbf{F}_T$, we have some freedom (they are not uniquely determined by $a$ and $b$, since (e.g.) one single $a$ can be written as $u\mathbf{E}_T$ for several values of $u$), and it is conceivable that the resulting value $[1]\left(u\mathbf{E}_T S\left(v\right)\right)$ could depend on these choices; but then we would get different values for the same $f\left(a, b\right)$ depending on our choices of $u$ and $v$, and this would render $f$ undefined.

So we need to show that this cannot happen. In other words, we need to show that the element $[1]\left(u\mathbf{E}_T S\left(v\right)\right)$ on the right hand side of (380) depends only on $u\mathbf{E}_T$ and $v\mathbf{F}_T$ but not on the specific values of $u$ and $v$. In other words, we need to show that if $u, v \in \mathcal{A}$ and $u', v' \in \mathcal{A}$ satisfy $u\mathbf{E}_T = u'\mathbf{E}_T$ and $v\mathbf{F}_T = v'\mathbf{F}_T$, then $[1]\left(u\mathbf{E}_T S\left(v\right)\right) = [1]\left(u'\mathbf{E}_T S\left(v'\right)\right)$.

Let us show this. Indeed, let $u, v \in \mathcal{A}$ and $u', v' \in \mathcal{A}$ be such that $u\mathbf{E}_T = u'\mathbf{E}_T$ and $v\mathbf{F}_T = v'\mathbf{F}_T$. Recall that $S$ is an algebra anti-morphism. Thus, $S\left(v\mathbf{F}_T\right) = \underbrace{S\left(\mathbf{F}_T\right)}_{=\mathbf{E}_T} S\left(v\right) = \mathbf{E}_T S\left(v\right)$. Likewise, $S\left(v'\mathbf{F}_T\right) = \mathbf{E}_T S\left(v'\right)$. From $S\left(v\mathbf{F}_T\right) = \mathbf{E}_T S\left(v\right)$, we obtain

$$\mathbf{E}_T S\left(v\right) = S\left(\underbrace{v\mathbf{F}_T}_{=v'\mathbf{F}_T}\right) = S\left(v'\mathbf{F}_T\right) = \mathbf{E}_T S\left(v'\right).$$

But now,

$$[1]\left(\underbrace{u\mathbf{E}_T}_{=u'\mathbf{E}_T} S(v)\right) = [1]\left(u'\underbrace{\mathbf{E}_T S(v)}_{=\mathbf{E}_T S(v')}\right) = [1]\left(u'\mathbf{E}_T S(v')\right).$$

Forget that we fixed $u, v$ and $u', v'$. Thus, we have shown that if $u, v \in \mathcal{A}$ and $u', v' \in \mathcal{A}$ satisfy $u\mathbf{E}_T = u'\mathbf{E}_T$ and $v\mathbf{F}_T = v'\mathbf{F}_T$, then $[1](u\mathbf{E}_T S(v)) = [1](u'\mathbf{E}_T S(v'))$. In other words, the right hand side of (380) depends only on $u\mathbf{E}_T$ and on $v\mathbf{F}_T$ but not on the specific choices of $u$ and $v$. Therefore, the map $f : \mathcal{A}\mathbf{E}_T \times \mathcal{A}\mathbf{F}_T \to \mathbf{k}$ is well-defined (in the sense that its output value does not depend on how exactly we write its two inputs as $u\mathbf{E}_T$ and $v\mathbf{F}_T$).

It remains to prove that this map is a bilinear form. In other words, it remains to prove that it is $\mathbf{k}$-bilinear. But this is straightforward[130]. Thus, Claim 2 is proved. □

*Claim 3:* The bilinear form $f : \mathcal{A}\mathbf{E}_T \times \mathcal{A}\mathbf{F}_T \to \mathbf{k}$ is $S_n$-invariant.

[130]Here is the proof in some detail:

We show that the map $f$ is $\mathbf{k}$-bilinear by verifying that it satisfies all six axioms from Remark 5.19.8. Let us only verify the first of these six axioms (the other five are analogous).

Thus, we must show that $f(a,\ b_1+b_2) = f(a,b_1) + f(a,b_2)$ for all $a \in \mathcal{A}\mathbf{E}_T$ and $b_1, b_2 \in \mathcal{A}\mathbf{F}_T$.

Indeed, let $a \in \mathcal{A}\mathbf{E}_T$ and $b_1, b_2 \in \mathcal{A}\mathbf{F}_T$. Then, we can write $a$ as $a = u\mathbf{E}_T$ for some $u \in \mathcal{A}$ (since $a \in \mathcal{A}\mathbf{E}_T$). Consider this $u$. Furthermore, we can write $b_1$ as $b_1 = v_1\mathbf{F}_T$ for some $v_1 \in \mathcal{A}$ (since $b_1 \in \mathcal{A}\mathbf{F}_T$), and we can write $b_2$ as $b_2 = v_2\mathbf{F}_T$ for some $v_2 \in \mathcal{A}$ (since $b_2 \in \mathcal{A}\mathbf{F}_T$). Consider these $v_1$ and $v_2$.

From $a = u\mathbf{E}_T$ and $b_1 = v_1\mathbf{F}_T$, we obtain

$$f(a,b_1) = f(u\mathbf{E}_T,\ v_1\mathbf{F}_T) = [1](u\mathbf{E}_T S(v_1)) \qquad \text{(by (380), applied to } v = v_1\text{)}.$$

In other words, $[1](u\mathbf{E}_T S(v_1)) = f(a,b_1)$. Likewise, $[1](u\mathbf{E}_T S(v_2)) = f(a,b_2)$. Furthermore, adding the equalities $b_1 = v_1\mathbf{F}_T$ and $b_2 = v_2\mathbf{F}_T$ together, we obtain $b_1 + b_2 = v_1\mathbf{F}_T + v_2\mathbf{F}_T = (v_1+v_2)\mathbf{F}_T$. From this and from $a = u\mathbf{E}_T$, we obtain

$$\begin{aligned}
f(a,\ b_1+b_2) &= f(u\mathbf{E}_T,\ (v_1+v_2)\mathbf{F}_T) \\
&= [1]\left(u\mathbf{E}_T \qquad \underbrace{S(v_1+v_2)}_{\substack{=S(v_1)+S(v_2) \\ \text{(since the map } S \text{ is } \mathbf{k}\text{-linear)}}}\right) \qquad \text{(by (380), applied to } v = v_1+v_2\text{)} \\
&= [1]\underbrace{(u\mathbf{E}_T(S(v_1)+S(v_2)))}_{=u\mathbf{E}_T S(v_1)+u\mathbf{E}_T S(v_2)} = [1](u\mathbf{E}_T S(v_1) + u\mathbf{E}_T S(v_2)) \\
&= \underbrace{[1](u\mathbf{E}_T S(v_1))}_{=f(a,b_1)} + \underbrace{[1](u\mathbf{E}_T S(v_2))}_{=f(a,b_2)} = f(a,b_1) + f(a,b_2).
\end{aligned}$$

Forget that we fixed $a, b_1, b_2$. We thus have shown that $f(a,\ b_1+b_2) = f(a,b_1) + f(a,b_2)$ for all $a \in \mathcal{A}\mathbf{E}_T$ and $b_1, b_2 \in \mathcal{A}\mathbf{F}_T$. This is the first of the six axioms that must hold in order for $f$ to be $\mathbf{k}$-bilinear. The remaining five axioms are just as easy to check using the same kind of reasoning. Thus, we have shown that $f$ is $\mathbf{k}$-bilinear.

*Proof of Claim 3.* We must show that every $g \in S_n$ and $a \in \mathcal{A}\mathbf{E}_T$ and $b \in \mathcal{A}\mathbf{F}_T$ satisfy the equality $f(ga,\ gb) = f(a,\ b)$ (since this is what it means for $f$ to be $S_n$-invariant, according to Definition 5.19.27).

So let us prove this. Fix $g \in S_n$ and $a \in \mathcal{A}\mathbf{E}_T$ and $b \in \mathcal{A}\mathbf{F}_T$. We must show that $f(ga,\ gb) = f(a,\ b)$.

Write $a$ as $a = u\mathbf{E}_T$ for some $u \in \mathcal{A}$ (we can do this, since $a \in \mathcal{A}\mathbf{E}_T$). Write $b$ as $b = v\mathbf{F}_T$ for some $v \in \mathcal{A}$ (we can do this, since $b \in \mathcal{A}\mathbf{F}_T$). Since $S$ is an algebra anti-morphism, we have $S(gv) = S(v)\,S(g) = S(v)\,g^{-1}$ (since the definition of $S$ yields $S(g) = g^{-1}$).

However, (380) (applied to $gu$ and $gv$ instead of $u$ and $v$) yields

$$\begin{aligned} f(gu\mathbf{E}_T,\ gv\mathbf{F}_T) &= [1]\left(gu\mathbf{E}_T\ \underbrace{S(gv)}_{=S(v)g^{-1}}\right) = [1]\left(gu\mathbf{E}_T S(v)\,g^{-1}\right) \\ &= [1]\left(u\mathbf{E}_T S(v)\right) \qquad \text{(by (376), applied to } \mathbf{a} = u\mathbf{E}_T S(v)\text{)} \\ &= f(u\mathbf{E}_T,\ v\mathbf{F}_T) \qquad \text{(by (380))}. \end{aligned}$$

In view of $a = u\mathbf{E}_T$ and $b = v\mathbf{F}_T$, we can rewrite this as $f(ga,\ gb) = f(a,\ b)$. But this is precisely what we desired to show. Thus, Claim 3 is proved. □

Now, recall that our bilinear form $f : \mathcal{A}\mathbf{E}_T \times \mathcal{A}\mathbf{F}_T \to \mathbf{k}$ gives rise to two $\mathbf{k}$-linear maps $f_L : \mathcal{A}\mathbf{E}_T \to \operatorname{Hom}(\mathcal{A}\mathbf{F}_T, \mathbf{k}) = (\mathcal{A}\mathbf{F}_T)^*$ and $f_R : \mathcal{A}\mathbf{F}_T \to \operatorname{Hom}(\mathcal{A}\mathbf{E}_T, \mathbf{k}) = (\mathcal{A}\mathbf{E}_T)^*$, as explained in Definition 5.19.7. We shall show that these two maps $f_L$ and $f_R$ are bijective. Indeed, we first claim that they are injective:

*Claim 4:* The $\mathbf{k}$-linear map $f_L : \mathcal{A}\mathbf{E}_T \to (\mathcal{A}\mathbf{F}_T)^*$ is injective.

*Proof of Claim 4.* The map $f_L$ is $\mathbf{k}$-linear. Thus, in order to prove its injectivity, it suffices to show that its kernel $\operatorname{Ker}(f_L)$ is 0. In other words, it suffices to show that $\mathbf{a} = 0$ for each $\mathbf{a} \in \operatorname{Ker}(f_L)$. This is what we shall now show.

Fix $\mathbf{a} \in \operatorname{Ker}(f_L)$. Thus, $\mathbf{a} \in \mathcal{A}\mathbf{E}_T$ and $f_L(\mathbf{a}) = 0$. We must show that $\mathbf{a} = 0$.

Write $\mathbf{a}$ as $\mathbf{a} = u\mathbf{E}_T$ for some $u \in \mathcal{A}$ (this can be done, since $\mathbf{a} \in \mathcal{A}\mathbf{E}_T$).

Let $w \in S_n$ be any permutation. We shall show that $[w]\,\mathbf{a} = 0$. Indeed, we have $[1]\left(\mathbf{a}w^{-1}\right) = [w]\,\mathbf{a}$ (by (377), applied to $g = w$). Furthermore, the definition of the antipode $S$ yields $S(w) = w^{-1}$. We have $w\mathbf{F}_T \in \mathcal{A}\mathbf{F}_T$ (since $w \in \mathcal{A}$). But the definition of $f_L$ yields

$$\begin{aligned} (f_L(\mathbf{a}))(w\mathbf{F}_T) &= f\left(\underbrace{\mathbf{a}}_{=u\mathbf{E}_T},\ w\mathbf{F}_T\right) = f(u\mathbf{E}_T,\ w\mathbf{F}_T) \\ &= [1]\left(\underbrace{u\mathbf{E}_T}_{=\mathbf{a}}\ \underbrace{S(w)}_{=w^{-1}}\right) \qquad \text{(by (380), applied to } v = w\text{)} \\ &= [1]\left(\mathbf{a}w^{-1}\right) = [w]\,\mathbf{a}, \end{aligned}$$

so that

$$[w]\,\mathbf{a} = \underbrace{(f_L(\mathbf{a}))}_{=0}(w\mathbf{F}_T) = 0\,(w\mathbf{F}_T) = 0.$$

Forget that we fixed $w$. We thus have shown that $[w]\,\mathbf{a} = 0$ for all $w \in S_n$. In other words, all coordinates of $\mathbf{a}$ are 0 (when $\mathbf{a}$ is expanded in the standard basis of $\mathbf{k}[S_n]$). Hence, $\mathbf{a} = 0$. As explained above, this completes the proof of Claim 4. □

*Claim 5:* The $\mathbf{k}$-linear map $f_R : \mathcal{A}\mathbf{F}_T \to (\mathcal{A}\mathbf{E}_T)^*$ is injective.

*Proof of Claim 5.* The map $f_R$ is $\mathbf{k}$-linear. Thus, in order to prove its injectivity, it suffices to show that its kernel $\operatorname{Ker}(f_R)$ is 0. In other words, it suffices to show that $\mathbf{a} = 0$ for each $\mathbf{a} \in \operatorname{Ker}(f_R)$. This is what we shall now show.

Fix $\mathbf{a} \in \operatorname{Ker}(f_R)$. Thus, $\mathbf{a} \in \mathcal{A}\mathbf{F}_T$ and $f_R(\mathbf{a}) = 0$. We must show that $\mathbf{a} = 0$.

Write $\mathbf{a}$ as $\mathbf{a} = v\mathbf{F}_T$ for some $v \in \mathcal{A}$ (this can be done, since $\mathbf{a} \in \mathcal{A}\mathbf{F}_T$). Thus,

$$\begin{aligned} S(\mathbf{a}) &= S(v\mathbf{F}_T) = \underbrace{S(\mathbf{F}_T)}_{=\mathbf{E}_T} S(v) \qquad \text{(since } S \text{ is an algebra anti-morphism)} \\ &= \mathbf{E}_T S(v). \end{aligned} \tag{381}$$

Let $w \in S_n$ be any permutation. We shall show that $[w]\,\mathbf{a} = 0$. Indeed, we have $[w]\,\mathbf{a} = [1]\,(wS(\mathbf{a}))$ (by (378), applied to $g = w$). We have $w\mathbf{E}_T \in \mathcal{A}\mathbf{E}_T$ (since $w \in \mathcal{A}$). But the definition of $f_R$ yields

$$\begin{aligned} (f_R(\mathbf{a}))(w\mathbf{E}_T) &= f\left(w\mathbf{E}_T,\ \underbrace{\mathbf{a}}_{=v\mathbf{F}_T}\right) = f(w\mathbf{E}_T,\ v\mathbf{F}_T) \\ &= [1]\left(w\underbrace{\mathbf{E}_T S(v)}_{\substack{=S(\mathbf{a}) \\ \text{(by (381))}}}\right) \qquad \text{(by (380), applied to } u = w) \\ &= [1]\,(wS(\mathbf{a})) = [w]\,\mathbf{a} \qquad \text{(since } [w]\,\mathbf{a} = [1]\,(wS(\mathbf{a}))). \end{aligned}$$

Hence,

$$[w]\,\mathbf{a} = \underbrace{(f_R(\mathbf{a}))}_{=0}(w\mathbf{E}_T) = 0\,(w\mathbf{E}_T) = 0.$$

Forget that we fixed $w$. We thus have shown that $[w]\,\mathbf{a} = 0$ for all $w \in S_n$. In other words, all coordinates of $\mathbf{a}$ are 0 (when $\mathbf{a}$ is expanded in the standard basis of $\mathbf{k}[S_n]$). Hence, $\mathbf{a} = 0$. As we explained above, this completes the proof of Claim 5. □

Now, recall that $\mathbf{k}$ is a field. Hence, $\mathcal{A} = \mathbf{k}[S_n]$ is a finite-dimensional $\mathbf{k}$-vector space (of dimension $|S_n| = n!$). Its subspaces $\mathcal{A}\mathbf{E}_T$ and $\mathcal{A}\mathbf{F}_T$ are therefore finite-dimensional as well. Moreover, the $\mathbf{k}$-linear maps $f_L : \mathcal{A}\mathbf{E}_T \to (\mathcal{A}\mathbf{F}_T)^*$ and $f_R : \mathcal{A}\mathbf{F}_T \to (\mathcal{A}\mathbf{E}_T)^*$ are injective (by Claim 4 and Claim 5). Hence, Lemma 5.19.13 (applied to $U = \mathcal{A}\mathbf{E}_T$ and $V = \mathcal{A}\mathbf{F}_T$ and $\alpha = f_L$ and $\beta = f_R$) yields that these two maps $f_L$ and $f_R$ are bijective. In other words, the maps $f_L$ and $f_R$ are invertible.

But the bilinear form $f$ is $S_n$-invariant (by Claim 3), and thus the map $f_R$ is a morphism of $S_n$-representations (by Proposition 5.19.29 **(a)**, applied to $G = S_n$ and $U = \mathcal{A}\mathbf{E}_T$ and $V = \mathcal{A}\mathbf{F}_T$). Since $f_R$ is invertible, we thus conclude that $f_R$ is an isomorphism of $S_n$-representations (since an invertible morphism of representations is an isomorphism). Hence, $(\mathcal{A}\mathbf{E}_T)^* \cong \mathcal{A}\mathbf{F}_T$ as $S_n$-representations. As we explained above, this proves Theorem 5.19.33.

(The above proof is somewhat inspired by Fayers's work [Fayers03, §4], though it is more direct.) □

### 5.19.6. The dual of a Specht module: The case of a skew Young diagram

Can we lift the "$\mathbf{k}$ is a field" requirement in Theorem 5.19.33? As I said, I don't know – in general. At least, we can lift it when $D$ is a skew Young diagram:

**Theorem 5.19.34.** Let $D$ be a skew Young diagram with $|D| = n$**.** Then, the Specht module $\mathcal{S}^D$ satisfies

$$\left(\mathcal{S}^D\right)^* \cong \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}} \text{ as } S_n\text{-representations.}$$

*Proof of Theorem 5.19.34.* We proceed as in the above proof of Theorem 5.19.33, up until the point where we used the assumption "$\mathbf{k}$ is a field". Since we no longer have this assumption, we must find a new reason why $f_R$ is bijective.

For any diagram $E$, we let $\mathrm{SYT}(E)$ be the set of all standard $n$-tableaux of shape $E$. We shall first show that both $\mathbf{k}$-modules $\mathcal{A}\mathbf{E}_T$ and $\mathcal{A}\mathbf{F}_T$ have finite bases indexed by this set $\mathrm{SYT}(D)$.

Indeed, recall Definition 5.15.2 and Definition 5.15.4. Recall also that if $P$ is any $n$-tableau of shape $D$, then $P\mathbf{r}$ is an $n$-tableau of shape $\mathbf{r}(D)$ (see Definition 5.15.10). Thus, in particular, $T\mathbf{r}$ is an $n$-tableau of shape $\mathbf{r}(D)$. Now, for each $P \in \mathrm{SYT}(D)$, we define the two elements

$$\mathbf{U}_P := \mathbf{E}_{P,T} \qquad \text{and} \tag{382}$$

$$\mathbf{V}_P := T_{\mathrm{sign}}\left(\mathbf{E}_{P\mathbf{r},T\mathbf{r}}\right) \tag{383}$$

of $\mathcal{A}$. Recall that $T_{\mathrm{sign}}$ is a $\mathbf{k}$-algebra automorphism of $\mathbf{k}[S_n]$ (by Theorem 3.11.5 **(a)**) and satisfies

$$T_{\mathrm{sign}}\left(\mathbf{E}_{T\mathbf{r}}\right) = \mathbf{F}_T$$

(as we saw in the proof of Lemma 5.18.15).

Recall that $D$ is a skew Young diagram. Thus, $\mathbf{r}(D)$ is a skew Young diagram as well (by Remark 5.1.12)[131].

Proposition 5.15.12 **(c)** (with the variable $T$ renamed as $Q$) says that the map

$$\begin{aligned} \mathrm{SYT}(D) &\to \mathrm{SYT}(\mathbf{r}(D)), \\ Q &\mapsto Q\mathbf{r} \end{aligned} \tag{384}$$

is a bijection.

Now, we claim the following:[132]

*Claim 6:* The family $(\mathbf{U}_P)_{P\in\mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_T$.

*Proof of Claim 6.* From (382), we obtain $(\mathbf{U}_P)_{P\in\mathrm{SYT}(D)} = (\mathbf{E}_{P,T})_{P\in\mathrm{SYT}(D)}$.

But Proposition 5.15.9 shows that the family $(\mathbf{E}_{P,T})_{P\in\mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_T$ (since $D$ is a skew Young diagram). In view of $(\mathbf{U}_P)_{P\in\mathrm{SYT}(D)} = (\mathbf{E}_{P,T})_{P\in\mathrm{SYT}(D)}$, we can rewrite this as follows: The family $(\mathbf{U}_P)_{P\in\mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_T$. This proves Claim 6. □

*Claim 7:* The family $\left(\mathbf{V}_Q\right)_{Q\in\mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{F}_T$.

*Proof of Claim 7.* For each $Q \in \mathrm{SYT}(D)$, we have $\mathbf{V}_Q = T_{\mathrm{sign}}\left(\mathbf{E}_{Q\mathbf{r},T\mathbf{r}}\right)$ (by (383), applied to $P = Q$). In other words, $\left(\mathbf{V}_Q\right)_{Q\in\mathrm{SYT}(D)} = \left(T_{\mathrm{sign}}\left(\mathbf{E}_{Q\mathbf{r},T\mathbf{r}}\right)\right)_{Q\in\mathrm{SYT}(D)}$.

Furthermore, $T_{\mathrm{sign}}(\mathbf{k}[S_n]) = \mathbf{k}[S_n]$ (since $T_{\mathrm{sign}}$ is an automorphism). In other words, $T_{\mathrm{sign}}(\mathcal{A}) = \mathcal{A}$ (since $\mathcal{A} = \mathbf{k}[S_n]$). Now, since $T_{\mathrm{sign}}$ is a $\mathbf{k}$-algebra morphism, we have

$$T_{\mathrm{sign}}(\mathcal{A}\mathbf{E}_{T\mathbf{r}}) = \underbrace{T_{\mathrm{sign}}(\mathcal{A})}_{=\mathcal{A}} \cdot \underbrace{T_{\mathrm{sign}}(\mathbf{E}_{T\mathbf{r}})}_{=\mathbf{F}_T} = \mathcal{A}\mathbf{F}_T.$$

But $T\mathbf{r}$ is an $n$-tableau of shape $\mathbf{r}(D)$ (since $T$ is an $n$-tableau of shape $D$). Hence, Proposition 5.15.9 (applied to $\mathbf{r}(D)$ and $T\mathbf{r}$ instead of $D$ and $T$) shows that the family $(\mathbf{E}_{P,T\mathbf{r}})_{P\in\mathrm{SYT}(\mathbf{r}(D))}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_{T\mathbf{r}}$ (since $\mathbf{r}(D)$ is a skew Young diagram).

Using the bijection (384), we can reindex this family as $\left(\mathbf{E}_{Q\mathbf{r},T\mathbf{r}}\right)_{Q\in\mathrm{SYT}(D)}$ (by substituting $Q\mathbf{r}$ for $P$ in this family). Thus, we conclude that the latter family $\left(\mathbf{E}_{Q\mathbf{r},T\mathbf{r}}\right)_{Q\in\mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_{T\mathbf{r}}$ as well (since reindexing a basis using a bijection yields a basis again).

But $T_{\mathrm{sign}}$ is a $\mathbf{k}$-module isomorphism (being a $\mathbf{k}$-algebra automorphism), and sends the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_{T\mathbf{r}}$ to $\mathcal{A}\mathbf{F}_T$ (since $T_{\mathrm{sign}}(\mathcal{A}\mathbf{E}_{T\mathbf{r}}) = \mathcal{A}\mathbf{F}_T$). Hence, it sends

[131] See the proof of Lemma 5.15.17 for the details of this argument.

[132] We are numbering the claims starting from 6 since we are continuing the proof of Theorem 5.19.33.

any basis of $\mathcal{A}\mathbf{E}_{T\mathbf{r}}$ to a basis of $\mathcal{A}\mathbf{F}_T$ (since a $\mathbf{k}$-module isomorphism sends any basis to a basis). Therefore, since $\left(\mathbf{E}_{Q\mathbf{r},T\mathbf{r}}\right)_{Q\in\mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_{T\mathbf{r}}$, we conclude that the family $\left(T_{\mathrm{sign}}\left(\mathbf{E}_{Q\mathbf{r},T\mathbf{r}}\right)\right)_{Q\in\mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{F}_T$ (since this family $\left(T_{\mathrm{sign}}\left(\mathbf{E}_{Q\mathbf{r},T\mathbf{r}}\right)\right)_{Q\in\mathrm{SYT}(D)}$ is the image of $\left(\mathbf{E}_{Q\mathbf{r},T\mathbf{r}}\right)_{Q\in\mathrm{SYT}(D)}$ under $T_{\mathrm{sign}}$). In view of $\left(\mathbf{V}_Q\right)_{Q\in\mathrm{SYT}(D)} = \left(T_{\mathrm{sign}}\left(\mathbf{E}_{Q\mathbf{r},T\mathbf{r}}\right)\right)_{Q\in\mathrm{SYT}(D)}$, we can rewrite this as follows: The family $\left(\mathbf{V}_Q\right)_{Q\in\mathrm{SYT}(D)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{F}_T$. This proves Claim 7. □

We have now found finite bases $\left(\mathbf{U}_P\right)_{P\in\mathrm{SYT}(D)}$ and $\left(\mathbf{V}_Q\right)_{Q\in\mathrm{SYT}(D)}$ of the two $\mathbf{k}$-modules $\mathcal{A}\mathbf{E}_T$ and $\mathcal{A}\mathbf{F}_T$, respectively (in Claim 6 and Claim 7). Thus, we can construct the Gram matrix[133]

$$\left(f\left(\mathbf{U}_P,\ \mathbf{V}_Q\right)\right)_{(P,Q)\in\mathrm{SYT}(D)\times\mathrm{SYT}(D)}$$

of the bilinear form $f : \mathcal{A}\mathbf{E}_T \times \mathcal{A}\mathbf{F}_T \to \mathbf{k}$ with respect to these two bases. Let us denote this matrix by $M_{\mathbf{k}}$, in order to make its dependence on $\mathbf{k}$ explicit. Note that this Gram matrix $M_{\mathbf{k}}$ is a square matrix, since both its rows and its columns are indexed by $\mathrm{SYT}\left(D\right)$. Thus, it has a determinant $\det\left(M_{\mathbf{k}}\right)$. This determinant determines whether $f_R$ is invertible:

> *Claim 8:* The map $f_R : \mathcal{A}\mathbf{F}_T \to \left(\mathcal{A}\mathbf{E}_T\right)^*$ is invertible if and only if the element $\det\left(M_{\mathbf{k}}\right)$ of $\mathbf{k}$ is invertible.

*Proof of Claim 8.* Proposition 5.19.11 **(b)** (applied to $U = \mathcal{A}\mathbf{E}_T$ and $V = \mathcal{A}\mathbf{F}_T$ and $I = \mathrm{SYT}\left(D\right)$ and $J = \mathrm{SYT}\left(D\right)$ and $\left(u_i\right)_{i\in I} = \left(\mathbf{U}_P\right)_{P\in\mathrm{SYT}(D)}$ and $\left(v_j\right)_{j\in J} = \left(\mathbf{V}_Q\right)_{Q\in\mathrm{SYT}(D)}$) shows that the $\mathbf{k}$-linear map $f_R : \mathcal{A}\mathbf{F}_T \to \left(\mathcal{A}\mathbf{E}_T\right)^*$ is invertible if and only if the matrix $\left(f\left(\mathbf{U}_P,\ \mathbf{V}_Q\right)\right)_{(P,Q)\in\mathrm{SYT}(D)\times\mathrm{SYT}(D)} \in \mathbf{k}^{\mathrm{SYT}(D)\times\mathrm{SYT}(D)}$ is invertible. Hence, we have the following chain of logical equivalences:

$$\begin{aligned}
&\left(\text{the map } f_R : \mathcal{A}\mathbf{F}_T \to \left(\mathcal{A}\mathbf{E}_T\right)^* \text{ is invertible}\right) \\
&\iff \left(\text{the matrix } \left(f\left(\mathbf{U}_P,\ \mathbf{V}_Q\right)\right)_{(P,Q)\in\mathrm{SYT}(D)\times\mathrm{SYT}(D)} \text{ is invertible}\right) \\
&\iff \left(\text{the matrix } M_{\mathbf{k}} \text{ is invertible}\right) \\
&\qquad\qquad \left(\text{since } \left(f\left(\mathbf{U}_P,\ \mathbf{V}_Q\right)\right)_{(P,Q)\in\mathrm{SYT}(D)\times\mathrm{SYT}(D)} = M_{\mathbf{k}}\right) \\
&\iff \left(\text{the element } \det\left(M_{\mathbf{k}}\right) \text{ of } \mathbf{k} \text{ is invertible}\right)
\end{aligned}$$

(since a well-known fact (e.g., [Grinbe15, Theorem 6.110 **(a)**] or [Loehr11, Corollary 9.51 and Remark 9.52]) says that a square matrix over a commutative ring is invertible if and only if its determinant is invertible). This proves Claim 8. □

[133] See Definition 5.19.12 for the meaning of "Gram matrix".

Now, let us take a closer look at how the Gram matrix $M_{\mathbf{k}}$ depends on the ring $\mathbf{k}$. We claim that it essentially does not: i.e., the entries of $M_{\mathbf{k}}$ are the same integers for all $\mathbf{k}$, just mapped into $\mathbf{k}$ using the canonical ring morphisms from $\mathbb{Z}$ to $\mathbf{k}$. In fact, we can describe these entries explicitly:

*Claim 9:* Let $P \in \mathrm{SYT}(D)$ and $Q \in \mathrm{SYT}(D)$. Then,

$$f\left(\mathbf{U}_P,\ \mathbf{V}_Q\right) = \sum_{\substack{(c,r) \in \mathcal{C}(T) \times \mathcal{R}(T); \\ w_{Q,T} = w_{P,T} cr}} (-1)^c (-1)^{w_{Q,T}}.$$

*Proof of Claim 9.* From (382), we obtain

$$\mathbf{U}_P = \mathbf{E}_{P,T} = w_{P,T} \mathbf{E}_T$$

(by (270), applied to $T$ instead of $Q$). Furthermore, Proposition 5.15.15 **(c)** (applied to $T$ instead of $P$) yields $w_{Q,T} = w_{Q\mathbf{r},T\mathbf{r}}$, so that $w_{Q\mathbf{r},T\mathbf{r}} = w_{Q,T}$. However, (270) (applied to $\mathbf{r}(D)$, $Q\mathbf{r}$ and $T\mathbf{r}$ instead of $D$, $P$ and $Q$) yields

$$\mathbf{E}_{Q\mathbf{r},T\mathbf{r}} = \underbrace{w_{Q\mathbf{r},T\mathbf{r}}}_{= w_{Q,T}} \mathbf{E}_{T\mathbf{r}} = w_{Q,T} \mathbf{E}_{T\mathbf{r}}.$$

Now, from (383) (applied to $Q$ instead of $P$), we obtain

$$\begin{aligned}
\mathbf{V}_Q &= T_{\mathrm{sign}} \left( \underbrace{\mathbf{E}_{Q\mathbf{r},T\mathbf{r}}}_{= w_{Q,T} \mathbf{E}_{T\mathbf{r}}} \right) = T_{\mathrm{sign}} \left( w_{Q,T} \mathbf{E}_{T\mathbf{r}} \right) \\
&= \underbrace{T_{\mathrm{sign}}\left(w_{Q,T}\right)}_{\substack{= (-1)^{w_{Q,T}} w_{Q,T} \\ \text{(by the definition} \\ \text{of } T_{\mathrm{sign}}\text{, since } w_{Q,T} \in S_n\text{)}}} \cdot \underbrace{T_{\mathrm{sign}}\left(\mathbf{E}_{T\mathbf{r}}\right)}_{= \mathbf{F}_T} \qquad \left(\text{since } T_{\mathrm{sign}} \text{ is a } \mathbf{k}\text{-algebra morphism}\right) \\
&= (-1)^{w_{Q,T}} w_{Q,T} \mathbf{F}_T.
\end{aligned}$$

Now,

$$\begin{aligned}
f\left( \underbrace{\mathbf{U}_P}_{= w_{P,T} \mathbf{E}_T},\ \underbrace{\mathbf{V}_Q}_{= (-1)^{w_{Q,T}} w_{Q,T} \mathbf{F}_T} \right) &= f\left( w_{P,T} \mathbf{E}_T,\ (-1)^{w_{Q,T}} w_{Q,T} \mathbf{F}_T \right) \\
&= [1]\left( w_{P,T} \mathbf{E}_T S\left( (-1)^{w_{Q,T}} w_{Q,T} \right) \right)
\end{aligned}$$

(by (380), applied to $u = w_{P,T}$ and $v = (-1)^{w_{Q,T}} w_{Q,T}$). In view of

$$\begin{aligned}
w_{P,T}\mathbf{E}_T & \underbrace{S\left((-1)^{w_{Q,T}} w_{Q,T}\right)}_{\substack{=(-1)^{w_{Q,T}} S\left(w_{Q,T}\right) \\ \text{(since the map } S \text{ is } \mathbf{k}\text{-linear)}}} \\
&= w_{P,T} \underbrace{\mathbf{E}_T}_{\substack{= \sum_{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T)} (-1)^c cr \\ \text{(by (234))}}} (-1)^{w_{Q,T}} \underbrace{S\left(w_{Q,T}\right)}_{\substack{=w_{Q,T}^{-1} \\ \text{(by the definition of } S\text{,} \\ \text{since } w_{Q,T}\in S_n\text{)}}} \\
&= w_{P,T} \left( \sum_{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T)} (-1)^c cr \right) (-1)^{w_{Q,T}} w_{Q,T}^{-1} \\
&= \sum_{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T)} \underbrace{w_{P,T} (-1)^c cr (-1)^{w_{Q,T}} w_{Q,T}^{-1}}_{=(-1)^c (-1)^{w_{Q,T}} w_{P,T} cr w_{Q,T}^{-1}} \\
&= \sum_{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T)} (-1)^c (-1)^{w_{Q,T}} w_{P,T} cr w_{Q,T}^{-1},
\end{aligned}$$

we can rewrite this as

$$\begin{aligned}
f\left(\mathbf{U}_P,\ \mathbf{V}_Q\right) &= [1]\left(\sum_{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T)} (-1)^c\,(-1)^{w_{Q,T}}\, w_{P,T}crw_{Q,T}^{-1}\right) \\
&= \sum_{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T)} (-1)^c\,(-1)^{w_{Q,T}}\cdot \underbrace{[1]\left(w_{P,T}crw_{Q,T}^{-1}\right)}_{=\left[w_{Q,T}\right](w_{P,T}cr)} \\
&\qquad \text{(by (377), applied to } g=w_{Q,T} \text{ and } \mathbf{a}=w_{P,T}cr\text{)} \\
&= \sum_{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T)} (-1)^c\,(-1)^{w_{Q,T}}\cdot \underbrace{\left[w_{Q,T}\right]\left(w_{P,T}cr\right)}_{=\delta_{w_{Q,T},w_{P,T}cr}} \\
&\qquad \text{(since } [u]v=\delta_{u,v} \text{ for any } u,v\in S_n\text{)} \\
&\qquad \left(\begin{array}{c}\text{since addition and scaling in } \mathbf{k}\left[S_n\right] \\ \text{are coordinatewise}\end{array}\right) \\
&= \sum_{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T)} (-1)^c\,(-1)^{w_{Q,T}}\,\delta_{w_{Q,T},w_{P,T}cr} \\
&= \sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ w_{Q,T}=w_{P,T}cr}} (-1)^c\,(-1)^{w_{Q,T}} \underbrace{\delta_{w_{Q,T},w_{P,T}cr}}_{=1} \\
&\qquad \text{(since } 1=w_{P,T}crw_{Q,T}^{-1}\text{)} \\
&\quad + \sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ w_{Q,T}\neq w_{P,T}cr}} (-1)^c\,(-1)^{w_{Q,T}} \underbrace{\delta_{w_{Q,T},w_{P,T}cr}}_{=0} \\
&\qquad \text{(since } 1\neq w_{P,T}crw_{Q,T}^{-1}\text{)} \\
&\qquad \left(\begin{array}{c}\text{since each } (c,r)\in\mathcal{C}\left(T\right)\times\mathcal{R}\left(T\right) \text{ satisfies} \\ \text{either } w_{Q,T}=w_{P,T}cr \text{ or } w_{Q,T}\neq w_{P,T}cr \\ \text{(but not both)}\end{array}\right) \\
&= \sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ w_{Q,T}=w_{P,T}cr}} (-1)^c\,(-1)^{w_{Q,T}} + \underbrace{\sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ w_{Q,T}\neq w_{P,T}cr}} (-1)^c\,(-1)^{w_{Q,T}}\,0}_{=0} \\
&= \sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ w_{Q,T}=w_{P,T}cr}} (-1)^c\,(-1)^{w_{Q,T}}\,.
\end{aligned}$$

This proves Claim 9. □

We shall now use a "base change" argument somewhat similar to the one in the proof of Proposition 5.12.10 but even stranger, in that we will use finite fields $\mathbb{Z}/p$ for prime numbers $p$.

We let $\varphi_{\mathbf{k}}$ denote the canonical ring morphism from $\mathbb{Z}$ to $\mathbf{k}$ (sending each integer $m$ to $m\cdot 1_{\mathbf{k}}$). (This was denoted $f$ in some prior proofs, but the letter $f$ now means something different.) We can now relate the determinant of the Gram matrix $M_{\mathbf{k}}$ defined over the ring $\mathbf{k}$ with the determinant of the Gram matrix $M_{\mathbb{Z}}$ defined over the ring $\mathbb{Z}$:

*Claim 10:* We have

$$\det\left(M_{\mathbf{k}}\right) = \varphi_{\mathbf{k}}\left(\det\left(M_{\mathbb{Z}}\right)\right).$$

*Proof of Claim 10.* We shall compare respective entries of the matrices $M_{\mathbb{Z}}$ and $M_{\mathbf{k}}$. Indeed, fix $P \in \mathrm{SYT}\left(D\right)$ and $Q \in \mathrm{SYT}\left(D\right)$. Then, by the definition of $M_{\mathbf{k}}$, we know that the $(P,Q)$-th entry of the matrix $M_{\mathbf{k}}$ is

$$f\left(\mathbf{U}_P,\ \mathbf{V}_Q\right) = \sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ w_{Q,T}=w_{P,T}cr}} \left(-1\right)^{c} \left(-1\right)^{w_{Q,T}} \tag{385}$$

(by Claim 9). But the same argument (applied to the ring $\mathbb{Z}$ instead of $\mathbf{k}$) shows that the $(P,Q)$-th entry of the matrix $M_{\mathbb{Z}}$ is

$$\sum_{\substack{(c,r)\in\mathcal{C}(T)\times\mathcal{R}(T);\\ w_{Q,T}=w_{P,T}cr}} \left(-1\right)^{c} \left(-1\right)^{w_{Q,T}} \tag{386}$$

as well, but now interpreted as an actual integer rather than as an element of $\mathbf{k}$. Clearly, the ring morphism $\varphi_{\mathbf{k}} : \mathbb{Z} \to \mathbf{k}$ sends the sum in (386) to the sum in (385) (since this morphism $\varphi_{\mathbf{k}}$ sends each integer $m$ to the corresponding element $m \cdot 1_{\mathbf{k}}$ of $\mathbf{k}$). In other words, the morphism $\varphi_{\mathbf{k}}$ sends the $(P,Q)$-th entry of the matrix $M_{\mathbb{Z}}$ to the $(P,Q)$-th entry of the matrix $M_{\mathbf{k}}$ (since the former entry is the sum in (386), whereas the latter entry is the sum in (385)).

Forget that we fixed $P$ and $Q$. We thus have shown that for each $P \in \mathrm{SYT}\left(D\right)$ and $Q \in \mathrm{SYT}\left(D\right)$, the morphism $\varphi_{\mathbf{k}}$ sends the $(P,Q)$-th entry of the matrix $M_{\mathbb{Z}}$ to the $(P,Q)$-th entry of the matrix $M_{\mathbf{k}}$. In other words, the morphism $\varphi_{\mathbf{k}}$ sends each entry of the matrix $M_{\mathbb{Z}}$ to the corresponding entry of the matrix $M_{\mathbf{k}}$. Since $\varphi_{\mathbf{k}}$ is a ring morphism, this entails that $\varphi_{\mathbf{k}}$ sends the determinant of $M_{\mathbb{Z}}$ to the determinant of $M_{\mathbf{k}}$ (since ring morphisms respect determinants). In other words, $\varphi_{\mathbf{k}}\left(\det\left(M_{\mathbb{Z}}\right)\right) = \det\left(M_{\mathbf{k}}\right)$. This proves Claim 10. □

Now comes the actually surprising part of the proof:

*Claim 11:* The integer $\det\left(M_{\mathbb{Z}}\right)$ is invertible in $\mathbb{Z}$.

*Proof of Claim 11.* Assume the contrary. Thus, $\det\left(M_{\mathbb{Z}}\right)$ is not invertible in $\mathbb{Z}$. Hence, $\det\left(M_{\mathbb{Z}}\right)$ is neither 1 nor $-1$ (since the integers 1 and $-1$ are invertible in $\mathbb{Z}$). Therefore, $\det\left(M_{\mathbb{Z}}\right)$ is divisible by some prime number $p$ (since any integer other than 1 and $-1$ is divisible by some prime). Consider this $p$. The canonical ring morphism $\varphi_{\mathbb{Z}/p} : \mathbb{Z} \to \mathbb{Z}/p$ sends each integer that is divisible by $p$ to 0. Hence, $\varphi_{\mathbb{Z}/p}$ sends $\det\left(M_{\mathbb{Z}}\right)$ to 0 (since $\det\left(M_{\mathbb{Z}}\right)$ is divisible by $p$). However, applying Claim 10 to the field $\mathbb{Z}/p$ instead of $\mathbf{k}$, we obtain

$$\det\left(M_{\mathbb{Z}/p}\right) = \varphi_{\mathbb{Z}/p}\left(\det\left(M_{\mathbb{Z}}\right)\right) = 0$$

(since $\varphi_{\mathbb{Z}/p}$ sends $\det\left(M_{\mathbb{Z}}\right)$ to 0).

However, in the above proof of Theorem 5.19.33, we have shown that the map $f_R$ is invertible **if $\mathbf{k}$ is a field**. We cannot apply this fact to our ring $\mathbf{k}$, since $\mathbf{k}$ might not be a field. However, we can apply it to $\mathbb{Z}/p$ instead of $\mathbf{k}$ (since $\mathbb{Z}/p$ is a field). Thus, we find that the map $f_R$ defined over the field $\mathbb{Z}/p$ instead of $\mathbf{k}$ is invertible.

But we can also apply Claim 8 to $\mathbb{Z}/p$ instead of $\mathbf{k}$. As a result, we see that the map $f_R$ defined over the field $\mathbb{Z}/p$ instead of $\mathbf{k}$ is invertible if and only if the element $\det\left(M_{\mathbb{Z}/p}\right)$ of $\mathbb{Z}/p$ is invertible. Hence, the element $\det\left(M_{\mathbb{Z}/p}\right)$ of $\mathbb{Z}/p$ is invertible (since the map $f_R$ defined over the field $\mathbb{Z}/p$ instead of $\mathbf{k}$ is invertible). In other words, the element 0 of $\mathbb{Z}/p$ is invertible (since $\det\left(M_{\mathbb{Z}/p}\right) = 0$). This contradicts the fact that the element 0 of $\mathbb{Z}/p$ is clearly not invertible (since 0 is not invertible in any field).

This contradiction shows that our assumption was false. Hence, Claim 11 is proved. □

Now, we are almost done. Claim 11 shows that the integer $\det\left(M_{\mathbb{Z}}\right)$ is invertible in $\mathbb{Z}$. Hence, its image $\varphi_{\mathbf{k}}\left(\det\left(M_{\mathbb{Z}}\right)\right)$ is invertible in $\mathbf{k}$ (since $\varphi_{\mathbf{k}}$ is a ring morphism and thus sends invertible elements to invertible elements). In other words, $\det\left(M_{\mathbf{k}}\right)$ is invertible in $\mathbf{k}$ (since Claim 10 says that $\det\left(M_{\mathbf{k}}\right) = \varphi_{\mathbf{k}}\left(\det\left(M_{\mathbb{Z}}\right)\right)$).

In other words, the element $\det\left(M_{\mathbf{k}}\right)$ of $\mathbf{k}$ is invertible. By Claim 8, this entails that the map $f_R : \mathcal{A}\mathbf{F}_T \to \left(\mathcal{A}\mathbf{E}_T\right)^*$ is invertible.

From here, we can finish the proof of Theorem 5.19.34 in the exact same way as we finished the proof of Theorem 5.19.33 after we showed that the map $f_R$ is invertible. □

The above proof of Theorem 5.19.34 works not only when $D$ is a skew Young diagram, but also more generally if we merely assume that the $\mathbf{k}$-modules $\mathcal{A}\mathbf{E}_T$ and $\mathcal{A}\mathbf{F}_T$ (or, equivalently, $\mathcal{S}^D$ and $\mathcal{S}^{\mathbf{r}(D)}$) have bases independent on $\mathbf{k}$ (that is, bases indexed by combinatorial objects that are "respected" by the canonical ring morphisms $\varphi_{\mathbf{k}} : \mathbb{Z} \to \mathbf{k}$). This holds for various diagrams that are not skew Young diagrams, such as the diagram from Exercise 5.9.2. Whether this holds for arbitrary $D$ is an open question. Thus, we do not know how far Theorem 5.19.34 can actually be generalized.

**Exercise 5.19.8.** There is an alternative proof of Theorem 5.19.34, which avoids the use of base change but instead argues by computing $\det\left(M_{\mathbf{k}}\right)$. Namely, using the notations from our above proof of Theorem 5.19.34:

**(a)** [3] Show that every $P, Q \in \mathrm{SYT}\left(D\right)$ satisfy

$$f\left(\mathbf{U}_P,\ \mathbf{V}_Q\right) = \sum_{\substack{(c,r)\in\mathcal{C}(P)\times\mathcal{R}(P);\\ Q=crP}} \left(-1\right)^c \left(-1\right)^{w_{Q,T}}.$$

**(b)** [2] Show that $f\left(\mathbf{U}_P,\ \mathbf{V}_Q\right) = 0$ for every $P, Q \in \mathrm{SYT}\left(D\right)$ that satisfy $\overline{P} < \overline{Q}$ in the Young last letter order (recall Definition 5.7.5).

**(c)** $\boxed{1}$ Show that $f(\mathbf{U}_P, \mathbf{V}_P) = (-1)^{w_{P,T}}$ for all $P \in \mathrm{SYT}(D)$.

**(d)** $\boxed{1}$ Conclude that the matrix $M_{\mathbf{k}}$ is triangular and has determinant 1, and thus is invertible.

### 5.19.7. Consequences in characteristic 0

Under certain special conditions, the dual of a Specht module $\mathcal{S}^D$ has an even simpler description than the one given by Theorem 5.19.34. Namely, if $D$ is the Young diagram $Y(\lambda)$ of a partition $\lambda$, and if the number $h^\lambda$ from Definition 5.12.11 is invertible in $\mathbf{k}$, then the dual of $\mathcal{S}^D = \mathcal{S}^\lambda$ is just $\mathcal{S}^\lambda$ itself:

**Theorem 5.19.35.** Let $\lambda$ be a partition of $n$. Assume that the integer $h^\lambda$ (defined in Definition 5.12.11) is invertible in $\mathbf{k}$. Then, $\left(\mathcal{S}^\lambda\right)^* \cong \mathcal{S}^\lambda$ as $S_n$-representations.

*Proof.* Recall that $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)}$. Likewise, $\mathcal{S}^{\lambda^t} = \mathcal{S}^{Y\left(\lambda^t\right)} = \mathcal{S}^{\mathbf{r}(Y(\lambda))}$ (since the definition of the conjugate partition $\lambda^t$ yields $Y\left(\lambda^t\right) = \mathbf{r}(Y(\lambda))$).

But $Y(\lambda) = Y(\lambda/\varnothing)$ is a skew Young diagram of size $|Y(\lambda)| = |\lambda| = n$ (since $\lambda$ is a partition of $n$). Hence, Theorem 5.19.34 (applied to $D = Y(\lambda)$) yields

$$\left(\mathcal{S}^{Y(\lambda)}\right)^* \cong \left(\mathcal{S}^{\mathbf{r}(Y(\lambda))}\right)^{\mathrm{sign}} \text{ as } S_n\text{-representations.}$$

In other words,

$$\left(\mathcal{S}^\lambda\right)^* \cong \left(\mathcal{S}^{\lambda^t}\right)^{\mathrm{sign}} \text{ as } S_n\text{-representations}$$

(since $\mathcal{S}^\lambda = \mathcal{S}^{Y(\lambda)}$ and $\mathcal{S}^{\lambda^t} = \mathcal{S}^{\mathbf{r}(Y(\lambda))}$). But Theorem 5.18.13 yields $\left(\mathcal{S}^\lambda\right)^{\mathrm{sign}} \cong \mathcal{S}^{\lambda^t}$. Thus, $\mathcal{S}^{\lambda^t} \cong \left(\mathcal{S}^\lambda\right)^{\mathrm{sign}}$.

By Corollary 5.18.9, this entails $\left(\mathcal{S}^{\lambda^t}\right)^{\mathrm{sign}} \cong \left(\left(\mathcal{S}^\lambda\right)^{\mathrm{sign}}\right)^{\mathrm{sign}} = \mathcal{S}^\lambda$ (by Proposition 5.18.6). Altogether, we thus find $\left(\mathcal{S}^\lambda\right)^* \cong \left(\mathcal{S}^{\lambda^t}\right)^{\mathrm{sign}} \cong \mathcal{S}^\lambda$. This proves Theorem 5.19.35. □

**Corollary 5.19.36** (self-duality theorem for $S_n$-representations)**.** Let $\mathbf{k}$ be a field of characteristic 0. Let $V$ be any $S_n$-representation that is finite-dimensional as a $\mathbf{k}$-vector space. Then, $V^* \cong V$ as $S_n$-representations.

*Proof sketch.* We shall say that an $S_n$-representation $M$ is *self-dual* if it satisfies $M^* \cong M$. Thus, our goal is to prove that $V$ is self-dual.

Clearly, if $M$ is a self-dual $S_n$-representation, then any $S_n$-representation $N$ that is isomorphic to $M$ must also be self-dual (since $N \cong M$ entails $N^* \cong M^* \cong M \cong N$). In other words, self-duality is preserved under isomorphism.

If $\lambda$ is any partition of $n$, then the positive integer $h^\lambda$ (defined in Definition 5.12.11) is invertible in $\mathbf{k}$ (since $\mathbf{k}$ has characteristic 0), and thus Theorem 5.19.35 shows that $\left(\mathcal{S}^\lambda\right)^* \cong \mathcal{S}^\lambda$; but this is just saying that the Specht module $\mathcal{S}^\lambda$ is self-dual (as an $S_n$-representation). In other words, we have shown the following claim:

> *Claim 1:* If $\lambda$ is any partition of $n$, then the Specht module $\mathcal{S}^\lambda$ is self-dual.

On the other hand, if $P$ and $Q$ are two self-dual $S_n$-representations, then Proposition 5.19.23 yields

$$(P \oplus Q)^* \cong \underbrace{P^*}_{\substack{\cong P \\ \text{(since } P \text{ is self-dual)}}} \oplus \underbrace{Q^*}_{\substack{\cong Q \\ \text{(since } Q \text{ is self-dual)}}} \cong P \oplus Q,$$

which shows that $P \oplus Q$ is again self-dual. In other words, any direct sum of two self-dual $S_n$-representations is again self-dual. Hence, by induction, we conclude that any direct sum of finitely many self-dual $S_n$-representations is again self-dual.

But Corollary 5.13.9 shows that $V$ is isomorphic to a direct sum of finitely many copies of Specht modules of the form $\mathcal{S}^\lambda$ with $\lambda$ being partitions of $n$. Each of these Specht modules $\mathcal{S}^\lambda$ is self-dual (by Claim 1). Hence, their direct sum is self-dual again (since any direct sum of finitely many self-dual $S_n$-representations is again self-dual). In other words, $V$ is self-dual (since $V$ is isomorphic to this direct sum, but self-duality is preserved under isomorphism). This proves Corollary 5.19.36. □

**Remark 5.19.37.** Corollary 5.19.36 would fail if $V$ was allowed to be infinite-dimensional (here, $V^* \cong V$ will often fail even on the level of $\mathbf{k}$-vector spaces). It can also fail if $\mathbf{k}$ is not a field but merely a commutative $\mathbb{Q}$-algebra.

**Remark 5.19.38.** Corollary 5.19.36 can be generalized: Let $G$ be a finite group such that each $g \in G$ is conjugate to its inverse $g^{-1}$. (Such groups $G$ are said to be *ambivalent*.) Let $\mathbf{k}$ be a field of characteristic 0. Let $V$ be any $G$-representation that is finite-dimensional as a $\mathbf{k}$-vector space. Then, $V^* \cong V$.

This result can be applied to $G = S_n$ because Theorem 3.13.1 shows that each $g \in S_n$ is conjugate to $g^{-1}$. A proof of this generalization is implicit in [Lorenz18, proof of Lemma 3.24].

**Exercise 5.19.9.** [3] A bilinear form $f : V \times V \to \mathbf{k}$ for a $\mathbf{k}$-module $V$ is said to be *symmetric* if all $u, v \in V$ satisfy $f(u, v) = f(v, u)$.

Let $\lambda$ be a partition of $n$. Assume that $h^\lambda$ is invertible in $\mathbf{k}$. Prove that there exists a symmetric $S_n$-invariant bilinear form $f : \mathcal{S}^\lambda \times \mathcal{S}^\lambda \to \mathbf{k}$ such that the map $f_R : \mathcal{S}^\lambda \to \left(\mathcal{S}^\lambda\right)^*$ is bijective.

### 5.19.8. Characters of duals

Recall the notion of a character (Definition 5.17.1). We can easily describe the character of the dual of a $G$-representation in general:

**Proposition 5.19.39.** Let $G$ be a group. Let $V$ be a representation of $G$ over $\mathbf{k}$. Assume that $V$ is free of finite rank (i.e., has a finite basis) as a $\mathbf{k}$-module. Then, the characters $\chi_V$ and $\chi_{V^*}$ of $V$ and $V^*$ are related by the following formula:

$$\chi_{V^*}(g) = \chi_V\left(g^{-1}\right) \qquad \text{for each } g \in G.$$

For $G = S_n$, we can simplify this by replacing $g^{-1}$ by $g$:

**Corollary 5.19.40.** Let $V$ be a representation of $S_n$ over $\mathbf{k}$. Assume that $V$ is free of finite rank (i.e., has a finite basis) as a $\mathbf{k}$-module. Then, the characters $\chi_V$ and $\chi_{V^*}$ of $V$ and $V^*$ are equal: i.e., we have $\chi_V = \chi_{V^*}$.

**Exercise 5.19.10.** [2] Prove Proposition 5.19.39 and Corollary 5.19.40.

**Exercise 5.19.11.** [3] Many texts on representation theory restrict themselves to the case when $\mathbf{k} = \mathbb{C}$ and the group $G$ is finite. In such texts, a version of Proposition 5.19.39 (for $\mathbf{k} = \mathbb{C}$ and for $G$ finite) is often found with the equality $\chi_{V^*}(g) = \chi_V\left(g^{-1}\right)$ replaced by $\chi_{V^*}(g) = \overline{\chi_V(g)}$, where $\overline{z}$ denotes the complex conjugate of a complex number $z$.

Derive this version from Proposition 5.19.39.

[**Hint:** Argue that the endomorphism $\tau_g$ of $V$ has finite order in the group $\mathrm{GL}(V)$, and thus all its eigenvalues lie on the unit circle.]

### 5.19.9. Proving Theorem 5.18.16

We can now prove Theorem 5.18.16, which we left unproved in the previous section:

*Proof of Theorem 5.18.16.* Theorem 5.19.33 shows that $\left(\mathcal{S}^D\right)^* \cong \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}}$ as $S_n$-representations. But Corollary 5.19.36 (applied to $V = \mathcal{S}^D$) yields $\left(\mathcal{S}^D\right)^* \cong \mathcal{S}^D$. Hence, $\mathcal{S}^D \cong \left(\mathcal{S}^D\right)^* \cong \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}}$. By Corollary 5.18.9, this entails

$$\left(\mathcal{S}^D\right)^{\mathrm{sign}} \cong \left(\left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}}\right)^{\mathrm{sign}} = \mathcal{S}^{\mathbf{r}(D)} \qquad \text{(by Proposition 5.18.6)}.$$

This proves Theorem 5.18.16. □

## 5.20. Application: Proof of antipodal conjugacy

With the Artin–Wedderburn theorem (Theorem 5.14.1) and the self-duality theorem (Corollary 5.19.36), we now have most of the tools needed to prove Theorem 3.13.2. Thus we will give the proof in the present section, albeit somewhat tersely.

### 5.20.1. Conjugacy and the Taussky–Zassenhaus theorem

The proof requires a few auxiliary facts. One is a linear-algebraic result by Taussky and Zassenhaus from 1959 [TauZas59, Theorem 1]:[134]

**Theorem 5.20.1** (Taussky–Zassenhaus theorem)**.** Assume that $\mathbf{k}$ is a field. Let $A \in \mathbf{k}^{m \times m}$ be any square matrix over $\mathbf{k}$. Then, the matrix $A$ is similar (i.e., conjugate in the matrix ring $\mathbf{k}^{m \times m}$) to its transpose $A^T$.

Note that this theorem would fail for $\mathbf{k} = \mathbb{Z}$ (see `https://math.stackexchange.com/questions/1732276/is-a-matrix-over-a-pid-similar-to-its-transpose` for a counterexample), so the "$\mathbf{k}$ is a field" assumption is important.

We can give the Taussky–Zassenhaus theorem a more conceptual formulation by replacing matrices by linear maps. Similarity is then replaced by a slightly generalized notion of conjugacy, defined as follows:

**Definition 5.20.2.** Let $U$ and $V$ be two $\mathbf{k}$-modules. Then, two endomorphisms $f \in \operatorname{End} U$ and $g \in \operatorname{End} V$ will be called *conjugate* (to each other) if there exists some $\mathbf{k}$-module isomorphism $\phi : U \to V$ satisfying $g = \phi \circ f \circ \phi^{-1}$ (that is, $g \circ \phi = \phi \circ f$). We will write "$f \sim g$" for this.

Note that this notion of conjugacy generalizes the usual concept of conjugacy in an endomorphism ring. Indeed, if $V$ is any $\mathbf{k}$-module, then two endomorphisms $f, g \in \operatorname{End} V$ satisfy $f \sim g$ (in the sense of Definition 5.20.2) if and only if $f$ and $g$ are conjugate in the ring $\operatorname{End} V$ (because a $\mathbf{k}$-module isomorphism $\phi : V \to V$ is the same thing as an invertible element $\phi$ of $\operatorname{End} V$).

It is easy to see that conjugacy of endomorphisms is an equivalence relation. In particular, it is symmetric (i.e., if $\alpha \sim \beta$, then $\beta \sim \alpha$) and transitive (i.e., if $\alpha \sim \beta$ and $\beta \sim \gamma$, then $\alpha \sim \gamma$).

Conjugacy of endomorphisms is the coordinate-free version of similarity of matrices:

**Proposition 5.20.3.** Let $U$ and $V$ be two free $\mathbf{k}$-modules with finite bases $(u_1, u_2, \ldots, u_m)$ and $(v_1, v_2, \ldots, v_m)$, respectively. Let $f \in \operatorname{End} U$ and $g \in \operatorname{End} V$ be two endomorphisms. Let $A \in \mathbf{k}^{m \times m}$ be the matrix representing

[134] See `https://math.stackexchange.com/questions/62497/matrix-is-conjugate-to-its-own-transpose` or [Elman22, Corollary 45.14] for a short proof using the Smith normal form over the polynomial ring $\mathbf{k}[x]$.

the endomorphism $f$ with respect to the basis $(u_1, u_2, \ldots, u_m)$ of $U$. Let $B \in \mathbf{k}^{m \times m}$ be the matrix representing the endomorphism $g$ with respect to the basis $(v_1, v_2, \ldots, v_m)$ of $V$. Then, the conjugacy $f \sim g$ holds if and only if the matrices $A$ and $B$ are similar (i.e., conjugate in the matrix ring $\mathbf{k}^{m \times m}$).

*Proof sketch.* This is basic linear algebra. For the "$\Longrightarrow$" direction, assume that $f \sim g$ holds. Thus, there exists some $\mathbf{k}$-module isomorphism $\phi : U \to V$ satisfying $g = \phi \circ f \circ \phi^{-1}$. Consider this $\phi$, and let $P \in \mathbf{k}^{m \times m}$ be the matrix representing it with respect to the bases $(u_1, u_2, \ldots, u_m)$ and $(v_1, v_2, \ldots, v_m)$ of $U$ and $V$. Then, this matrix $P$ is invertible (since it represents the invertible map $\phi$) and satisfies $B = PAP^{-1}$ (since $g = \phi \circ f \circ \phi^{-1}$). Hence, the matrices $A$ and $B$ are similar. This proves the "$\Longrightarrow$" direction of Proposition 5.20.3. To prove the "$\Longleftarrow$" direction, reverse the argument (recalling that any $m \times m$-matrix represents some linear map $\phi : U \to V$ with respect to the bases $(u_1, u_2, \ldots, u_m)$ and $(v_1, v_2, \ldots, v_m)$). □

We can now restate Theorem 5.20.1 as follows:

**Corollary 5.20.4** (Taussky–Zassenhaus theorem in coordinate-free form)**.** Assume that $\mathbf{k}$ is a field. Let $V$ be a finite-dimensional $\mathbf{k}$-vector space. Let $f \in \operatorname{End} V$ be any endomorphism of $V$. Then, the endomorphism $f \in \operatorname{End} V$ is conjugate to its dual $f^* \in \operatorname{End}(V^*)$.

*Proof sketch.* Fix a basis $(v_1, v_2, \ldots, v_m)$ of $V$. Let $(v_1^*, v_2^*, \ldots, v_m^*)$ be its dual basis, which of course is a basis of $V^*$. Let $A$ be the matrix representing the endomorphism $f$ with respect to the basis $(v_1, v_2, \ldots, v_m)$ of $V$. Then, the matrix representing its dual $f^*$ with respect to the dual basis $(v_1^*, v_2^*, \ldots, v_m^*)$ is the transpose $A^T$ of $A$ (by Proposition 5.19.4 **(d)**). Hence, the conjugacy $f \sim f^*$ holds if and only if the matrices $A$ and $A^T$ are similar (by Proposition 5.20.3). But Theorem 5.20.1 shows that the matrices $A$ and $A^T$ are indeed similar. Thus, $f \sim f^*$ holds. This proves Corollary 5.20.4. □

Another source of conjugate endomorphisms are isomorphic representations of a group or, more generally, isomorphic modules over a $\mathbf{k}$-algebra:

**Proposition 5.20.5.** Let $R$ be a $\mathbf{k}$-algebra. Let $U$ and $V$ be two isomorphic left $R$-modules. Let $\mathbf{a} \in R$. Then,

$$(\tau_{\mathbf{a}} \text{ on } U) \sim (\tau_{\mathbf{a}} \text{ on } V).$$

Here, "$\tau_{\mathbf{a}}$ on $U$" means the $\mathbf{k}$-linear map $U \to U,\ u \mapsto \mathbf{a} \cdot u$ (viewed as an endomorphism of the $\mathbf{k}$-module $U$), whereas "$\tau_{\mathbf{a}}$ on $V$" means the $\mathbf{k}$-linear map $V \to V,\ v \mapsto \mathbf{a} \cdot v$ (viewed as an endomorphism of the $\mathbf{k}$-module $V$).

*Proof.* There exists a left $R$-module isomorphism $\phi : U \to V$ (since $U$ and $V$ are isomorphic). Consider this $\phi$. Thus,

$$\phi(\mathbf{a} \cdot u) = \mathbf{a} \cdot \phi(u) \qquad \text{for each } u \in U \tag{387}$$

(since $\phi$ is a left $R$-module morphism). In other words,

$$\phi \circ (\tau_{\mathbf{a}} \text{ on } U) = (\tau_{\mathbf{a}} \text{ on } V) \circ \phi$$

[135]. Since $\phi$ is invertible, we can solve this equality for $(\tau_{\mathbf{a}} \text{ on } V)$, obtaining

$$(\tau_{\mathbf{a}} \text{ on } V) = \phi \circ (\tau_{\mathbf{a}} \text{ on } U) \circ \phi^{-1}.$$

Since $\phi$ is a $\mathbf{k}$-module isomorphism (being an $R$-module isomorphism), we thus conclude that $(\tau_{\mathbf{a}} \text{ on } U) \sim (\tau_{\mathbf{a}} \text{ on } V)$ (by Definition 5.20.2). This proves Proposition 5.20.5. □

### 5.20.2. Using Artin–Wedderburn

Our next lemma is a simple consequence of the Artin–Wedderburn theorem:

**Lemma 5.20.6.** Assume that $n!$ is invertible in $\mathbf{k}$. Let $\mathbf{a}, \mathbf{b} \in \mathbf{k}[S_n]$. Assume that

$$\rho_\lambda(\mathbf{a}) \text{ is conjugate to } \rho_\lambda(\mathbf{b}) \text{ in } \operatorname{End}\left(\mathcal{S}^\lambda\right) \tag{388}$$

for each partition $\lambda$ of $n$, where the morphism $\rho_\lambda$ is defined in Theorem 5.14.1. Then, $\mathbf{a}$ is conjugate to $\mathbf{b}$ in $\mathbf{k}[S_n]$.

*Proof.* Let $P$ be the set of all partitions of $n$. Consider the map

$$\begin{aligned} \mathrm{AH} : \mathbf{k}[S_n] &\to \underbrace{\prod_{\lambda \in P} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)}_{\text{a direct product of } \mathbf{k}\text{-algebras}}, \\ \mathbf{c} &\mapsto (\rho_\lambda(\mathbf{c}))_{\lambda \in P}. \end{aligned}$$

By Theorem 5.14.1, this map AH is a $\mathbf{k}$-algebra isomorphism[136]. Hence, in particular, this map AH is injective, and its inverse $\mathrm{AH}^{-1}$ is again a $\mathbf{k}$-algebra isomorphism.

[135] *Proof.* Let $u \in U$. Then,

$$(\phi \circ (\tau_{\mathbf{a}} \text{ on } U))(u) = \phi((\tau_{\mathbf{a}} \text{ on } U)(u)) = \phi(\mathbf{a} \cdot u)$$

(since the definition of "$\tau_{\mathbf{a}}$ on $U$" yields $(\tau_{\mathbf{a}} \text{ on } U)(u) = \mathbf{a} \cdot u$) and

$$((\tau_{\mathbf{a}} \text{ on } V) \circ \phi)(u) = (\tau_{\mathbf{a}} \text{ on } V)(\phi(u)) = \mathbf{a} \cdot \phi(u)$$

(by the definition of "$\tau_{\mathbf{a}}$ on $V$"). The right hand sides of these two equalities are equal (by (387)). Hence, so are their left hand sides. In other words, $(\phi \circ (\tau_{\mathbf{a}} \text{ on } U))(u) = ((\tau_{\mathbf{a}} \text{ on } V) \circ \phi)(u)$.

Having proved this equality for all $u \in U$, we thus conclude that $\phi \circ (\tau_{\mathbf{a}} \text{ on } U) = (\tau_{\mathbf{a}} \text{ on } V) \circ \phi$.

[136] The notation in Theorem 5.14.1 differs slightly from the one we are using here, in that we wrote "$\lambda$ is a partition of $n$" instead of the short notation "$\lambda \in P$", and in that we used the letter $\mathbf{a}$ (rather than $\mathbf{c}$) for the variable.

Now, let $\lambda \in P$. Thus, $\lambda$ is a partition of $n$ (by the definition of $P$). Hence, $\rho_\lambda(\mathbf{a})$ is conjugate to $\rho_\lambda(\mathbf{b})$ in $\operatorname{End}\left(\mathcal{S}^\lambda\right)$ (by our assumption (388)). In other words, there exists some invertible element $\psi_\lambda \in \operatorname{End}\left(\mathcal{S}^\lambda\right)$ such that

$$\rho_\lambda(\mathbf{a}) = \psi_\lambda \cdot \rho_\lambda(\mathbf{b}) \cdot \psi_\lambda^{-1} \tag{389}$$

(by the definition of "conjugate" in a ring[137]). Consider this $\psi_\lambda$.

Forget that we fixed $\lambda$. We thus have constructed an invertible element $\psi_\lambda \in \operatorname{End}\left(\mathcal{S}^\lambda\right)$ satisfying (389) for each $\lambda \in P$.

Next, define the family

$$\psi := (\psi_\lambda)_{\lambda \in P} \in \prod_{\lambda \in P} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right).$$

Then, $\psi$ is a family of invertible elements (since $\psi_\lambda$ is invertible for each $\lambda \in P$), and thus is itself invertible (since the multiplication in the direct product $\prod_{\lambda \in P} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right)$ is entrywise), with inverse

$$\psi^{-1} = \left(\psi_\lambda^{-1}\right)_{\lambda \in P}$$

(since $\psi = (\psi_\lambda)_{\lambda \in P}$ and again since the multiplication is entrywise).

Define an element $\mathbf{c} \in \mathbf{k}[S_n]$ by $\mathbf{c} := \mathrm{AH}^{-1}(\psi)$. Then, $\mathbf{c}$ is the image of the invertible element $\psi$ under the $\mathbf{k}$-algebra isomorphism $\mathrm{AH}^{-1}$; therefore, $\mathbf{c}$ itself is invertible (since $\mathbf{k}$-algebra isomorphisms send invertible elements to invertible elements). From $\mathbf{c} = \mathrm{AH}^{-1}(\psi)$, we obtain

$$\mathrm{AH}(\mathbf{c}) = \psi = (\psi_\lambda)_{\lambda \in P}.$$

Moreover, since AH is a $\mathbf{k}$-algebra morphism, we have

$$\mathrm{AH}\left(\mathbf{c}^{-1}\right) = \left(\underbrace{\mathrm{AH}(\mathbf{c})}_{=\psi}\right)^{-1} = \psi^{-1} = \left(\psi_\lambda^{-1}\right)_{\lambda \in P}.$$

[137] Note that the multiplication $\cdot$ on the right hand side of (389) is composition of endomorphisms, so we could just as well write $\circ$ for it.

Since AH is a $\mathbf{k}$-algebra morphism, we furthermore have

$$\begin{aligned}
&\operatorname{AH}\left(\mathbf{c}\mathbf{b}\mathbf{c}^{-1}\right) \\
&= \operatorname{AH}(\mathbf{c}) \cdot \operatorname{AH}(\mathbf{b}) \cdot \operatorname{AH}\left(\mathbf{c}^{-1}\right) \\
&= (\psi_\lambda)_{\lambda\in P} \cdot (\rho_\lambda(\mathbf{b}))_{\lambda\in P} \cdot \left(\psi_\lambda^{-1}\right)_{\lambda\in P} \\
&\qquad\qquad \left(\begin{array}{c} \text{since } \operatorname{AH}(\mathbf{c}) = (\psi_\lambda)_{\lambda\in P} \text{ and } \operatorname{AH}(\mathbf{b}) = (\rho_\lambda(\mathbf{b}))_{\lambda\in P} \\ \text{(by the definition of AH) and } \operatorname{AH}\left(\mathbf{c}^{-1}\right) = \left(\psi_\lambda^{-1}\right)_{\lambda\in P} \end{array}\right) \\
&= \left(\psi_\lambda \cdot \rho_\lambda(\mathbf{b}) \cdot \psi_\lambda^{-1}\right)_{\lambda\in P} \\
&\qquad\qquad \left(\text{since the multiplication in } \prod_{\lambda\in P} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^\lambda\right) \text{ is entrywise}\right) \\
&= (\rho_\lambda(\mathbf{a}))_{\lambda\in P} \qquad \left(\begin{array}{c} \text{since (389) yields } \psi_\lambda \cdot \rho_\lambda(\mathbf{b}) \cdot \psi_\lambda^{-1} = \rho_\lambda(\mathbf{a}) \\ \text{for each } \lambda \in P \end{array}\right) \\
&= \operatorname{AH}(\mathbf{a}) \qquad \text{(by the definition of AH)}.
\end{aligned}$$

Since AH is injective, we thus conclude that $\mathbf{c}\mathbf{b}\mathbf{c}^{-1} = \mathbf{a}$. In other words, $\mathbf{a} = \mathbf{c}\mathbf{b}\mathbf{c}^{-1}$. Hence, the elements $\mathbf{a}$ and $\mathbf{b}$ of $\mathbf{k}[S_n]$ are conjugate. Lemma 5.20.6 is thus proved. □

### 5.20.3. The proof

We can now finally prove Theorem 3.13.2:

*Proof of Theorem 3.13.2 (sketched).* We must show that $\mathbf{a}$ is conjugate to $S(\mathbf{a})$ in $\mathbf{k}[S_n]$. We know that $n!$ is invertible in $\mathbf{k}$ (since $\mathbf{k}$ is a field of characteristic 0). Hence, by Lemma 5.20.6 (applied to $\mathbf{b} = S(\mathbf{a})$), it suffices to show that $\rho_\lambda(\mathbf{a})$ is conjugate to $\rho_\lambda(S(\mathbf{a}))$ in $\operatorname{End}\left(\mathcal{S}^\lambda\right)$ for each partition $\lambda$ of $n$.

So let $\lambda$ be any partition of $n$. We must prove that $\rho_\lambda(\mathbf{a})$ is conjugate to $\rho_\lambda(S(\mathbf{a}))$ in $\operatorname{End}\left(\mathcal{S}^\lambda\right)$.

If $G$ is any group, if $\mathbf{b} \in \mathbf{k}[G]$ is any element, and if $U$ is any left $\mathbf{k}[G]$-module, then we let "$\tau_{\mathbf{b}}$ on $U$" denote the $\mathbf{k}$-linear map $U \to U$, $u \mapsto \mathbf{b} \cdot u$ (viewed as an endomorphism of the $\mathbf{k}$-module $U$). This generalizes the notations "$\tau_{\mathbf{a}}$ on $V^*$" and "$\tau_{S(\mathbf{a})}$ on $V$" from Proposition 5.19.21 **(b)**.

Note that the endomorphism $\rho_\lambda(\mathbf{a})$ of $\mathcal{S}^\lambda$ is precisely the endomorphism "$\tau_{\mathbf{a}}$ on $\mathcal{S}^\lambda$" (since both endomorphisms send each $v \in \mathcal{S}^\lambda$ to $\mathbf{a} \cdot v$). In other words,

$$\rho_\lambda(\mathbf{a}) = \left(\tau_{\mathbf{a}} \text{ on } \mathcal{S}^\lambda\right). \tag{390}$$

The same reasoning (applied to $S(\mathbf{a})$ instead of $\mathbf{a}$) shows that

$$\rho_\lambda(S(\mathbf{a})) = \left(\tau_{S(\mathbf{a})} \text{ on } \mathcal{S}^\lambda\right). \tag{391}$$

Corollary 5.20.4 (applied to $V = \mathcal{S}^\lambda$ and $f = \rho_\lambda(S(\mathbf{a}))$) shows that the endomorphism $\rho_\lambda(S(\mathbf{a}))$ of $\mathcal{S}^\lambda$ is conjugate to its dual $(\rho_\lambda(S(\mathbf{a})))^*$. In other words,

$$\rho_\lambda(S(\mathbf{a})) \sim (\rho_\lambda(S(\mathbf{a})))^*. \tag{392}$$

But Proposition 5.19.21 **(b)** (applied to $G = S_n$ and $V = \mathcal{S}^\lambda$) yields that

$$\left(\tau_{\mathbf{a}} \text{ on } \left(\mathcal{S}^\lambda\right)^*\right) = \left(\tau_{S(\mathbf{a})} \text{ on } \mathcal{S}^\lambda\right)^* = (\rho_\lambda(S(\mathbf{a})))^*$$

(since (391) yields $\left(\tau_{S(\mathbf{a})} \text{ on } \mathcal{S}^\lambda\right) = \rho_\lambda(S(\mathbf{a}))$).

However, the integer $h^\lambda$ (defined in Definition 5.12.11) is a positive integer, and thus is invertible in $\mathbf{k}$ (again since $\mathbf{k}$ is a field of characteristic 0). Hence, Theorem 5.19.35 shows that the $S_n$-representations $\mathcal{S}^\lambda$ and $\left(\mathcal{S}^\lambda\right)^*$ are isomorphic. In other words, the left $\mathbf{k}[S_n]$-modules $\mathcal{S}^\lambda$ and $\left(\mathcal{S}^\lambda\right)^*$ are isomorphic. Hence, Proposition 5.20.5 (applied to $R = \mathbf{k}[S_n]$ and $U = \mathcal{S}^\lambda$ and $V = \left(\mathcal{S}^\lambda\right)^*$) yields

$$\left(\tau_{\mathbf{a}} \text{ on } \mathcal{S}^\lambda\right) \sim \left(\tau_{\mathbf{a}} \text{ on } \left(\mathcal{S}^\lambda\right)^*\right).$$

Now, we have the following chain of equivalences:

$$\rho_\lambda(\mathbf{a}) = \left(\tau_{\mathbf{a}} \text{ on } \mathcal{S}^\lambda\right) \sim \left(\tau_{\mathbf{a}} \text{ on } \left(\mathcal{S}^\lambda\right)^*\right) = (\rho_\lambda(S(\mathbf{a})))^* \sim \rho_\lambda(S(\mathbf{a}))$$

(by (392)). Hence, $\rho_\lambda(\mathbf{a}) \sim \rho_\lambda(S(\mathbf{a}))$ (since conjugacy of endomorphisms is an equivalence relation). In other words, $\rho_\lambda(\mathbf{a})$ is conjugate to $\rho_\lambda(S(\mathbf{a}))$ in $\operatorname{End}\left(\mathcal{S}^\lambda\right)$. As we explained, this proves Theorem 3.13.2. □

Another proof of Theorem 3.13.2 can be given using the seminormal form, which we might study later.

## 5.21. The $\mathbf{F}_U w_{U,V} \mathbf{E}_V$ basis

Throughout this section, we will use the following conventions:

**Convention 5.21.1.** **(a)** We let $\mathcal{A} = \mathbf{k}[S_n]$ as usual.

**(b)** If $P$ and $Q$ are two $n$-tableaux of the same shape $D$, then the permutation $w_{P,Q} \in S_n$ is defined as in Definition 5.15.2. (That is, it is the unique permutation $w \in S_n$ that satisfies $wQ = P$.)

**(c)** If $T$ is an $n$-tableau of any shape, then we set

$$\mathbf{E}_T := \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \in \mathcal{A} \qquad \text{and} \qquad \mathbf{F}_T := \nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T} \in \mathcal{A}.$$

(See Proposition 5.11.19 for more about $\mathbf{F}_T$.)

### 5.21.1. The basis

We shall next construct another basis of $\mathcal{A} = \mathbf{k}[S_n]$ when $n!$ is invertible in $\mathbf{k}$. While similar to Young's natural basis $(\mathbf{E}_{U,V})_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \mathrm{SYT}(\lambda)}$ from Corollary 5.15.28, it will interact with the antipode $S$ in an even simpler way:

**Theorem 5.21.2.** For any partition $\lambda$ of $n$, let $\mathrm{SYT}(\lambda)$ be the set of all standard $n$-tableaux of shape $Y(\lambda)$. Assume that $n!$ is invertible in $\mathbf{k}$. Then, the family $(\mathbf{F}_U w_{U,V} \mathbf{E}_V)_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \mathrm{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}$.

**Example 5.21.3.** Let $n = 3$, and consider the two $n$-tableaux $U := 12\backslash\backslash 3$ and $V := 13\backslash\backslash 2$ of shape $(2,1)$. Then,

$$\underbrace{\mathbf{F}_U}_{=\nabla_{\mathrm{Row}\,U}\nabla^-_{\mathrm{Col}\,U}} w_{U,V} \underbrace{\mathbf{E}_V}_{=\nabla^-_{\mathrm{Col}\,V}\nabla_{\mathrm{Row}\,V}} = \underbrace{\nabla_{\mathrm{Row}\,U}}_{=1+t_{1,2}} \underbrace{\nabla^-_{\mathrm{Col}\,U}}_{=1-t_{1,3}} \underbrace{w_{U,V}}_{=t_{2,3}} \underbrace{\nabla^-_{\mathrm{Col}\,V}}_{=1-t_{1,2}} \underbrace{\nabla_{\mathrm{Row}\,V}}_{=1+t_{1,3}}$$
$$= -2 - 2s_1 + 4s_2 - 2t_{1,3} + 4\,\mathrm{cyc}_{1,2,3} - 2\,\mathrm{cyc}_{1,3,2}\,.$$

In total, denoting the four standard 3-tableaux of partition shapes by

$$P := 123, \qquad U := 12\backslash\backslash 3, \qquad V := 13\backslash\backslash 2, \qquad N := 1\backslash\backslash 2\backslash\backslash 3,$$

the family $(\mathbf{F}_U w_{U,V} \mathbf{E}_V)_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \mathrm{SYT}(\lambda)}$ looks as follows:

$$\begin{aligned}
\mathbf{F}_P w_{P,P} \mathbf{E}_P &= 6\nabla = 6 + 6s_1 + 6s_2 + 6t_{1,3} + 6\,\mathrm{cyc}_{1,2,3} + 6\,\mathrm{cyc}_{1,3,2};\\
\mathbf{F}_U w_{U,U} \mathbf{E}_U &= 4 + 4s_1 - 2s_2 - 2t_{1,3} - 2\,\mathrm{cyc}_{1,2,3} - 2\,\mathrm{cyc}_{1,3,2};\\
\mathbf{F}_U w_{U,V} \mathbf{E}_V &= -2 - 2s_1 + 4s_2 - 2t_{1,3} + 4\,\mathrm{cyc}_{1,2,3} - 2\,\mathrm{cyc}_{1,3,2};\\
\mathbf{F}_V w_{V,U} \mathbf{E}_U &= -2 - 2s_1 + 4s_2 - 2t_{1,3} - 2\,\mathrm{cyc}_{1,2,3} + 4\,\mathrm{cyc}_{1,3,2};\\
\mathbf{F}_V w_{V,V} \mathbf{E}_V &= 4 - 2s_1 - 2s_2 + 4t_{1,3} - 2\,\mathrm{cyc}_{1,2,3} - 2\,\mathrm{cyc}_{1,3,2};\\
\mathbf{F}_N w_{N,N} \mathbf{E}_N &= 6\nabla^- = 6 - 6s_1 - 6s_2 - 6t_{1,3} + 6\,\mathrm{cyc}_{1,2,3} + 6\,\mathrm{cyc}_{1,3,2}\,.
\end{aligned}$$

The proof of Theorem 5.21.2 will require some work. First, however, let us reveal how the antipode $S$ acts on this basis:

**Proposition 5.21.4.** Let $\lambda$ be a partition of $n$. Let $U$ and $V$ be any two $n$-tableaux of shape $Y(\lambda)$. Then,

$$S(\mathbf{F}_U w_{U,V} \mathbf{E}_V) = \mathbf{F}_V w_{V,U} \mathbf{E}_U.$$

*Proof.* Proposition 5.11.19 **(a)** (applied to $T = V$) shows that $S(\mathbf{E}_V) = \mathbf{F}_V$. Applying the map $S$ to this equality, we obtain $S(S(\mathbf{E}_V)) = S(\mathbf{F}_V)$. But

$S\left(S\left(\mathbf{E}_V\right)\right) = \mathbf{E}_V$ (since Theorem 3.11.14 **(b)** says that $S \circ S = \mathrm{id}$). Comparing these two equalities, we find $S\left(\mathbf{F}_V\right) = \mathbf{E}_V$. Applying the same argument to $U$ instead of $V$, we obtain $S\left(\mathbf{F}_U\right) = \mathbf{E}_U$.

Proposition 5.15.15 **(b)** (applied to $D = Y\left(\lambda\right)$ and $P = U$ and $Q = V$) yields $w_{V,U} = w_{U,V}^{-1}$. However, the definition of the antipode $S$ yields $S\left(w_{U,V}\right) = w_{U,V}^{-1}$. Comparing these two equalities, we find $S\left(w_{U,V}\right) = w_{V,U}$.

We know from Theorem 3.11.14 **(a)** that the antipode $S$ is an algebra antimorphism. Thus,

$$S\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right) = \underbrace{S\left(\mathbf{E}_V\right)}_{=\mathbf{F}_V} \cdot \underbrace{S\left(w_{U,V}\right)}_{=w_{V,U}} \cdot \underbrace{S\left(\mathbf{F}_U\right)}_{=\mathbf{E}_U} = \mathbf{F}_V w_{V,U} \mathbf{E}_U.$$

This proves Proposition 5.21.4. □

The following exercise can be used to speed up the computation of the family $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \mathrm{SYT}(\lambda)}$, as it mostly reduces this computation to picking an arbitrary $n$-tableau $T$ of shape $Y\left(\lambda\right)$ for each partition $\lambda$ of $n$ and computing the products $\nabla_{\mathrm{Row}\, T} \nabla_{\mathrm{Col}\, T}^{-} \nabla_{\mathrm{Row}\, T}$ for these chosen $T$'s:

**Exercise 5.21.1.** [2] Let $\lambda$ be a partition of $n$. Let $T$, $U$ and $V$ be any three $n$-tableaux of shape $Y\left(\lambda\right)$. Prove that

$$\mathbf{F}_U w_{U,V} \mathbf{E}_V = \left|\mathcal{C}\left(T\right)\right| \cdot w_{U,T} \nabla_{\mathrm{Row}\, T} \nabla_{\mathrm{Col}\, T}^{-} \nabla_{\mathrm{Row}\, T} w_{T,V}.$$

(See Definition 5.3.1 for the meaning of $\mathcal{C}\left(T\right)$.)

### 5.21.2. Proof of the basis property

We now take aim at the proof of Theorem 5.21.2. A few simple lemmas pave the way:

**Lemma 5.21.5.** Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $Y\left(\lambda\right)$. Then,

$$\mathbf{F}_T \mathbf{E}_T \mathbf{F}_T = \left|\mathcal{C}\left(T\right)\right| \cdot \left|\mathcal{R}\left(T\right)\right| \cdot h^\lambda \cdot \mathbf{F}_T.$$

(See Definition 5.12.11 for the meaning of $h^\lambda$, and Definition 5.3.1 for the meanings of $\mathcal{R}\left(T\right)$ and $\mathcal{C}\left(T\right)$.)

*Proof.* We know that $T$ is an $n$-tableau of shape $Y\left(\lambda\right)$, that is, an $n$-tableau of shape $\lambda$. Hence, Proposition 5.11.19 **(b)** yields

$$\mathbf{F}_T^2 = \underbrace{\frac{n!}{f^\lambda}}_{\substack{=h^\lambda \\ \text{(by the definition of } h^\lambda\text{)}}} \mathbf{F}_T = h^\lambda \cdot \mathbf{F}_T. \tag{393}$$

On the other hand, $\mathbf{F}_T = \nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T}$ (by the definition of $\mathbf{F}_T$) and $\mathbf{E}_T = \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$ (by the definition of $\mathbf{E}_T$). Hence,

$$\begin{aligned}
\mathbf{F}_T \mathbf{E}_T \mathbf{F}_T &= \left(\nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T}\right) \left(\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}\right) \left(\nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T}\right) \\
&= \nabla_{\operatorname{Row} T} \quad \underbrace{\left(\nabla^-_{\operatorname{Col} T}\right)^2}_{\substack{=|\mathcal{C}(T)| \cdot \nabla^-_{\operatorname{Col} T} \\ \text{(by Corollary 5.5.6 \textbf{(b)})}}} \quad \underbrace{\left(\nabla_{\operatorname{Row} T}\right)^2}_{\substack{=|\mathcal{R}(T)| \cdot \nabla_{\operatorname{Row} T} \\ \text{(by Corollary 5.5.6 \textbf{(a)})}}} \quad \nabla^-_{\operatorname{Col} T} \\
&= \nabla_{\operatorname{Row} T} |\mathcal{C}(T)| \cdot \nabla^-_{\operatorname{Col} T} \cdot |\mathcal{R}(T)| \cdot \nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T} \\
&= |\mathcal{C}(T)| \cdot |\mathcal{R}(T)| \cdot \underbrace{\nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T}}_{=\mathbf{F}_T} \cdot \underbrace{\nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T}}_{=\mathbf{F}_T} \\
&= |\mathcal{C}(T)| \cdot |\mathcal{R}(T)| \cdot \underbrace{\mathbf{F}_T \mathbf{F}_T}_{\substack{=\mathbf{F}_T^2 = h^\lambda \cdot \mathbf{F}_T \\ \text{(by (393))}}} = |\mathcal{C}(T)| \cdot |\mathcal{R}(T)| \cdot h^\lambda \cdot \mathbf{F}_T.
\end{aligned}$$

This proves Lemma 5.21.5. □

**Lemma 5.21.6.** Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $Y(\lambda)$. Assume that $n!$ is invertible in $\mathbf{k}$. Then,

$$\mathbf{F}_T \mathcal{A} \mathbf{E}_T = \mathbf{k} \cdot \mathbf{F}_T \mathbf{E}_T.$$

(Here, as usual, for any $\mathbf{a} \in \mathcal{A}$, we let $\mathbf{k} \cdot \mathbf{a}$ denote the $\mathbf{k}$-submodule $\{u\mathbf{a} \mid u \in \mathbf{k}\}$ of $\mathcal{A}$, that is, the $\mathbf{k}$-linear span of $\{\mathbf{a}\}$.)

*Proof.* We have $\mathbf{F}_T \mathbf{E}_T = \mathbf{F}_T \underbrace{1_{\mathcal{A}}}_{\in \mathcal{A}} \mathbf{E}_T \in \mathbf{F}_T \mathcal{A} \mathbf{E}_T$ and thus

$$\mathbf{k} \cdot \underbrace{\mathbf{F}_T \mathbf{E}_T}_{\in \mathbf{F}_T \mathcal{A} \mathbf{E}_T} \in \mathbf{k} \cdot \mathbf{F}_T \mathcal{A} \mathbf{E}_T \subseteq \mathbf{F}_T \mathcal{A} \mathbf{E}_T \tag{394}$$

(since $\mathbf{F}_T \mathcal{A} \mathbf{E}_T$ is a $\mathbf{k}$-module).

On the other hand, Lemma 5.21.5 yields

$$\mathbf{F}_T \mathbf{E}_T \mathbf{F}_T = |\mathcal{C}(T)| \cdot |\mathcal{R}(T)| \cdot h^\lambda \cdot \mathbf{F}_T, \tag{395}$$

where $h^\lambda$ and $\mathcal{R}(T)$ and $\mathcal{C}(T)$ are defined as in Lemma 5.21.5.

But $h^\lambda = \dfrac{n!}{f^\lambda}$ (by the definition of $h^\lambda$). Thus, $h^\lambda$ is a divisor of $n!$. Hence, $h^\lambda$ is invertible in $\mathbf{k}$ (since $n!$ is invertible in $\mathbf{k}$). Furthermore, $\mathcal{R}(T)$ is a subgroup of $S_n$ (by its definition). Hence, Lagrange's theorem shows that $|\mathcal{R}(T)|$ is a divisor of $|S_n| = n!$. Thus, $|\mathcal{R}(T)|$ is invertible in $\mathbf{k}$ (since $n!$ is invertible in $\mathbf{k}$). Similarly, we can see that $|\mathcal{C}(T)|$ is invertible in $\mathbf{k}$. Now, we can divide the equality (395) by the three scalars $|\mathcal{C}(T)|$, $|\mathcal{R}(T)|$ and $h^\lambda$ (since we have

just showed that these scalars $|\mathcal{C}(T)|$, $|\mathcal{R}(T)|$ and $h^\lambda$ are invertible in $\mathbf{k}$). As a result, we obtain

$$\frac{1}{|\mathcal{C}(T)| \cdot |\mathcal{R}(T)| \cdot h^\lambda} \cdot \mathbf{F}_T \mathbf{E}_T \mathbf{F}_T = \mathbf{F}_T. \tag{396}$$

Now, we shall show that $\mathbf{F}_T \mathcal{A} \mathbf{E}_T \subseteq \mathbf{k} \cdot \mathbf{F}_T \mathbf{E}_T$. Indeed, fix $\mathbf{b} \in \mathcal{A}$. Then, $\mathbf{F}_T \mathbf{b} \in \mathcal{A} = \mathbf{k}[S_n]$. Hence, Proposition 5.11.5 (applied to $\mathbf{a} = \mathbf{F}_T \mathbf{b}$) shows that $\mathbf{E}_T \mathbf{F}_T \mathbf{b} \mathbf{E}_T = \kappa \mathbf{E}_T$ for some some scalar $\kappa \in \mathbf{k}$. Consider this $\kappa$. We have

$$\begin{aligned}
\mathbf{b}\mathbf{E}_T &= \frac{1}{|\mathcal{C}(T)| \cdot |\mathcal{R}(T)| \cdot h^\lambda} \cdot \underbrace{\mathbf{F}_T}_{= \frac{1}{|\mathcal{C}(T)| \cdot |\mathcal{R}(T)| \cdot h^\lambda} \cdot \mathbf{F}_T \mathbf{E}_T \mathbf{F}_T \text{ (by (396))}} \underbrace{\mathbf{E}_T \mathbf{F}_T \mathbf{b} \mathbf{E}_T}_{= \kappa \mathbf{E}_T} \\
&= \frac{1}{|\mathcal{C}(T)| \cdot |\mathcal{R}(T)| \cdot h^\lambda} \cdot \mathbf{F}_T \cdot \kappa \mathbf{E}_T \\
&= \underbrace{\frac{\kappa}{|\mathcal{C}(T)| \cdot |\mathcal{R}(T)| \cdot h^\lambda}}_{\in \mathbf{k}} \cdot \mathbf{F}_T \mathbf{E}_T \\
&\in \mathbf{k} \cdot \mathbf{F}_T \mathbf{E}_T.
\end{aligned}$$

Forget that we fixed $\mathbf{b}$. We thus have shown that $\mathbf{F}_T \mathbf{b} \mathbf{E}_T \in \mathbf{k} \cdot \mathbf{F}_T \mathbf{E}_T$ for each $\mathbf{b} \in \mathcal{A}$. In other words,

$$\{\mathbf{F}_T \mathbf{b} \mathbf{E}_T \mid \mathbf{b} \in \mathcal{A}\} \subseteq \mathbf{k} \cdot \mathbf{F}_T \mathbf{E}_T.$$

In other words, $\mathbf{F}_T \mathcal{A} \mathbf{E}_T \subseteq \mathbf{k} \cdot \mathbf{F}_T \mathbf{E}_T$ (since $\mathbf{F}_T \mathcal{A} \mathbf{E}_T = \{\mathbf{F}_T \mathbf{b} \mathbf{E}_T \mid \mathbf{b} \in \mathcal{A}\}$). Combining this with (394), we obtain $\mathbf{F}_T \mathcal{A} \mathbf{E}_T = \mathbf{k} \cdot \mathbf{F}_T \mathbf{E}_T$. Thus, Lemma 5.21.6 is proved. □

**Lemma 5.21.7.** Let $\lambda$ be a partition of $n$. Let $U$ and $V$ be two $n$-tableaux of shape $Y(\lambda)$. Assume that $n!$ is invertible in $\mathbf{k}$. Then,

$$\mathbf{F}_U \mathcal{A} \mathbf{E}_V = \mathbf{k} \cdot \mathbf{F}_U w_{U,V} \mathbf{E}_V.$$

*Proof.* Proposition 5.15.7 (applied to $D = Y(\lambda)$ and $P = U$ and $Q = V$) yields

$$\mathbf{E}_{U,V} = w_{U,V} \mathbf{E}_V = \mathbf{E}_U w_{U,V}.$$

Hence, in particular, $w_{U,V} \mathbf{E}_V = \mathbf{E}_U w_{U,V}$.

Moreover, $w_{U,V}$ is a permutation in $S_n$, thus an invertible element of $\mathcal{A}$ (since any element of a group $G$ is invertible in the group algebra $\mathbf{k}[G]$). Thus, $\mathcal{A} w_{U,V} = \mathcal{A}$ (since any invertible element $\mathbf{p}$ of $\mathcal{A}$ satisfies $\mathcal{A}\mathbf{p} = \mathcal{A}$). Hence,

$\mathcal{A} = \mathcal{A} w_{U,V}$, so that

$$\mathbf{F}_U \underbrace{\mathcal{A}}_{=\mathcal{A}w_{U,V}} \mathbf{E}_V = \mathbf{F}_U \mathcal{A} \underbrace{w_{U,V} \mathbf{E}_V}_{=\mathbf{E}_U w_{U,V}} = \underbrace{\mathbf{F}_U \mathcal{A} \mathbf{E}_U}_{\substack{=\mathbf{k}\cdot\mathbf{F}_U\mathbf{E}_U \\ \text{(by Lemma 5.21.6,} \\ \text{applied to } T=U)}} w_{U,V} = \mathbf{k} \cdot \mathbf{F}_U \underbrace{\mathbf{E}_U w_{U,V}}_{=w_{U,V}\mathbf{E}_V}$$
$$= \mathbf{k} \cdot \mathbf{F}_U w_{U,V} \mathbf{E}_V.$$

This proves Lemma 5.21.7. □

Our next two lemmas are general linear-algebraic facts about free modules of finite rank. When $\mathbf{k}$ is a field, they are both well-known from any good course on linear algebra, but they are in fact true for any commutative ring $\mathbf{k}$.

**Lemma 5.21.8.** Let $r \in \mathbb{N}$. Let $V$ and $W$ be two free $\mathbf{k}$-modules of rank $r$. Let $f : V \to W$ be a surjective $\mathbf{k}$-linear map. Then, $f$ is a $\mathbf{k}$-module isomorphism.

*Proof.* Both $\mathbf{k}$-modules $V$ and $W$ are free of rank $r$. Hence, they have bases $(v_1, v_2, \ldots, v_r)$ and $(w_1, w_2, \ldots, w_r)$ of size $r$. Consider these bases.

The map $f$ is surjective. In other words, $W = f(V)$. Hence, for each $i \in \{1, 2, \ldots, r\}$, there exists some $x_i \in V$ such that $w_i = f(x_i)$ (since $w_i \in W = f(V)$). Consider this $x_i$. Thus, we have constructed $r$ vectors $x_1, x_2, \ldots, x_r$ in $V$.

Recall that $(w_1, w_2, \ldots, w_r)$ is a basis of the $\mathbf{k}$-module $W$. Hence, we can define a $\mathbf{k}$-linear map $g : W \to V$ by setting

$$(g(w_i) = x_i \qquad \text{for each } i \in \{1, 2, \ldots, r\}) \tag{397}$$

(since a $\mathbf{k}$-linear map from a free $\mathbf{k}$-module can be defined by choosing its values on the elements of a given basis). Consider this map $g$.

Now, for each $i \in \{1, 2, \ldots, r\}$, we have

$$\begin{aligned}(f \circ g)(w_i) &= f(g(w_i)) \\ &= f(x_i) \qquad (\text{since (397) yields } g(w_i) = x_i) \\ &= w_i \qquad (\text{since } x_i \text{ was defined to satisfy } w_i = f(x_i)) \\ &= \mathrm{id}_W(w_i).\end{aligned}$$

Thus, the two $\mathbf{k}$-linear maps $f \circ g : W \to W$ and $\mathrm{id}_W : W \to W$ agree on each of the $r$ vectors $w_1, w_2, \ldots, w_r$. In other words, these two $\mathbf{k}$-linear maps agree on the basis $(w_1, w_2, \ldots, w_r)$ of $W$. Hence, these two maps must be identical (because if two $\mathbf{k}$-linear maps agree on a basis of their domain, then they must be identical). In other words, $f \circ g = \mathrm{id}_W$. Hence, Lemma 5.14.2 yields that $g \circ f = \mathrm{id}_V$. Combining this with $f \circ g = \mathrm{id}_W$, we conclude that the maps $f$ and $g$ are mutually inverse. Hence, the map $f$ is invertible, and hence a $\mathbf{k}$-module isomorphism (since it is $\mathbf{k}$-linear). This proves Lemma 5.21.8. □

**Lemma 5.21.9.** Let $r \in \mathbb{N}$. Let $M$ be a free $\mathbf{k}$-module of rank $r$. Let $(m_i)_{i \in I}$ be a family of vectors in $M$ with $|I| = r$. Assume that this family $(m_i)_{i \in I}$ spans $M$. Then, $(m_i)_{i \in I}$ is a basis of $M$.

*Proof.* The set $I$ has $r$ elements (since $|I| = r$), and is only used for indexing our family $(m_i)_{i \in I}$. Thus, we can WLOG assume that $I = \{1, 2, \ldots, r\}$ (since we can otherwise rename the $r$ elements of $I$ as $1, 2, \ldots, r$ and accordingly rename the entries of our family $(m_i)_{i \in I}$, without changing the essence of anything). Assume this. Then, the family $(m_i)_{i \in I} = (m_i)_{i \in \{1,2,\ldots,r\}}$ is just an $r$-tuple $(m_1, m_2, \ldots, m_r)$. We have assumed that this family $(m_i)_{i \in I}$ spans $M$. In other words, the $r$-tuple $(m_1, m_2, \ldots, m_r)$ spans $M$.

But $M$ is a free $\mathbf{k}$-module of rank $r$. Hence, $M$ has a basis $(b_1, b_2, \ldots, b_r)$ consisting of $r$ vectors. Consider this basis.

Thus, $(b_1, b_2, \ldots, b_r)$ is a basis of $M$. Hence, we can define a $\mathbf{k}$-linear map $f : M \to M$ by setting

$$(f(b_i) = m_i \qquad \text{for each } i \in \{1, 2, \ldots, r\})$$

(since a $\mathbf{k}$-linear map from a free $\mathbf{k}$-module can be defined by choosing its values on the elements of a given basis). Consider this map $f$.

The definition of $f$ shows that $f(b_i) = m_i$ for each $i \in \{1, 2, \ldots, r\}$. Hence,

$$(f(b_1), f(b_2), \ldots, f(b_r)) = (m_1, m_2, \ldots, m_r). \tag{398}$$

Since $(b_1, b_2, \ldots, b_r)$ is a basis of $M$, we have $M = \operatorname{span}\{b_1, b_2, \ldots, b_r\}$. Applying the map $f$ to this equality, we obtain

$$\begin{aligned} f(M) &= f(\operatorname{span}\{b_1, b_2, \ldots, b_r\}) \\ &= \operatorname{span} \underbrace{\{f(b_1), f(b_2), \ldots, f(b_r)\}}_{\substack{= \{m_1, m_2, \ldots, m_r\} \\ \text{(by (398))}}} \\ &\qquad \left( \begin{array}{c} \text{since the map } f \text{ is } \mathbf{k}\text{-linear, and thus} \\ \text{respects } \mathbf{k}\text{-linear combinations} \end{array} \right) \\ &= \operatorname{span}\{m_1, m_2, \ldots, m_r\} = M \end{aligned}$$

(since the $r$-tuple $(m_1, m_2, \ldots, m_r)$ spans $M$). In other words, the map $f$ is surjective. Hence, Lemma 5.21.8 (applied to $V = M$ and $W = M$) shows that this map $f$ is a $\mathbf{k}$-module isomorphism. Therefore, $f$ must send any basis of $M$ to a basis of $M$ (since a $\mathbf{k}$-module isomorphism always sends a basis of its domain to a basis of its target). Since $(b_1, b_2, \ldots, b_r)$ is a basis of $M$, we thus conclude that $(f(b_1), f(b_2), \ldots, f(b_r))$ is a basis of $M$ as well (since $f$ sends the basis $(b_1, b_2, \ldots, b_r)$ to $(f(b_1), f(b_2), \ldots, f(b_r))$). In other words, $(m_1, m_2, \ldots, m_r)$ is a basis of $M$ (since (398) shows that $(m_1, m_2, \ldots, m_r) = (f(b_1), f(b_2), \ldots, f(b_r))$). In other words, $(m_i)_{i \in I}$ is a basis of $M$ (since the family $(m_i)_{i \in I}$ is just the $r$-tuple $(m_1, m_2, \ldots, m_r)$). This proves Lemma 5.21.9. $\square$

Next, we state a version of Theorem 5.21.2 for a specific partition:

**Proposition 5.21.10.** Fix a partition $\lambda$ of $n$. Let $\operatorname{SYT}(\lambda)$ be the set of all standard $n$-tableaux of shape $Y(\lambda)$. Assume that $n!$ is invertible in $\mathbf{k}$. Then, the family $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{U,V \in \operatorname{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$.

*Proof.* Recall Definition 5.12.11. In particular, $h^\lambda = \dfrac{n!}{f^\lambda}$ (by the definition of $h^\lambda$). Thus, $h^\lambda$ is a divisor of $n!$. Hence, $h^\lambda$ is invertible in $\mathbf{k}$ (since $n!$ is invertible in $\mathbf{k}$). Furthermore, note that $|\operatorname{SYT}(\lambda)| = f^\lambda$ [138].

Let $D = Y(\lambda)$. Thus, the standard $n$-tableaux of shape $D$ are just the standard $n$-tableaux of shape $Y(\lambda)$. Hence, the set $\operatorname{SYT}(D)$ in Theorem 5.15.20 (defined as the set of all standard $n$-tableaux of shape $D$) is precisely our set $\operatorname{SYT}(\lambda)$ (defined as the set of all standard $n$-tableaux of shape $Y(\lambda)$). Hence, Theorem 5.15.20 (with $\operatorname{SYT}(D)$ renamed as $\operatorname{SYT}(\lambda)$) shows that

$$\begin{aligned} \mathcal{A}\mathbf{E}_\lambda &= \bigoplus_{T \in \operatorname{SYT}(\lambda)} \mathcal{A}\mathbf{E}_T \qquad \text{(an internal direct sum)} \\ &= \sum_{T \in \operatorname{SYT}(\lambda)} \mathcal{A}\mathbf{E}_T \end{aligned} \tag{399}$$

(since direct sums are sums).

On the other hand, Proposition 5.12.3 yields $\mathbf{E}_\lambda \in Z(\mathbf{k}[S_n])$. Hence, Proposition 3.13.3 (applied to $\mathbf{a} = \mathbf{E}_\lambda$) yields $S(\mathbf{E}_\lambda) = \mathbf{E}_\lambda$. Furthermore, $S : \mathcal{A} \to \mathcal{A}$ is an involution (by Theorem 3.11.14 **(b)**), and thus a bijection. Hence, $S(\mathcal{A}) = \mathcal{A}$. Moreover, $S$ is an algebra anti-morphism (by Theorem 3.11.14 **(a)**). Therefore,

$$S(\mathcal{A}\mathbf{E}_\lambda) = \underbrace{S(\mathbf{E}_\lambda)}_{=\mathbf{E}_\lambda} \underbrace{S(\mathcal{A})}_{=\mathcal{A}} = \mathbf{E}_\lambda \mathcal{A}.$$

Hence,

$$\begin{aligned} \mathbf{E}_\lambda \mathcal{A} = S(\mathcal{A}\mathbf{E}_\lambda) &= S\left(\sum_{T \in \operatorname{SYT}(\lambda)} \mathcal{A}\mathbf{E}_T\right) \qquad \text{(by (399))} \\ &= \sum_{T \in \operatorname{SYT}(\lambda)} \underbrace{S(\mathbf{E}_T)}_{=\mathbf{F}_T} \underbrace{S(\mathcal{A})}_{=\mathcal{A}} \qquad \begin{pmatrix} \text{since } S \text{ is an algebra} \\ \text{anti-morphism} \end{pmatrix} \\ &\qquad \text{(by Proposition 5.11.19 \textbf{(a)})} \\ &= \sum_{T \in \operatorname{SYT}(\lambda)} \mathbf{F}_T \mathcal{A}. \end{aligned}$$

[138] *Proof.* The set of all standard $n$-tableaux of shape $Y(\lambda)$ was denoted by $\operatorname{SYT}(Y(\lambda))$ in the proof of Corollary 5.15.13, but we are now denoting it by $\operatorname{SYT}(\lambda)$. Hence, $\operatorname{SYT}(Y(\lambda)) = \operatorname{SYT}(\lambda)$.

However, in our proof of Corollary 5.15.13, we saw that $f^\lambda = |\operatorname{SYT}(Y(\lambda))|$. In view of $\operatorname{SYT}(Y(\lambda)) = \operatorname{SYT}(\lambda)$, we can rewrite this as $f^\lambda = |\operatorname{SYT}(\lambda)|$. In other words, $|\operatorname{SYT}(\lambda)| = f^\lambda$.

Multiplying this equality with (399), we obtain

$$\begin{aligned}
(\mathbf{E}_\lambda \mathcal{A})(\mathcal{A}\mathbf{E}_\lambda) &= \left(\sum_{T\in \mathrm{SYT}(\lambda)} \mathbf{F}_T \mathcal{A}\right)\left(\sum_{T\in \mathrm{SYT}(\lambda)} \mathcal{A}\mathbf{E}_T\right) \\
&= \left(\sum_{U\in \mathrm{SYT}(\lambda)} \mathbf{F}_U \mathcal{A}\right)\left(\sum_{V\in \mathrm{SYT}(\lambda)} \mathcal{A}\mathbf{E}_V\right) \\
&\qquad \left(\begin{array}{l}\text{here, we have renamed the summation}\\ \text{indices } T \text{ and } T \text{ as } U \text{ and } V\end{array}\right) \\
&= \underbrace{\sum_{U\in \mathrm{SYT}(\lambda)} \sum_{V\in \mathrm{SYT}(\lambda)}}_{= \sum_{(U,V)\in \mathrm{SYT}(\lambda)\times \mathrm{SYT}(\lambda)}} \underbrace{(\mathbf{F}_U \mathcal{A})(\mathcal{A}\mathbf{E}_V)}_{=\mathbf{F}_U \mathcal{A}\mathcal{A}\mathbf{E}_V} \\
&= \sum_{(U,V)\in \mathrm{SYT}(\lambda)\times \mathrm{SYT}(\lambda)} \mathbf{F}_U \underbrace{\mathcal{A}\mathcal{A}}_{=\mathcal{A}} \mathbf{E}_V \\
&= \sum_{(U,V)\in \mathrm{SYT}(\lambda)\times \mathrm{SYT}(\lambda)} \underbrace{\mathbf{F}_U \mathcal{A}\mathbf{E}_V}_{=\mathbf{k}\cdot \mathbf{F}_U w_{U,V}\mathbf{E}_V \text{ (by Lemma 5.21.7)}} \\
&= \sum_{(U,V)\in \mathrm{SYT}(\lambda)\times \mathrm{SYT}(\lambda)} \underbrace{\mathbf{k}\cdot \mathbf{F}_U w_{U,V}\mathbf{E}_V}_{=\mathrm{span}\{\mathbf{F}_U w_{U,V}\mathbf{E}_V\} \text{ (since } \mathbf{k}\cdot v = \mathrm{span}\{v\} \text{ for any vector } v)} \\
&= \sum_{(U,V)\in \mathrm{SYT}(\lambda)\times \mathrm{SYT}(\lambda)} \mathrm{span}\left\{\mathbf{F}_U w_{U,V}\mathbf{E}_V\right\} \\
&= \mathrm{span}\left\{\mathbf{F}_U w_{U,V}\mathbf{E}_V \mid (U,V)\in \mathrm{SYT}(\lambda)\times \mathrm{SYT}(\lambda)\right\}.
\end{aligned}$$

In view of

$$(\mathbf{E}_\lambda \mathcal{A})(\mathcal{A}\mathbf{E}_\lambda) = \mathbf{E}_\lambda \underbrace{\mathcal{A}\mathcal{A}}_{=\mathcal{A}} \mathbf{E}_\lambda = \underbrace{\mathbf{E}_\lambda \mathcal{A}}_{=\mathcal{A}\mathbf{E}_\lambda \text{ (since } \mathbf{E}_\lambda \in Z(\mathbf{k}[S_n])\text{)}} \mathbf{E}_\lambda$$

$$= \mathcal{A} \underbrace{\mathbf{E}_\lambda \mathbf{E}_\lambda}_{=\mathbf{E}_\lambda^2 = \left(h^\lambda\right)^2 \mathbf{E}_\lambda \text{ (by Theorem 5.12.14)}} = \underbrace{\mathcal{A}\left(h^\lambda\right)^2}_{=\mathcal{A} \text{ (since } \left(h^\lambda\right)^2 \text{ is invertible in } \mathbf{k} \text{ (because } h^\lambda \text{ is invertible in } \mathbf{k})\text{)}} \mathbf{E}_\lambda = \mathcal{A}\mathbf{E}_\lambda,$$

we can rewrite this as

$$\mathcal{A}\mathbf{E}_\lambda = \mathrm{span}\left\{\mathbf{F}_U w_{U,V}\mathbf{E}_V \mid (U,V)\in \mathrm{SYT}(\lambda)\times \mathrm{SYT}(\lambda)\right\}.$$

In other words, the family $\left(\mathbf{F}_U w_{U,V}\mathbf{E}_V\right)_{(U,V)\in \mathrm{SYT}(\lambda)\times \mathrm{SYT}(\lambda)}$ spans the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$. This family has size

$$|\mathrm{SYT}(\lambda)\times \mathrm{SYT}(\lambda)| = \underbrace{|\mathrm{SYT}(\lambda)|}_{=f^\lambda} \cdot \underbrace{|\mathrm{SYT}(\lambda)|}_{=f^\lambda} = f^\lambda \cdot f^\lambda = \left(f^\lambda\right)^2.$$

But Lemma 5.15.29 shows that $\mathcal{A}\mathbf{E}_\lambda$ is a free $\mathbf{k}$-module of rank $\left(f^\lambda\right)^2$.

Hence, Lemma 5.21.9 (applied to $r = \left(f^\lambda\right)^2$ and $M = \mathcal{A}\mathbf{E}_\lambda$ and $I = \mathrm{SYT}(\lambda) \times \mathrm{SYT}(\lambda)$ and $(m_i)_{i\in I} = \left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{(U,V)\in \mathrm{SYT}(\lambda)\times\mathrm{SYT}(\lambda)}$) shows that the family $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{(U,V)\in \mathrm{SYT}(\lambda)\times\mathrm{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$ (since we have shown that this family spans $\mathcal{A}\mathbf{E}_\lambda$, and since its size is $|\mathrm{SYT}(\lambda) \times \mathrm{SYT}(\lambda)| = \left(f^\lambda\right)^2$). In other words, the family $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{U,V\in \mathrm{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$ (since "$U, V \in \mathrm{SYT}(\lambda)$" is just shorthand for "$(U, V) \in \mathrm{SYT}(\lambda) \times \mathrm{SYT}(\lambda)$"). This proves Proposition 5.21.10. □

We are now ready for the proof of Theorem 5.21.2:

*Proof of Theorem 5.21.2.* This is very similar to the proof of Corollary 5.15.28.

Corollary 5.12.25 shows that

$$\mathcal{A} = \bigoplus_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda \qquad \text{(an internal direct sum)}.$$

Thus, the sum $\sum_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda$ is direct.

From Proposition 5.21.10, we know that the family $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{U,V\in \mathrm{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$ whenever $\lambda$ is a partition of $n$. Thus, Observation 1 from our above proof of Corollary 5.15.28 can be applied to to $M = \mathcal{A}$ and $I = \{\text{partitions of } n\}$ and $N_\lambda = \mathcal{A}\mathbf{E}_\lambda$ and $\left(v_{\lambda,j}\right)_{j\in J_\lambda} = \left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{U,V\in \mathrm{SYT}(\lambda)}$ (since the sum $\sum_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda$ is direct, and since each $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$ has a basis $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{U,V\in \mathrm{SYT}(\lambda)}$). Thus, we conclude that the internal direct sum $\bigoplus_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda$ has a basis $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{\lambda \text{ is a partition of } n, \text{ and } U,V\in \mathrm{SYT}(\lambda)}$. In other words, the $\mathbf{k}$-module $\mathcal{A}$ has a basis $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{\lambda \text{ is a partition of } n, \text{ and } U,V\in \mathrm{SYT}(\lambda)}$ (since $\mathcal{A} = \bigoplus_{\lambda \text{ is a partition of } n} \mathcal{A}\mathbf{E}_\lambda$). This proves Theorem 5.21.2. □

The basis $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{\lambda \text{ is a partition of } n, \text{ and } U,V\in \mathrm{SYT}(\lambda)}$ of $\mathcal{A}$ (when $n!$ is invertible in $\mathbf{k}$) will be discussed further in Section 6.5.

### 5.21.3. A combinatorial consequence: the sum of all $f^\lambda$

Theorem 5.21.2 leads to a combinatorial identity in the vein of Corollary 5.12.18:

**Corollary 5.21.11.** We have

$$\sum_{\lambda \text{ is a partition of } n} f^\lambda = (\#\text{ of involutions } w \in S_n).$$

(We recall that an *involution* means a permutation $w$ satisfying $w^2 = \mathrm{id}$, or, equivalently, $w^{-1} = w$. Also recall Definition 5.12.11.)

**Example 5.21.12.** Let $n = 4$. Then, the partitions of $n$ are $(4)$, $(3,1)$, $(2,2)$, $(2,1,1)$ and $(1,1,1,1)$, and the corresponding numbers $f^\lambda$ are

$$f^{(4)} = 1, \qquad f^{(3,1)} = 3, \qquad f^{(2,2)} = 2, \qquad f^{(2,1,1)} = 3, \qquad f^{(1,1,1,1)} = 1.$$

Meanwhile, the involutions $w \in S_4$ are id, the six transpositions $t_{i,j}$ and the three products $t_{1,2}t_{3,4}$ and $t_{1,3}t_{2,4}$ and $t_{1,4}t_{2,3}$, which make altogether 10 involutions. Hence, Corollary 5.21.11 says that $1 + 3 + 2 + 3 + 1 = 10$. And this is indeed true.

To prove Corollary 5.21.11, we need a basic lemma about traces of matrices:

**Lemma 5.21.13.** Let $M$ be a free $\mathbf{k}$-module with basis $(m_i)_{i \in I}$, where $I$ is a finite set. Let $f : I \to I$ be any map. Let $\phi : M \to M$ be a $\mathbf{k}$-linear map that satisfies

$$\left(\phi(m_i) = m_{f(i)} \qquad \text{for all } i \in I\right). \tag{400}$$

(This condition actually determines $\phi$ uniquely, but this is immaterial here.) Then, the trace of this map $\phi$ is

$$\operatorname{Tr}\phi = (\#\text{ of all } i \in I \text{ satisfying } f(i) = i) \cdot 1_{\mathbf{k}}.$$

**Example 5.21.14.** Let $M = \mathbf{k}^5$; this is a free $\mathbf{k}$-module with basis $(e_1, e_2, e_3, e_4, e_5) = (e_i)_{i \in [5]}$. Let $f : [5] \to [5]$ be the map that sends the numbers $1, 2, 3, 4, 5$ to $2, 4, 3, 3, 5$, respectively. Let $\phi : M \to M$ be the $\mathbf{k}$-linear map that sends each basis vector $e_i$ to $e_{f(i)}$ (that is, that satisfies $\phi(e_i) = e_{f(i)}$ for all $i \in [5]$). Then, Lemma 5.21.13 (applied to $I = [5]$ and $m_i = e_i$) says that

$$\operatorname{Tr}\phi = \underbrace{(\#\text{ of all } i \in [5] \text{ satisfying } f(i) = i)}_{\substack{=2 \\ \text{(since these } i \in [5] \text{ are 3 and 5)}}} \cdot 1_{\mathbf{k}} = 2 \cdot 1_{\mathbf{k}}.$$

And indeed, this can be easily checked: The matrix that represents the $\mathbf{k}$-linear map $\phi$ with respect to the basis $(e_1, e_2, e_3, e_4, e_5)$ is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1_{\mathbf{k}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1_{\mathbf{k}} & 1_{\mathbf{k}} & 0 \\ 0 & 1_{\mathbf{k}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1_{\mathbf{k}} \end{pmatrix},$$

and thus its trace is $0 + 0 + 1_{\mathbf{k}} + 0 + 1_{\mathbf{k}} = 2 \cdot 1_{\mathbf{k}}$. This example illustrates why Lemma 5.21.13 holds in general: The matrix that represents the $\mathbf{k}$-linear map $\phi$ with respect to the basis $(e_1, e_2, e_3, e_4, e_5)$ (or $(m_i)_{i \in I}$ in the general case) consists only of 0's and $1_{\mathbf{k}}$'s, and the $1_{\mathbf{k}}$'s lie precisely in the $(f(i), i)$-th

cells of the matrix (for all $i \in I$). Thus, the diagonal entries of this matrix are 0's and $1_{\mathbf{k}}$'s, with a $1_{\mathbf{k}}$ for each $i \in I$ that satisfies $f(i) = i$. Hence, the sum of these diagonal entries (i.e., the trace of the matrix) is a sum of $1_{\mathbf{k}}$'s for each $i \in I$ that satisfies $f(i) = i$. Of course, this sum is precisely $(\#\text{ of all } i \in I \text{ satisfying } f(i) = i) \cdot 1_{\mathbf{k}}$.

*Proof of Lemma 5.21.13.* Let us first assume that $I = [r]$ for some $r \in \mathbb{N}$. (We will discuss the general case further below.)

We have $I = [r]$. Thus, the family $(m_i)_{i \in I} = (m_i)_{i \in [r]}$ is just an $r$-tuple $(m_1, m_2, \ldots, m_r)$. Hence, $(m_1, m_2, \ldots, m_r)$ is a basis of $M$ (since $(m_i)_{i \in I}$ is a basis of $M$). Let $A$ be the matrix that represents the $\mathbf{k}$-linear map $\phi : M \to M$ with respect to this basis $(m_1, m_2, \ldots, m_r)$. Then, $\operatorname{Tr} \phi = \operatorname{Tr} A$ (by the definition of the trace of $\phi$).

For any vector $u \in M$ and any $i \in [r]$, we let $[m_i]\, u$ denote the $m_i$-coordinate of $u$ with respect to the basis $(m_1, m_2, \ldots, m_r)$ (that is, the coefficient of $m_i$ when $u$ is expanded in the basis $(m_1, m_2, \ldots, m_r)$ of $M$). Then, the matrix $A$ can be written explicitly as

$$A = \left( [m_i] \left( \phi \left( m_j \right) \right) \right)_{i,j \in [r]}$$

(since $A$ is the matrix that represents the $\mathbf{k}$-linear map $\phi : M \to M$ with respect to the basis $(m_1, m_2, \ldots, m_r)$). Hence,

$$\operatorname{Tr} A = \operatorname{Tr} \left( \left( [m_i] \left( \phi \left( m_j \right) \right) \right)_{i,j \in [r]} \right) = \sum_{i=1}^{r} [m_i] \left( \phi \left( m_i \right) \right)$$

(since the trace of a matrix is defined as the sum of its diagonal entries). In

view of $\operatorname{Tr}\phi = \operatorname{Tr} A$, we can rewrite this as

$$\begin{aligned}
\operatorname{Tr}\phi &= \sum_{i=1}^{r} [m_i] \left( \underbrace{\phi(m_i)}_{\substack{=m_{f(i)} \\ \text{(by (400))}}} \right) = \sum_{i=1}^{r} [m_i] \left( m_{f(i)} \right) = \sum_{i \in [r]} [m_i] \left( m_{f(i)} \right) \\
&= \sum_{\substack{i \in [r]; \\ f(i)=i}} \underbrace{[m_i] \left( m_{f(i)} \right)}_{\substack{=1_{\mathbf{k}} \\ \text{(since } f(i)=i\text{, and thus} \\ \text{the } m_i\text{-coordinate of } m_{f(i)} \text{ is } 1_{\mathbf{k}}\text{)}}} + \sum_{\substack{i \in [r]; \\ f(i) \neq i}} \underbrace{[m_i] \left( m_{f(i)} \right)}_{\substack{=0 \\ \text{(since } f(i) \neq i\text{, and thus} \\ \text{the } m_i\text{-coordinate of } m_{f(i)} \text{ is 0)}}} \\
&\qquad \left( \begin{array}{c} \text{since each } i \in [r] \text{ satisfies either } f(i) = i \\ \text{or } f(i) \neq i\text{, but not both} \end{array} \right) \\
&= \sum_{\substack{i \in [r]; \\ f(i)=i}} 1_{\mathbf{k}} + \underbrace{\sum_{\substack{i \in [r]; \\ f(i) \neq i}} 0}_{=0} = \sum_{\substack{i \in [r]; \\ f(i)=i}} 1_{\mathbf{k}} \\
&= (\#\text{ of all } i \in [r] \text{ satisfying } f(i) = i) \cdot 1_{\mathbf{k}} \\
&= (\#\text{ of all } i \in I \text{ satisfying } f(i) = i) \cdot 1_{\mathbf{k}} \qquad (\text{since } [r] = I).
\end{aligned}$$

Thus, Lemma 5.21.13 is proved in the particular case when $I = [r]$ for some $r \in \mathbb{N}$.

The general case can easily be reduced to this particular case: We just rename the elements of the set $I$ as $1, 2, \ldots, r$ (where $r = |I|$), and accordingly change the map $f : I \to I$ (by replacing both its inputs and its output values by $1, 2, \ldots, r$). [139] These operations do not change the number

$$(\#\text{ of all } i \in I \text{ satisfying } f(i) = i),$$

since the map $f$ is modified along with the set $I$. [140] Hence, the claim of Lemma 5.21.13 remains unaffected under this operation. But as a result of this operation, we obtain the situation when $I = [r]$, and we have already proved Lemma 5.21.13 in this situation. Hence, Lemma 5.21.13 must hold in the original (fully general) situation as well. □

[139] To make this rigorous: We pick any bijection $\beta : I \to [r]$, and we define a new map $f' : [r] \to [r]$ by $f' := \beta \circ f \circ \beta^{-1}$. Then, we apply Lemma 5.21.13 to the set $[r]$, the basis $\left( m_{\beta^{-1}(i)} \right)_{i \in [r]}$ and the map $f' : [r] \to [r]$ instead of the set $I$, the basis $(m_i)_{i \in I}$ and the map $f : I \to I$. This is allowed, since we have $\phi\left( m_{\beta^{-1}(i)} \right) = m_{\beta^{-1}(f'(i))}$ for all $i \in [r]$ (an easy consequence of (400) and $f' = \beta \circ f \circ \beta^{-1}$).

[140] To make this rigorous: We have

$$(\#\text{ of all } i \in I \text{ satisfying } f(i) = i) = (\#\text{ of all } i \in [r] \text{ satisfying } f'(i) = i),$$

where $f'$ is defined as in the previous footnote.

*Proof of Corollary 5.21.11.* Let $\mathbf{k}$ be the field $\mathbb{Q}$. Then, $n!$ is invertible in $\mathbf{k}$. Hence, Theorem 5.21.2 shows that the family $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \mathrm{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}$.

The antipode $S : \mathbf{k}[S_n] \to \mathbf{k}[S_n]$ is a $\mathbf{k}$-linear map from $\mathcal{A}$ to $\mathcal{A}$. We shall compute its trace $\operatorname{Tr} S$ in two different ways. Both times, we will apply Lemma 5.21.13 to $M = \mathcal{A}$ and $\phi = S$, but we will use two different bases $(m_i)_{i \in I}$ and two different maps $f$.

Here are our two computations of $\operatorname{Tr} S$:

*Claim 1:* We have

$$\operatorname{Tr} S = (\#\text{ of involutions } w \in S_n).$$

*Proof of Claim 1.* Define a map $f : S_n \to S_n$ by setting

$$f(w) = w^{-1} \qquad \text{for all } w \in S_n.$$

The $\mathbf{k}$-module $\mathcal{A} = \mathbf{k}[S_n]$ is free with basis $(w)_{w \in S_n}$. The map $f : S_n \to S_n$ and the $\mathbf{k}$-linear map $S : \mathcal{A} \to \mathcal{A}$ satisfy

$$S(w) = f(w) \qquad \text{for all } w \in S_n \tag{401}$$

(since both $S(w)$ and $f(w)$ are defined to be $w^{-1}$). In other words,

$$S(i) = f(i) \qquad \text{for all } i \in S_n$$

(here, we have renamed the variable $w$ as $i$ in (401)). Hence, Lemma 5.21.13 (applied to $M = \mathcal{A}$ and $I = S_n$ and $(m_i)_{i \in I} = (w)_{w \in S_n}$ and $\phi = S$) yields

$$\begin{aligned}
\operatorname{Tr} S &= (\#\text{ of all } i \in S_n \text{ satisfying } f(i) = i) \cdot \underbrace{1_{\mathbf{k}}}_{\substack{=1_{\mathbb{Q}} \\ (\text{since } \mathbf{k}=\mathbb{Q})}} \\
&= (\#\text{ of all } i \in S_n \text{ satisfying } f(i) = i) \\
&= \left( \#\text{ of all } w \in S_n \text{ satisfying } \underbrace{f(w)}_{\substack{=w^{-1} \\ (\text{by the definition of } f)}} = w \right) \\
&\qquad (\text{here, we have renamed the index } i \text{ as } w) \\
&= \left( \#\text{ of all } w \in S_n \text{ satisfying } w^{-1} = w \right) \\
&= (\#\text{ of involutions } w \in S_n)
\end{aligned}$$

(since the permutations $w \in S_n$ satisfying $w^{-1} = w$ are precisely the involutions $w \in S_n$). This proves Claim 1. $\square$

*Claim 2:* We have

$$\operatorname{Tr} S = \sum_{\lambda \text{ is a partition of } n} f^\lambda.$$

*Proof of Claim 2.* For each partition $\lambda$ of $n$, we let $\operatorname{SYT}(\lambda)$ denote the set of all standard $n$-tableaux of shape $Y(\lambda)$. Then, for each partition $\lambda$ of $n$, we have

$$|\operatorname{SYT}(\lambda)| = f^\lambda$$

(as we saw in the proof of Proposition 5.21.10).

Let

$$I = \bigsqcup_{\lambda \text{ is a partition of } n} (\operatorname{SYT}(\lambda) \times \operatorname{SYT}(\lambda)).$$

In other words, let $I$ be the set of all pairs $(U, V)$, where $\lambda$ is a partition of $n$ and where $U, V \in \operatorname{SYT}(\lambda)$.

If $(U, V) \in I$ is any pair, then the tableaux $U$ and $V$ have the same shape (since $(U, V) \in I$ means that $U, V \in \operatorname{SYT}(\lambda)$ for some partition $\lambda$ of $n$, and therefore both tableaux $U$ and $V$ have shape $Y(\lambda)$), and thus the permutation $w_{U,V}$ is well-defined. Thus, for each $(U, V) \in I$, we can define a vector

$$m_{(U,V)} := \mathbf{F}_U w_{U,V} \mathbf{E}_V \in \mathcal{A}.$$

Thus, we have defined a vector $m_i \in \mathcal{A}$ for each $i \in I$ (since each $i \in I$ can be written as a pair $(U, V)$). We therefore have defined a family $(m_i)_{i \in I}$ of vectors in $\mathcal{A}$. This family

$$\begin{aligned}
(m_i)_{i \in I} &= \left(m_{(U,V)}\right)_{(U,V) \in I} && \left(\begin{array}{c}\text{here, we have renamed the} \\ \text{index } i \text{ as } (U,V)\end{array}\right) \\
&= \left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{(U,V) \in I} && \left(\begin{array}{c}\text{since } m_{(U,V)} = \mathbf{F}_U w_{U,V} \mathbf{E}_V \\ \text{for each } (U,V) \in I\end{array}\right) \\
&= \left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \operatorname{SYT}(\lambda)} \\
& \qquad \left(\begin{array}{c}\text{since our definition of } I \text{ allows us to rewrite} \\ \text{the subscript ``} (U,V) \in I\text{''} \\ \text{as ``} \lambda \text{ is a partition of } n, \text{ and } U, V \in \operatorname{SYT}(\lambda)\text{''}\end{array}\right)
\end{aligned}$$

is a basis of the $\mathbf{k}$-module $\mathcal{A}$ (since we saw above that the family $\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right)_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \operatorname{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}$).

Define the map $f : I \to I$ by setting

$$f(U, V) = (V, U) \qquad \text{for all } (U, V) \in I.$$

(This is well-defined, since $U, V \in \operatorname{SYT}(\lambda)$ clearly entails $V, U \in \operatorname{SYT}(\lambda)$.) Then, for each $(U, V) \in I$, we have

$$\begin{aligned}
S\left(m_{(U,V)}\right) &= S\left(\mathbf{F}_U w_{U,V} \mathbf{E}_V\right) \qquad \left(\text{since } m_{(U,V)} = \mathbf{F}_U w_{U,V} \mathbf{E}_V\right) \\
&= \mathbf{F}_V w_{V,U} \mathbf{E}_U \qquad \left(\begin{array}{c}\text{by Proposition 5.21.4, since } U, V \in \operatorname{SYT}(\lambda) \\ \text{for some partition } \lambda \text{ of } n\end{array}\right) \\
&= m_{(V,U)} \qquad \left(\begin{array}{c}\text{since the definition of } m_{(V,U)} \\ \text{yields } m_{(V,U)} = \mathbf{F}_V w_{V,U} \mathbf{E}_U\end{array}\right) \\
&= m_{f(U,V)} \qquad (\text{since } (V,U) = f(U,V) \text{ (because } f(U,V) = (V,U))).
\end{aligned}$$

Renaming the variable $(U, V)$ as $i$ in this sentence, we obtain the following: For each $i \in I$, we have $S(m_i) = m_{f(i)}$.

Hence, we can apply Lemma 5.21.13 to $M = \mathcal{A}$ and $\phi = S$. Thus, we obtain

$$\begin{aligned}
\operatorname{Tr} S &= (\#\text{ of all } i \in I \text{ satisfying } f(i) = i) \cdot \underbrace{1_{\mathbf{k}}}_{\substack{=1_{\mathbb{Q}} \\ (\text{since } \mathbf{k}=\mathbb{Q})}} \\
&= (\#\text{ of all } i \in I \text{ satisfying } f(i) = i) \\
&= \left(\#\text{ of all } (U,V) \in I \text{ satisfying } \underbrace{f(U,V)}_{\substack{=(V,U) \\ (\text{by the definition of } f)}} = (U,V)\right) \\
&\qquad \left(\begin{array}{c}\text{here, we have renamed the index } i \text{ as } (U,V), \\ \text{since each } i \in I \text{ is a pair of the form } (U,V)\end{array}\right) \\
&= (\#\text{ of all } (U,V) \in I \text{ satisfying } (V,U) = (U,V)) \\
&= (\#\text{ of all } (U,V) \in I \text{ satisfying } U = V) \\
&\qquad (\text{since the condition "} (V,U) = (U,V) \text{" is equivalent to "} U = V") \\
&= \left(\#\text{ of all } (U,V) \in \bigsqcup_{\lambda \text{ is a partition of } n} (\operatorname{SYT}(\lambda) \times \operatorname{SYT}(\lambda)) \text{ satisfying } U = V\right) \\
&\qquad \left(\text{since } I = \bigsqcup_{\lambda \text{ is a partition of } n} (\operatorname{SYT}(\lambda) \times \operatorname{SYT}(\lambda))\right) \\
&= \sum_{\lambda \text{ is a partition of } n} \underbrace{(\#\text{ of all } (U,V) \in \operatorname{SYT}(\lambda) \times \operatorname{SYT}(\lambda) \text{ satisfying } U = V)}_{\substack{=(\#\text{ of all } U \in \operatorname{SYT}(\lambda)) \\ (\text{since a pair } (U,V) \text{ satisfying } U=V \\ \text{is uniquely determined by its first entry } U)}} \\
&= \sum_{\lambda \text{ is a partition of } n} \underbrace{(\#\text{ of all } U \in \operatorname{SYT}(\lambda))}_{=|\operatorname{SYT}(\lambda)|=f^\lambda} = \sum_{\lambda \text{ is a partition of } n} f^\lambda.
\end{aligned}$$

This proves Claim 2. □

Now, Claim 2 yields

$$\sum_{\lambda \text{ is a partition of } n} f^\lambda = \operatorname{Tr} S = (\#\text{ of involutions } w \in S_n)$$

(by Claim 1). Thus, Corollary 5.21.11 is proved. □

Corollary 5.21.11 is not as fundamental as Corollary 5.12.18, but still has uses in the combinatorics of permutations. It has combinatorial proofs as well – see, e.g., [Leeuwe11, Proposition 1.3.2 and the paragraph after its proof], [MenRem15, Exercise 5.7], [Fulton97, §4.3, Exercise 6 and the paragraph above it], [Aigner07, Theorem 8.26 b.], [Sagan01, Exercise 3.12.7 (c)], [Prasad15, Exercise 3.2.12], [Knuth98, §5.1.4, Corollary B], [Stanle23, Corollary 7.13.9], [Beissi87].

The following counting exercise gives some equivalent expressions for the right hand side of Corollary 5.21.11:

**Exercise 5.21.2.** **(a)** $\boxed{1}$ Prove that

$$(\#\text{ of involutions } w \in S_n) = \sum_{k=0}^{n} \binom{n}{2k} \cdot \underbrace{1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2k-1)}_{\substack{\text{This is the product of all} \\ \text{odd integers from 1 to } 2k-1}}$$
$$= \sum_{k=0}^{n} \frac{\binom{n}{2k}\binom{2k}{k} k!}{2^k}.$$

**(b)** $\boxed{1}$ Let $t(m) := (\#\text{ of involutions } w \in S_m)$ for each $m \in \mathbb{N}$. Prove that $t(n) = t(n-1) + (n-1)\, t(n-2)$ if $n \geq 2$.

The numbers $t(n) = (\#\text{ of involutions } w \in S_n)$ appear in the OEIS as Sequence A000085. See Remark 3.15.2 for a table of the first few.

## 5.22. The quotient presentation of a Specht module

### 5.22.1. Definitions and theorems

We have already seen several avatars of a Specht module $\mathcal{S}^D$: the original definition (Definition 5.4.1), the left ideal avatar (Theorem 5.5.13 **(b)**), the Specht–Vandermonde avatar (Theorem 5.6.1 **(b)**) and the letterplace avatar (Theorem 5.6.5 **(b)**). All these avatars have one thing in common: They present $\mathcal{S}^D$ as a submodule of some larger module (the Young module or the symmetric group algebra $\mathbf{k}[S_n]$ itself or a polynomial ring). What we have not seen so far is a presentation of $\mathcal{S}^D$ as a **quotient** module. Ideally, we would like to write $\mathcal{S}^D$ as

a quotient of a free module (over $\mathbf{k}$ or, better, over $\mathcal{A}$) by an explicitly generated submodule; this would be a presentation of $\mathcal{S}^D$ by generators and relations.

This is what we shall do now. In part, this is very easy: The Specht module $\mathcal{S}^D$ is already defined by a set of generators (viz., the polytabloids $\mathbf{e}_T$). Can we describe the relations between these generators? I don't know how to do this in general (the Garnir relations from Theorem 5.8.17 always hold, but they are not always sufficient to derive all relations between the $\mathbf{e}_T$), but at least it can be done when $D$ is a skew Young diagram.

In this section, we will see how this is done. We begin with some generalities that make sense for any diagram $D$ of size $|D| = n$.

Let $\operatorname{Tab}(D)$ be the set of all $n$-tableaux of shape $D$. As we said, the Specht module $\mathcal{S}^D$ is spanned (as a $\mathbf{k}$-module) by the polytabloids $\mathbf{e}_T$ with $T \in \operatorname{Tab}(D)$. There are $n!$ of these polytabloids (since there are $n!$ many $n$-tableaux $T \in \operatorname{Tab}(D)$). Thus, we can write $\mathcal{S}^D$ as a quotient of a free $\mathbf{k}$-module of rank $n!$, which has a basis $(e_T)_{T \in \operatorname{Tab}(D)}$ (do not mistake the $e_T$ for the $\mathbf{e}_T$). We shall call this free $\mathbf{k}$-module $\mathcal{T}^D$. The best way to define it (which gives it a left $\mathbf{k}[S_n]$-module structure "for free") is as a permutation module:

**Definition 5.22.1.** Let $D$ be any diagram with $|D| = n$.

As we know from Definition 5.3.9, the group $S_n$ acts on the set $\{n\text{-tableaux of shape } D\}$. Let us denote this set by $\operatorname{Tab}(D)$. Thus, $\operatorname{Tab}(D) = \{n\text{-tableaux of shape } D\}$ is a left $S_n$-set.

Let $\mathcal{T}^D$ be the permutation module[141] $\mathbf{k}^{(\operatorname{Tab}(D))}$ corresponding to this left $S_n$-set $\operatorname{Tab}(D)$. Thus, $\mathcal{T}^D$ is the free $\mathbf{k}$-module $\mathbf{k}^{(\operatorname{Tab}(D))}$ with its standard basis $(e_T)_{T \in \operatorname{Tab}(D)}$ (do not mistake these basis vectors $e_T$ for the polytabloids $\mathbf{e}_T$), and it is furthermore a left $\mathbf{k}[S_n]$-module whose action satisfies

$$we_T = e_{wT} \qquad \text{for each } w \in S_n \text{ and } T \in \operatorname{Tab}(D) \tag{402}$$

(by a general property of permutation modules – see (108)). Hence, it is a representation of $S_n$.

The basis vectors $e_T$ of $\mathcal{T}^D$ are not the polytabloids $\mathbf{e}_T$ from the Specht module $\mathcal{S}^D$, even though the rule for the $S_n$-action is the same for both types of vectors: The polytabloids $\mathbf{e}_T \in \mathcal{S}^D$ also satisfy

$$w\mathbf{e}_T = \mathbf{e}_{wT} \qquad \text{for each } w \in S_n \text{ and } T \in \operatorname{Tab}(D) \tag{403}$$

(since Lemma 5.4.6 **(a)** (applied to $u = w$) yields $\mathbf{e}_{wT} = w\mathbf{e}_T$). The equalities (402) and (403) look alike, and this will be used later on.

As a left $\mathbf{k}[S_n]$-module (i.e., as an $S_n$-representation), $\mathcal{T}^D$ is not particularly interesting:

[141] See Example 4.3.9 for the definition of a permutation module.

**Proposition 5.22.2.** Let $D$ be any diagram with $|D| = n$. Consider the left regular representation $\mathbf{k}[S_n]$ of $S_n$. Then,

$$\mathcal{T}^D \cong \mathbf{k}[S_n] \qquad \text{as } S_n\text{-representations.}$$

Specifically, we can obtain an isomorphism as follows: Fix any $n$-tableau $P \in \operatorname{Tab}(D)$. Let $\phi_P : \mathbf{k}[S_n] \to \mathcal{T}^D$ be the $\mathbf{k}$-linear map that sends each standard basis vector $w \in S_n$ to $e_{wP}$. Then, $\phi_P$ is an isomorphism of $S_n$-representations.

*Proof.* First, we shall show that $\phi_P$ is an isomorphism of $\mathbf{k}$-modules. Indeed, $P$ is an $n$-tableau of shape $D$ (since $P \in \operatorname{Tab}(D)$). Hence, the map

$$\begin{aligned} f : S_n &\to \{n\text{-tableaux of shape } D\}, \\ w &\mapsto w \rightharpoonup P \end{aligned}$$

is a bijection (by Proposition 5.3.13 **(d)**, applied to $T = P$). In other words, the map

$$\begin{aligned} f : S_n &\to \operatorname{Tab}(D), \\ w &\mapsto wP \end{aligned}$$

is a bijection (since $\{n\text{-tableaux of shape } D\} = \operatorname{Tab}(D)$ and since $w \rightharpoonup P = wP$ for each $w \in S_n$). Therefore, the family $(e_{wP})_{w \in S_n}$ is a reindexing of the family $(e_T)_{T \in \operatorname{Tab}(D)}$. Hence, the family $(e_{wP})_{w \in S_n}$ is a basis of the $\mathbf{k}$-module $\mathcal{T}^D$ (since the family $(e_T)_{T \in \operatorname{Tab}(D)}$ is a basis of $\mathcal{T}^D$).

Now, the $\mathbf{k}$-linear map $\phi_P : \mathbf{k}[S_n] \to \mathcal{T}^D$ sends the basis $(w)_{w \in S_n}$ of $\mathbf{k}[S_n]$ to the basis $(e_{wP})_{w \in S_n}$ of $\mathcal{T}^D$ (since it sends each $w$ to $e_{wP}$ by its definition). Hence, this map is an isomorphism of $\mathbf{k}$-modules (because any $\mathbf{k}$-linear map that sends a basis of its domain to a basis of its target must be an isomorphism), and thus is invertible.

Moreover, it is easy to show that the map $\phi_P$ is $S_n$-equivariant (this essentially follows from (402)[142]). Hence, this map $\phi_P$ is a morphism of $S_n$-representations

[142] In more detail: We must show that $\phi_P$ is $S_n$-equivariant. In other words, we must prove that all $g \in S_n$ and all $x \in \mathbf{k}[S_n]$ satisfy $\phi_P(g \rightharpoonup x) = g \rightharpoonup \phi_P(x)$. So let us prove this. Let $g \in S_n$ and $x \in \mathbf{k}[S_n]$. We must prove the equality $\phi_P(g \rightharpoonup x) = g \rightharpoonup \phi_P(x)$. Since both sides of this equality are $\mathbf{k}$-linear in $x$, we can WLOG assume that $x$ is an element of the standard basis $(w)_{w \in S_n}$ of $\mathbf{k}[S_n]$. Assume this. Thus, $x = w$ for some $w \in S_n$. Consider this $w$.

From $x = w$, we obtain $\phi_P(x) = \phi_P(w) = e_{wP}$ (by the definition of $\phi_P$) and $g \rightharpoonup x = g \rightharpoonup w = gw$. From $g \rightharpoonup x = gw$, we obtain $\phi_P(g \rightharpoonup x) = \phi_P(gw) = e_{gwP}$ (by the definition of $\phi_P$, since $gw \in S_n$). But (402) (applied to $g$ and $wP$ instead of $w$ and $T$) yields $ge_{wP} = e_{gwP}$. Comparing this with $\phi_P(g \rightharpoonup x) = e_{gwP}$, we obtain $\phi_P(g \rightharpoonup x) = ge_{wP} = g \rightharpoonup e_{wP}$. In view of $\phi_P(x) = e_{wP}$, we can rewrite this as $\phi_P(g \rightharpoonup x) = g \rightharpoonup \phi_P(x)$. This completes our proof that $\phi_P$ is $S_n$-equivariant.

(since it is $\mathbf{k}$-linear). In other words, the map $\phi_P$ is a morphism of left $\mathbf{k}[S_n]$-modules (by Proposition 4.3.18). Since $\phi_P$ is invertible, we thus conclude that $\phi_P$ is a left $\mathbf{k}[S_n]$-module isomorphism. In other words, $\phi_P$ is an isomorphism of $S_n$-representations. Hence, $\mathcal{T}^D \cong \mathbf{k}[S_n]$ as $S_n$-representations. Proposition 5.22.2 is thus proven.

(Alternatively, we could have argued that the above map $f : S_n \to \operatorname{Tab}(D)$ is an isomorphism of left $S_n$-sets, and thus (by the functoriality of the "permutation module" functor) induces an isomorphism between the corresponding permutation modules, which are $\mathbf{k}[S_n]$ and $\mathcal{T}^D$; this latter isomorphism is precisely our map $\phi_P$.) □

Another easy observation is the existence of a surjective morphism of $S_n$-representations from $\mathcal{T}^D$ to the Specht module $\mathcal{S}^D$:

**Proposition 5.22.3.** Let $D$ be any diagram with $|D| = n$. Let $\pi_D : \mathcal{T}^D \to \mathcal{S}^D$ be the $\mathbf{k}$-linear map that sends each standard basis vector $e_T$ of $\mathcal{T}^D$ to the polytabloid $\mathbf{e}_T \in \mathcal{S}^D$. Then, $\pi_D$ is a surjective morphism of $S_n$-representations.

*Proof.* Recall that $\mathcal{S}^D$ is the span of the polytabloids $\mathbf{e}_T$ for all $T \in \operatorname{Tab}(D)$ (by the definition of $\mathcal{S}^D$). This yields that the map $\pi_D$ is surjective[143]. It remains to prove that $\pi_D$ is a morphism of $S_n$-representations. It clearly suffices to show that $\pi_D$ is $S_n$-equivariant (since $\pi_D$ is $\mathbf{k}$-linear). But this is an easy consequence of comparing (402) with (403).[144] Thus, the proof of Proposition 5.22.3 is complete. □

[143] In more detail: The $\mathbf{k}$-module $\mathcal{T}^D$ is free with basis $(e_T)_{T \in \operatorname{Tab}(D)}$. Hence, $\mathcal{T}^D = \operatorname{span}_{\mathbf{k}}\{e_T \mid T \in \operatorname{Tab}(D)\}$. Thus,

$$\begin{aligned}
\pi_D\left(\mathcal{T}^D\right) &= \pi_D\left(\operatorname{span}_{\mathbf{k}}\{e_T \mid T \in \operatorname{Tab}(D)\}\right) \\
&= \operatorname{span}_{\mathbf{k}}\left\{ \underbrace{\pi_D(e_T)}_{\substack{=\mathbf{e}_T \\ \text{(by the definition of } \pi_D\text{)}}} \mid T \in \operatorname{Tab}(D) \right\} \\
&\qquad \left( \begin{array}{c} \text{since the map } \pi_D \text{ is } \mathbf{k}\text{-linear, and thus} \\ \text{respects } \mathbf{k}\text{-linear combinations} \end{array} \right) \\
&= \operatorname{span}_{\mathbf{k}}\{\mathbf{e}_T \mid T \in \operatorname{Tab}(D)\} \\
&= \operatorname{span}_{\mathbf{k}}\{\mathbf{e}_T \mid T \text{ is an } n\text{-tableau of shape } D\} \\
&\qquad (\text{since } \operatorname{Tab}(D) \text{ is the set of all } n\text{-tableaux of shape } D) \\
&= \mathcal{S}^D \qquad \left(\text{by the definition of } \mathcal{S}^D\right).
\end{aligned}$$

In other words, the map $\pi_D$ is surjective.

[144] In more detail: We must prove that $\pi_D$ is $S_n$-equivariant. In other words, we must prove that all $g \in S_n$ and all $x \in \mathcal{T}^D$ satisfy $\pi_D(g \rightharpoonup x) = g \rightharpoonup \pi_D(x)$. So let us prove this. Let $g \in S_n$ and $x \in \mathcal{T}^D$. We must prove the equality $\pi_D(g \rightharpoonup x) = g \rightharpoonup \pi_D(x)$. Since both sides of this equality are $\mathbf{k}$-linear in $x$, we can WLOG assume that $x$ is an element of the standard

The morphism $\pi_D : \mathcal{T}^D \to \mathcal{S}^D$ from Proposition 5.22.3 is surjective (by said proposition), and thus we have

$$\mathcal{S}^D = \operatorname{Im}(\pi_D) = \pi_D\left(\mathcal{T}^D\right) \cong \mathcal{T}^D / \operatorname{Ker}(\pi_D) \tag{404}$$

(by the first isomorphism theorem), and this is an isomorphism of $S_n$-representations (since $\pi_D$ is a morphism of $S_n$-representations). This can be viewed as a new avatar of $\mathcal{S}^D$. Moreover, since $\mathcal{T}^D$ is a free $\mathbf{k}$-module, this yields a presentation of the $\mathbf{k}$-module $\mathcal{S}^D$ by generators and relations (the generators being the projections of the standard basis vectors $e_T \in \mathcal{T}^D$ onto $\mathcal{T}^D / \operatorname{Ker}(\pi_D)$, while the relations are given by the elements of $\operatorname{Ker}(\pi_D)$), provided that we can explicitly express $\operatorname{Ker}(\pi_D)$ as a span of some concrete vectors (called "relators", since they determine the relations in this presentation). We do not know how to do this in general, but we shall do this in the case when $D$ is a skew Young diagram. The relevant vectors spanning $\operatorname{Ker}(\pi_D)$ will be the so-called *column relators* and *Garnir relators*. These relators can be defined for arbitrary $D$ and always belong to $\operatorname{Ker}(\pi_D)$, but they might not always suffice to span $\operatorname{Ker}(\pi_D)$ when $D$ is not a skew Young diagram.

We begin with their definition:

**Definition 5.22.4.** Let $D$ be any diagram with $|D| = n$. We define the following two types of elements of $\mathcal{T}^D$:

**(a)** A *column relator* (in $\mathcal{T}^D$) shall mean an element of the form

$$e_T - (-1)^u e_{uT} \in \mathcal{T}^D,$$

where $T \in \operatorname{Tab}(D)$ and $u \in \mathcal{C}(T)$.

**(b)** A *Garnir pentuple* (of shape $D$) shall mean a 5-tuple $(T, j, k, Z, L)$ whose entries $T, j, k, Z, L$ satisfy the assumptions of Theorem 5.8.17. Concretely, this means the following:

- The first entry $T$ of the 5-tuple is an $n$-tableau of shape $D$ (that is, we have $T \in \operatorname{Tab}(D)$).
- The next two entries $j$ and $k$ are two distinct integers.
- The fourth entry $Z$ is a subset of $\operatorname{Col}(j, T) \cup \operatorname{Col}(k, T)$ such that $|Z| > |D\langle j\rangle \cup D\langle k\rangle|$. (See Definition 5.8.1 for the meaning of $D\langle j\rangle$.)

basis $(e_T)_{T \in \operatorname{Tab}(D)}$ of $\mathcal{T}^D$. Assume this. Thus, $x = e_T$ for some $T \in \operatorname{Tab}(D)$. Consider this $T$.

From $x = e_T$, we obtain $\pi_D(x) = \pi_D(e_T) = \mathbf{e}_T$ (by the definition of $\pi_D$) and thus $g \rightharpoonup \pi_D(x) = g \rightharpoonup \mathbf{e}_T = g\mathbf{e}_T = \mathbf{e}_{gT}$ (by (403), applied to $w = g$). On the other hand, from $x = e_T$, we obtain $g \rightharpoonup x = g \rightharpoonup e_T = ge_T = e_{gT}$ (by (402), applied to $w = g$) and therefore $\pi_D(g \rightharpoonup x) = \pi_D\left(e_{gT}\right) = \mathbf{e}_{gT}$ (by the definition of $\pi_D$). Comparing this with $g \rightharpoonup \pi_D(x) = \mathbf{e}_{gT}$, we find $\pi_D(g \rightharpoonup x) = g \rightharpoonup \pi_D(x)$. This completes our proof that $\pi_D$ is $S_n$-equivariant.

- The final entry $L$ is a left transversal of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$. (See Proposition 3.7.4 for the definition of $S_{n,Z}$.)

**(c)** A *Garnir relator* (in $\mathcal{T}^D$) shall mean an element of the form

$$\sum_{x \in L} (-1)^x e_{xT} \in \mathcal{T}^D,$$

where $(T, j, k, Z, L)$ is a Garnir pentuple (of shape $D$).

**Example 5.22.5.** Let $n = 8$ and $D = Y(3,3,2)$ for this example. Then:

**(a)** There are 8! distinct $n$-tableaux $T$ of shape $D$, and for each such $T$ there are $|\mathcal{C}(T)| = 6 \cdot 6 \cdot 2 = 72$ possible permutations $u \in \mathcal{C}(T)$. Thus, there are altogether $8! \cdot 72 = 2\,903\,040$ many column relators in $\mathcal{T}^D$. Here are two of them (using the notation from Convention 5.2.4):

$$e_{123\backslash\backslash 456\backslash\backslash 78} + e_{423\backslash\backslash 156\backslash\backslash 78}$$

(corresponding to $T = 123\backslash\backslash 456\backslash\backslash 78$ and $u = t_{1,4} \in \mathcal{C}(T)$) and

$$e_{123\backslash\backslash 456\backslash\backslash 78} - e_{453\backslash\backslash 126\backslash\backslash 78}$$

(corresponding to $T = 123\backslash\backslash 456\backslash\backslash 78$ and $u = t_{1,4}t_{2,5} \in \mathcal{C}(T)$). Note that each column relator $e_T - (-1)^u e_{uT}$ with $u = \mathrm{id}$ is simply 0; the other column relators also have many linear dependencies (usually).

**(b)** There are 8! distinct $n$-tableaux $T$ of shape $D$, and for each such $T$ there are many possible Garnir pentuples $(T, j, k, Z, L)$ that start with $T$. One such pentuple has been given in Example 5.8.18:

$$T = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 6 & 3 & 7 \\ \hline 8 & 5 \\ \cline{1-2}\end{array} \quad \text{and} \quad j = 1 \quad \text{and} \quad k = 2 \quad \text{and}$$

$$Z = \{2,3,6,8\} \quad \text{and} \quad L = \{\mathrm{id},\ t_{2,6},\ t_{2,8},\ t_{3,6},\ t_{3,8},\ t_{2,6}t_{3,8}\}.$$

Hence,

$$\sum_{x \in L} (-1)^x e_{xT} = e_{\mathrm{id}\,T} - e_{t_{2,6}T} - e_{t_{2,8}T} - e_{t_{3,6}T} - e_{t_{3,8}T} + e_{t_{2,6}t_{3,8}T}$$

is a Garnir relator in $\mathcal{T}^D$.

As we already mentioned, both column and Garnir relators belong to $\mathrm{Ker}(\pi_D)$:

**Proposition 5.22.6.** Let $D$ be any diagram with $|D| = n$. Consider the morphism $\pi_D : \mathcal{T}^D \to \mathcal{S}^D$ from Proposition 5.22.3. Then:

**(a)** Each column relator in $\mathcal{T}^D$ belongs to $\operatorname{Ker}(\pi_D)$.

**(b)** Each Garnir relator in $\mathcal{T}^D$ belongs to $\operatorname{Ker}(\pi_D)$.

*Proof.* **(a)** This is a consequence of Lemma 5.4.6 **(c)** (rewritten in the equivalent form $\mathbf{e}_T - (-1)^u \mathbf{e}_{uT} = 0$). In more detail:

Let $\mathfrak{c}$ be a column relator in $\mathcal{T}^D$. We must prove that $\mathfrak{c} \in \operatorname{Ker}(\pi_D)$.

We know that $\mathfrak{c}$ is a column relator. In other words, $\mathfrak{c} = e_T - (-1)^u e_{uT}$ for some $T \in \operatorname{Tab}(D)$ and some $u \in \mathcal{C}(T)$ (by the definition of a column relator). Consider these $T$ and $u$. Then, Lemma 5.4.6 **(c)** yields $\mathbf{e}_{uT} = (-1)^u \mathbf{e}_T$. But applying the map $\pi_D$ to the equality $\mathfrak{c} = e_T - (-1)^u e_{uT}$, we obtain

$$\begin{aligned}
\pi_D(\mathfrak{c}) &= \pi_D\left(e_T - (-1)^u e_{uT}\right) \\
&= \underbrace{\pi_D(e_T)}_{\substack{=\mathbf{e}_T \\ \text{(by the definition of } \pi_D\text{)}}} - (-1)^u \underbrace{\pi_D(e_{uT})}_{\substack{=\mathbf{e}_{uT} \\ \text{(by the definition of } \pi_D\text{)}}} \qquad \text{(since } \pi_D \text{ is } \mathbf{k}\text{-linear)} \\
&= \mathbf{e}_T - (-1)^u \underbrace{\mathbf{e}_{uT}}_{=(-1)^u \mathbf{e}_T} = \mathbf{e}_T - \underbrace{(-1)^u (-1)^u}_{\substack{=(-1)^{2u}=1 \\ \text{(since } 2u \text{ is even)}}} \mathbf{e}_T \\
&= \mathbf{e}_T - \mathbf{e}_T = 0.
\end{aligned}$$

In other words, $\mathfrak{c} \in \operatorname{Ker}(\pi_D)$. Thus, Proposition 5.22.6 **(a)** is proved.

**(b)** This is a consequence of the Garnir relation (206). In more detail:

Let $\mathfrak{g}$ be a Garnir relator in $\mathcal{T}^D$. We must prove that $\mathfrak{g} \in \operatorname{Ker}(\pi_D)$.

We know that $\mathfrak{g}$ is a Garnir relator. Thus, $\mathfrak{g} = \sum_{x \in L} (-1)^x e_{xT}$ for some Garnir pentuple $(T, j, k, Z, L)$. Consider this $(T, j, k, Z, L)$. We know that $(T, j, k, Z, L)$ is a Garnir pentuple, so that its entries $T, j, k, Z, L$ satisfy the assumptions of Theorem 5.8.17. Hence, (206) says that

$$\sum_{x \in L} (-1)^x \mathbf{e}_{xT} = 0.$$

However, applying the map $\pi_D$ to the equality $\mathfrak{g} = \sum_{x \in L} (-1)^x e_{xT}$, we obtain

$$\begin{aligned}
\pi_D(\mathfrak{g}) &= \pi_D\left(\sum_{x \in L} (-1)^x e_{xT}\right) \\
&= \sum_{x \in L} (-1)^x \underbrace{\pi_D(e_{xT})}_{\substack{=\mathbf{e}_{xT} \\ \text{(by the definition of } \pi_D\text{)}}} \qquad \text{(since } \pi_D \text{ is } \mathbf{k}\text{-linear)} \\
&= \sum_{x \in L} (-1)^x \mathbf{e}_{xT} = 0.
\end{aligned}$$

That is, $\mathfrak{g} \in \operatorname{Ker}(\pi_D)$. Thus, Proposition 5.22.6 **(b)** is proved. □

The main result of this section is the following:

**Theorem 5.22.7.** Let $D$ be a skew Young diagram with $|D| = n$. Consider the morphism $\pi_D : \mathcal{T}^D \to \mathcal{S}^D$ from Proposition 5.22.3. Then,

$$\begin{aligned}&\operatorname{Ker}(\pi_D)\\ &= \operatorname{span}_{\mathbf{k}}\left(\left\{\text{all column relators in } \mathcal{T}^D\right\} \cup \left\{\text{all Garnir relators in } \mathcal{T}^D\right\}\right).\end{aligned}$$

We will prove this theorem in Subsection 5.22.2. But first, we shall generalize it by replacing the two (finite but rather large) sets $\left\{\text{all column relators in } \mathcal{T}^D\right\}$ and $\left\{\text{all Garnir relators in } \mathcal{T}^D\right\}$ by certain subsets $\mathfrak{C}$ and $\mathfrak{G}$ that can be much smaller (we shall not uniquely specify these subsets but only impose conditions on them). The following types of subsets do the trick:

**Definition 5.22.8.** Let $D$ be a skew Young diagram with $|D| = n$.

**(a)** A set $\mathfrak{C}$ of column relators (in $\mathcal{T}^D$) is said to be *triangulating* if it has the following property: For any $n$-tableau $T \in \operatorname{Tab}(D)$ that is not column-standard, there exists a $u \in \mathcal{C}(T)$ such that the $n$-tableau $uT$ is column-standard and such that $e_T - (-1)^u e_{uT} \in \mathfrak{C}$.

**(b)** A Garnir pentuple $(T, j, k, Z, L)$ is said to be *triangulating* if each $x \in L \setminus \mathcal{C}(T)$ satisfies $\widetilde{xT} > \widetilde{T}$ in the column last letter order. (For the definition of the column last letter order, see Definition 5.9.6 and the discussion that follows.)

**(c)** A set $\mathfrak{G}$ of Garnir relators (in $\mathcal{T}^D$) is said to be *triangulating* if it has the following property: For any column-standard $n$-tableau $T \in \operatorname{Tab}(D)$ that is not standard, there exists a triangulating Garnir pentuple $(T, j, k, Z, L)$ with first entry $T$ such that $\sum_{x \in L} (-1)^x e_{xT} \in \mathfrak{G}$.

Triangulating sets of column relators and of Garnir relators can be much smaller than the sets of all such relators: It is enough to have just one column relator for each tableau $T \in \operatorname{Tab}(D)$ that is not column-standard, and only one Garnir relator for each column-standard tableau $T$ that is not standard. Altogether, fewer than $n!$ relators are needed.

**Example 5.22.9.** For this example, let $n = 5$ and $D = Y(2,2,1)$. Then, there are 10 column-standard $n$-tableaux of shape $D$. Five of these are standard; the other five are (using the notation from Convention 5.2.4)

$$\begin{array}{|c|c|}\hline 1&2\\\hline 4&3\\\hline 5\\\cline{1-1}\end{array}\,,\quad \begin{array}{|c|c|}\hline 2&1\\\hline 3&5\\\hline 4\\\cline{1-1}\end{array}\,,\quad \begin{array}{|c|c|}\hline 2&1\\\hline 3&4\\\hline 5\\\cline{1-1}\end{array}\,,\quad \begin{array}{|c|c|}\hline 2&1\\\hline 4&3\\\hline 5\\\cline{1-1}\end{array}\,,\quad \begin{array}{|c|c|}\hline 3&1\\\hline 4&2\\\hline 5\\\cline{1-1}\end{array}$$

(where we used yellow background color to mark the rows that violate standardness). For a set $\mathfrak{G}$ of Garnir relators (in $\mathcal{T}^D$) to be triangulating, it must contain – for each of these five non-standard tableaux $T$ – a Garnir relator $\sum_{x\in L}(-1)^x e_{xT}$ coming from a triangulating Garnir pentuple $(T, j, k, Z, L)$ with first entry $T$. One possible choice of such Garnir relators is given as follows (note that $j$ and $k$ are always 1 and 2 here, and the set $Z$ is always the set of the entries highlighted in green, whereas the set $L$ is easily reconstructed from the resulting tableaux $xT$):

| $T$ | Garnir relator $\sum_{x\in L}(-1)^x e_{xT}$ |
|---|---|
| $\begin{array}{|c|c|}\hline 1&2\\\hline 4&3\\\hline 5\\\cline{1-1}\end{array}$ | $e_{12\backslash\backslash43\backslash\backslash5} - e_{12\backslash\backslash34\backslash\backslash5} - e_{14\backslash\backslash23\backslash\backslash5} - e_{15\backslash\backslash43\backslash\backslash2} - e_{12\backslash\backslash45\backslash\backslash3} + e_{14\backslash\backslash25\backslash\backslash3}$ |
| $\begin{array}{|c|c|}\hline 2&1\\\hline 3&5\\\hline 4\\\cline{1-1}\end{array}$ | $e_{21\backslash\backslash35\backslash\backslash4} - e_{12\backslash\backslash35\backslash\backslash4} - e_{23\backslash\backslash15\backslash\backslash4} - e_{24\backslash\backslash35\backslash\backslash1}$ |
| $\begin{array}{|c|c|}\hline 2&1\\\hline 3&4\\\hline 5\\\cline{1-1}\end{array}$ | $e_{21\backslash\backslash34\backslash\backslash5} - e_{12\backslash\backslash34\backslash\backslash5} - e_{23\backslash\backslash14\backslash\backslash5} - e_{25\backslash\backslash34\backslash\backslash1}$ |
| $\begin{array}{|c|c|}\hline 2&1\\\hline 4&3\\\hline 5\\\cline{1-1}\end{array}$ | $e_{21\backslash\backslash43\backslash\backslash5} - e_{12\backslash\backslash43\backslash\backslash5} - e_{24\backslash\backslash13\backslash\backslash5} - e_{25\backslash\backslash43\backslash\backslash1}$ |
| $\begin{array}{|c|c|}\hline 3&1\\\hline 4&2\\\hline 5\\\cline{1-1}\end{array}$ | $e_{31\backslash\backslash42\backslash\backslash5} - e_{13\backslash\backslash42\backslash\backslash5} - e_{34\backslash\backslash12\backslash\backslash5} - e_{35\backslash\backslash42\backslash\backslash1}$ |

(check that these Garnir pentuples are all triangulating!). Thus, if we take $\mathfrak{G}$ to be the set of these five Garnir relators, then $\mathfrak{G}$ is triangulating. There are

several alternative choices (in particular, for the last tableau $T$ in our list, we could have picked $Z = \{1, 2, 3, 4, 5\}$, thus obtaining a Garnir relator with 10 addends).

We now claim that triangulating sets are sufficient to span $\operatorname{Ker}(\pi_D)$:

**Theorem 5.22.10.** Let $D$ be a skew Young diagram with $|D| = n$. Consider the morphism $\pi_D : \mathcal{T}^D \to \mathcal{S}^D$ from Proposition 5.22.3. Let $\mathfrak{C}$ be a triangulating set of column relators (in $\mathcal{T}^D$). Let $\mathfrak{G}$ be a triangulating set of Garnir relators (in $\mathcal{T}^D$). Then,

$$\operatorname{Ker}(\pi_D) = \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G}).$$

This will be proved in Subsection 5.22.2.

As we said, Theorem 5.22.10 and Theorem 5.22.7 can be combined with (404) to yield explicit presentations of a skew-shaped Specht module $\mathcal{S}^D$ in terms of generators and relations. Let us make this explicit (for Theorem 5.22.10):

**Corollary 5.22.11.** Let $D$ be a skew Young diagram with $|D| = n$. Consider the morphism $\pi_D : \mathcal{T}^D \to \mathcal{S}^D$ from Proposition 5.22.3. Let $\mathfrak{C}$ be a triangulating set of column relators (in $\mathcal{T}^D$). Let $\mathfrak{G}$ be a triangulating set of Garnir relators (in $\mathcal{T}^D$). Then, $\operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})$ is an $S_n$-subrepresentation of $\mathcal{T}^D$, and we have

$$\mathcal{S}^D \cong \mathcal{T}^D / \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G}) \qquad \text{as } S_n\text{-representations.}$$

*Proof.* The map $\pi_D : \mathcal{T}^D \to \mathcal{S}^D$ from Proposition 5.22.3 is a morphism of $S_n$-representations, i.e., a morphism of left $\mathbf{k}[S_n]$-modules. Thus, its kernel $\operatorname{Ker}(\pi_D)$ is a left $\mathbf{k}[S_n]$-submodule of $\mathcal{T}^D$, that is, an $S_n$-subrepresentation of $\mathcal{T}^D$. But Theorem 5.22.10 yields

$$\operatorname{Ker}(\pi_D) = \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G}).$$

Thus, $\operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})$ is an $S_n$-subrepresentation of $\mathcal{T}^D$ (since $\operatorname{Ker}(\pi_D)$ is an $S_n$-subrepresentation of $\mathcal{T}^D$). Now, (404) becomes

$$\mathcal{S}^D \cong \mathcal{T}^D / \underbrace{\operatorname{Ker}(\pi_D)}_{=\operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})} = \mathcal{T}^D / \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G}) \qquad \text{as } S_n\text{-representations.}$$

Thus, Corollary 5.22.11 is proved. □

### 5.22.2. Proofs

We shall now provide the missing proofs of Theorem 5.22.10 and Theorem 5.22.7 that we promised in Subsection 5.22.1.

To prove Theorem 5.22.10, we shall study the quotient module $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})$, showing in particular that it has a basis $(\overline{e_T})_{T \in \operatorname{SYT}(D)}$ analogous to the standard basis $(\mathbf{e}_T)_{T \in \operatorname{SYT}(D)}$ of the Specht module $\mathcal{S}^D$. More precisely, we will only need to show that the family $(\overline{e_T})_{T \in \operatorname{SYT}(D)}$ spans the quotient $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})$ (its linear independence, while true, won't be needed). This will give us an injective morphism from $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})$ to $\mathcal{S}^D$ compatible with the projection $\pi_D : \mathcal{T}^D \to \mathcal{S}^D$, and this will easily yield Theorem 5.22.10. This proof will be somewhat long, but almost all of it will be a calque of our above proof of the standard basis theorem (Theorem 5.9.1), and the rest is easy enough that it could be given as a $\boxed{2}$-point exercise. Yet, the quest for completeness is forcing me to present this proof in full, for what it is worth.

**Convention 5.22.12.** Throughout Subsection 5.22.2, we shall use the following notations:

- We fix a skew Young diagram $D$ with $|D| = n$.
- We let $\mathfrak{C}$ be a triangulating set of column relators (in $\mathcal{T}^D$), and we let $\mathfrak{G}$ be a triangulating set of Garnir relators (in $\mathcal{T}^D$).
- We let $\operatorname{SYT}(D)$ denote the set of all standard tableaux of shape $D$. In other words, $\operatorname{SYT}(D)$ is the set of all standard $n$-tableaux of shape $D$ (by Proposition 5.2.6). Thus, $\operatorname{SYT}(D) \subseteq \operatorname{Tab}(D)$.
- We consider the quotient $\mathbf{k}$-module $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})$. For each $v \in \mathcal{T}^D$, we let $\overline{v}$ be the projection of $v$ on this quotient module. In particular, the projections $\overline{e_T}$ of the standard basis vectors $e_T \in \mathcal{T}^D$ will be called the *quasi-polytabloids* (to stress their analogy with the polytabloids $\mathbf{e}_T \in \mathcal{S}^D$). The quasi-polytabloids $\overline{e_T}$ with $T \in \operatorname{SYT}(D)$ will be called the *standard quasi-polytabloids*.

The first step on our way to Theorem 5.22.10 is a restatement of Proposition 5.9.5:

**Lemma 5.22.13.** Each $n$-column-tabloid of shape $D$ contains a unique column-standard $n$-tableau[145].

*Proof.* Proposition 5.9.5 shows that there is a bijection

$$\begin{aligned} &\text{from } \{\text{column-standard } n\text{-tableaux of shape } D\} \\ &\text{to } \{n\text{-column-tabloids of shape } D\} \end{aligned}$$

that sends each column-standard $n$-tableau $T$ to its column-tabloid $\widetilde{T}$. Hence, each $n$-column-tabloid of shape $D$ has a unique preimage under this bijection (because this is

[145] The word "contains" here should be understood literally, as a containment of an element in a set (since an $n$-column-tabloid is an equivalence class of $n$-tableaux of shape $D$ with respect to column-equivalence, and thus is a set of $n$-tableaux).

what being a bijection means). In other words, for each $n$-column-tabloid $C$ of shape $D$, there exists a unique column-standard $n$-tableau $T$ of shape $D$ such that $\widetilde{T} = C$. In other words, for each $n$-column-tabloid $C$ of shape $D$, there exists a unique column-standard $n$-tableau $T$ of shape $D$ such that $T \in C$ (because the statement "$\widetilde{T} = C$" is equivalent to "$T \in C$"). In other words, each $n$-column-tabloid $C$ of shape $D$ contains a unique column-standard $n$-tableau $T$. This proves Lemma 5.22.13. □

Next, we show a lemma saying that the quasi-polytabloids $\overline{e_T} \in \mathcal{T}^D / \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})$ corresponding to column-equivalent $n$-tableaux $T$ are equal up to sign:

**Lemma 5.22.14.** Let $P$ and $Q$ be two column-equivalent $n$-tableaux of shape $D$. Then, $\overline{e_P} = \pm\overline{e_Q}$ in $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})$.

*Proof.* Recall that an $n$-column-tabloid is an equivalence class of $n$-tableaux of shape $D$ with respect to column-equivalence. Hence, the $n$-tableaux $P$ and $Q$ belong to the same $n$-column-tabloid (since they are column-equivalent). In other words, $\widetilde{P} = \widetilde{Q}$.

However, Lemma 5.22.13 says that each $n$-column-tabloid of shape $D$ contains a unique column-standard $n$-tableau. Thus, in particular, the $n$-column-tabloid $\widetilde{P} = \widetilde{Q}$ contains a unique column-standard $n$-tableau. Let us denote this unique column-standard $n$-tableau by $K$. Hence, $K \in \widetilde{P} = \widetilde{Q}$, so that $\widetilde{K} = \widetilde{Q} = \widetilde{P}$. Furthermore, $K$ is defined as the **unique** column-standard $n$-tableau contained in the $n$-column-tabloid $\widetilde{P}$. Hence, if $W \in \widetilde{P}$ is any column-standard $n$-tableau, then

$$W = K. \tag{405}$$

We will now prove that

$$\overline{e_P} = \pm\overline{e_K}. \tag{406}$$

*Proof of (406):* We must show that $\overline{e_P} = \pm\overline{e_K}$.

The $n$-tableau $P$ is clearly contained in the $n$-column-tabloid $\widetilde{P}$. If this $n$-tableau $P$ is furthermore column-standard, then (405) (applied to $W = P$) shows that $P = K$ and therefore $\overline{e_P} = \overline{e_K} = \pm\overline{e_K}$, so that (406) is proved. Hence, for the rest of this proof, we WLOG assume that $P$ is **not** column-standard.

Now recall that the set $\mathfrak{C}$ of column relators is triangulating. Hence, (by Definition 5.22.8 **(a)**) we conclude that for any $n$-tableau $T \in \operatorname{Tab}(D)$ that is not column-standard, there exists a $u \in \mathcal{C}(T)$ such that the $n$-tableau $uT$ is column-standard and such that $e_T - (-1)^u e_{uT} \in \mathfrak{C}$. Applying this to $T = P$, we conclude that there exists a $u \in \mathcal{C}(P)$ such that the $n$-tableau $uP$ is column-standard and such that $e_P - (-1)^u e_{uP} \in \mathfrak{C}$ (since $P \in \operatorname{Tab}(D)$ is not column-standard). Consider this $u$. Note that

$$e_P - (-1)^u e_{uP} \in \mathfrak{C} \subseteq \mathfrak{C} \cup \mathfrak{G} \subseteq \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G}).$$

In other words, $\overline{e_P - (-1)^u e_{uP}} = 0$ in $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})$. Hence, $\overline{e_P} - (-1)^u \overline{e_{uP}} = \overline{e_P - (-1)^u e_{uP}} = 0$, so that

$$\overline{e_P} = \underbrace{(-1)^u}_{=\pm 1} \overline{e_{uP}} = \pm\overline{e_{uP}}. \tag{407}$$

But $u \in \mathcal{C}(P)$. Hence, the $n$-tableau $uP = u \rightharpoonup P$ is column-equivalent to $P$ (by Proposition 5.3.12 **(b)**, applied to $w = u$ and $T = P$), and thus belongs to the same

$n$-column-tabloid as $P$. In other words, $uP \in \widetilde{P}$. Since $uP$ is furthermore column-standard, we thus conclude that $uP = K$ (by (405), applied to $W = uP$). In view of this, we can rewrite (407) as $\overline{e_P} = \pm\overline{e_K}$. Thus, (406) is proved. □

We have now shown that $\overline{e_P} = \pm\overline{e_K}$. Since $P$ and $Q$ play analogous roles in our situation, we can likewise show that $\overline{e_Q} = \pm\overline{e_K}$ (the $\pm$ sign here is not necessarily the same as in $\overline{e_P} = \pm\overline{e_K}$). Hence, $\overline{e_K} = \pm\overline{e_Q}$. Combining our results, we obtain

$$\overline{e_P} = \pm \underbrace{\overline{e_K}}_{=\pm\overline{e_Q}} = \pm\left(\pm\overline{e_Q}\right) = \pm\overline{e_Q}.$$

Thus, Lemma 5.22.14 is proved. □

Next we show an analogue of Lemma 5.9.13:

**Lemma 5.22.15.** Let $T$ be an $n$-tableau of shape $D$ that is column-standard but not standard. Then, $\overline{e_T}$ can be written as a $\mathbb{Z}$-linear combination of quasi-polytabloids $\overline{e_P}$, where the subscripts $P$ are $n$-tableaux of shape $D$ satisfying $\widetilde{P} > \widetilde{T}$ (with respect to the column last letter order).

*Proof.* The set $\mathfrak{G}$ of Garnir relators is triangulating. Therefore, by Definition 5.22.8 **(c)**, we conclude that there exists a triangulating Garnir pentuple $(T, j, k, Z, L)$ with first entry $T$ such that $\sum_{x\in L} (-1)^x e_{xT} \in \mathfrak{G}$ (since $T \in \operatorname{Tab}(D)$ is column-standard but not standard). Consider this Garnir pentuple $(T, j, k, Z, L)$. By the definition of a Garnir pentuple, it thus satisfies the following:

- The entries $j$ and $k$ are two distinct integers.
- The entry $Z$ is a subset of $\operatorname{Col}(j, T) \cup \operatorname{Col}(k, T)$ such that $|Z| > |D\langle j\rangle \cup D\langle k\rangle|$.
- The entry $L$ is a left transversal of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$.

Moreover, the Garnir pentuple $(T, j, k, Z, L)$ is triangulating. In other words, each $x \in L \setminus \mathcal{C}(T)$ satisfies

$$\widetilde{xT} > \widetilde{T} \qquad \text{in the column last letter order} \tag{408}$$

(by Definition 5.22.8 **(b)**).

However,

$$\sum_{x\in L} (-1)^x e_{xT} \in \mathfrak{G} \subseteq \mathfrak{C} \cup \mathfrak{G} \subseteq \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G}).$$

In other words, $\overline{\sum_{x\in L} (-1)^x e_{xT}} = 0$ in $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}}(\mathfrak{C} \cup \mathfrak{G})$. Hence,

$$0 = \overline{\sum_{x\in L} (-1)^x e_{xT}} = \sum_{x\in L} (-1)^x \overline{e_{xT}}. \tag{409}$$

Now, recall that $L$ is a left transversal of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$. Thus, each left coset of the subgroup $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$ contains exactly one element of $L$ (by Definition 5.8.9

**(a)**). Hence, in particular, the trivial left coset $1\left(S_{n,Z}\cap\mathcal{C}\left(T\right)\right)$ of $S_{n,Z}\cap\mathcal{C}\left(T\right)$ in $S_{n,Z}$ contains exactly one element of $L$. Let $v\in L$ be this element. Thus, $v\in1\left(S_{n,Z}\cap\mathcal{C}\left(T\right)\right)=S_{n,Z}\cap\mathcal{C}\left(T\right)\subseteq\mathcal{C}\left(T\right)$. Therefore, the $n$-tableau $vT=v\rightharpoonup T$ is column-equivalent to $T$ (by Proposition 5.3.12 **(b)**, applied to $w=v$). Thus, Lemma 5.22.14 (applied to $P=vT$ and $Q=T$) yields

$$\overline{e_{vT}}=\pm\overline{e_T}\qquad\text{in }\mathcal{T}^D/\operatorname{span}_{\mathbf{k}}\left(\mathfrak{C}\cup\mathfrak{G}\right).$$

Now, recall that $v$ is the only element of $L$ that is contained in $1\left(S_{n,Z}\cap\mathcal{C}\left(T\right)\right)$ (by the definition of $v$). Thus, each $x\in L\setminus\{v\}$ satisfies $x\notin1\left(S_{n,Z}\cap\mathcal{C}\left(T\right)\right)$ and therefore

$$\widetilde{xT}>\widetilde{T}\qquad\text{in the column last letter order}\tag{410}$$

(because combining $x\in L\setminus\{v\}\subseteq L\subseteq S_{n,Z}$ with $x\notin1\left(S_{n,Z}\cap\mathcal{C}\left(T\right)\right)=S_{n,Z}\cap\mathcal{C}\left(T\right)$, we obtain $x\in S_{n,Z}\setminus\left(S_{n,Z}\cap\mathcal{C}\left(T\right)\right)=S_{n,Z}\setminus\mathcal{C}\left(T\right)$ and therefore $x\notin\mathcal{C}\left(T\right)$, so that $x\in L\setminus\mathcal{C}\left(T\right)$ (since $x\in L$ and $x\notin\mathcal{C}\left(T\right)$), and therefore (408) shows that $\widetilde{xT}>\widetilde{T}$ in the column last letter order). But (409) becomes

$$\begin{aligned}0&=\sum_{x\in L}\left(-1\right)^x\overline{e_{xT}}=\underbrace{\left(-1\right)^v}_{=\pm1}\underbrace{\overline{e_{vT}}}_{=\pm\overline{e_T}}+\sum_{x\in L\setminus\{v\}}\left(-1\right)^x\overline{e_{xT}}\\&\qquad\qquad\left(\begin{array}{c}\text{here, we have split off the addend for }x=v\\\text{from the sum, since }v\in L\end{array}\right)\\&=\underbrace{\left(\pm1\right)\left(\pm\overline{e_T}\right)}_{=\pm\overline{e_T}}+\sum_{x\in L\setminus\{v\}}\left(-1\right)^x\overline{e_{xT}}=\pm\overline{e_T}+\sum_{x\in L\setminus\{v\}}\left(-1\right)^x\overline{e_{xT}}.\end{aligned}$$

Moving the $\pm\overline{e_T}$ addend onto the left hand side here, we obtain

$$\overline{e_T}=\pm\sum_{x\in L\setminus\{v\}}\left(-1\right)^x\overline{e_{xT}}.\tag{411}$$

But (410) tells us that each $x\in L\setminus\{v\}$ satisfies $\widetilde{xT}>\widetilde{T}$ in the column last letter order. Hence, each of the vectors $\overline{e_{xT}}$ on the right hand side of (411) has the form $\overline{e_P}$ for some $n$-tableau $P$ of shape $D$ satisfying $\widetilde{P}>\widetilde{T}$ in the column last letter order (namely, for $P=xT$). Thus, the equality (411) shows that $\overline{e_T}$ can be written as a $\mathbb{Z}$-linear combination of quasi-polytabloids $\overline{e_P}$, where the subscripts $P$ are $n$-tableaux of shape $D$ satisfying $\widetilde{P}>\widetilde{T}$ (with respect to the column last letter order). This proves Lemma 5.22.15. □

The following lemma is an analogue of Proposition 5.9.14:

**Lemma 5.22.16.** Let $T$ be an $n$-tableau of shape $D$. Then, $\overline{e_T}$ is a $\mathbb{Z}$-linear combination of standard quasi-polytabloids $\overline{e_Q}$, where $Q$ ranges over all standard tableaux of shape $D$.

*Proof.* Forget that we fixed $T$. We must show that each quasi-polytabloid $\overline{e_T}$ is a $\mathbb{Z}$-linear combination of standard quasi-polytabloids.

In this proof, we will not consider any diagrams other than $D$. Thus, the word "$n$-tableau" will always mean "$n$-tableau of shape $D$". Likewise for the word "$n$-column-tabloid".

Define a $\mathbb{Z}$-submodule $\overline{\mathcal{Q}}^D$ of $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G})$ by

$$\overline{\mathcal{Q}}^D := \operatorname{span}_{\mathbb{Z}} \left\{ \overline{e_Q} \mid Q \text{ is a standard } n\text{-tableau} \right\}.$$

Thus, $\overline{\mathcal{Q}}^D$ is the $\mathbb{Z}$-linear span of all standard quasi-polytabloids. We shall show that

$$\overline{e_T} \in \overline{\mathcal{Q}}^D \qquad \text{for all } n\text{-tableaux } T. \tag{412}$$

Indeed, let us define the *depth* $\operatorname{dep} \widetilde{T}$ of an $n$-column-tabloid $\widetilde{T}$ to be the # of all $n$-column-tabloids $\widetilde{P}$ that satisfy $\widetilde{P} > \widetilde{T}$ (with respect to the column last letter order). This is clearly a well-defined nonnegative integer (since there are only finitely many $n$-column-tabloids $\widetilde{P}$). Moreover, if two $n$-column-tabloids $\widetilde{S}$ and $\widetilde{T}$ satisfy $\widetilde{S} > \widetilde{T}$ (with respect to the column last letter order), then

$$\operatorname{dep} \widetilde{S} < \operatorname{dep} \widetilde{T} \tag{413}$$

(indeed, this was already proved during the proof of Proposition 5.9.14).

Now, we shall prove (412) by strong induction on $\operatorname{dep} \widetilde{T}$:

*Induction step:* Fix an $n$-tableau $U$, and assume (as the induction hypothesis) that (412) is already proved for all $n$-tableaux $T$ satisfying $\operatorname{dep} \widetilde{T} < \operatorname{dep} \widetilde{U}$. Our goal is now to prove (412) for $T = U$. In other words, our goal is to prove that $\overline{e_U} \in \overline{\mathcal{Q}}^D$.

The tableau $U$ may or may not be column-standard. However, it is easy to find a column-standard $n$-tableau $S$ that is column-equivalent to $U$ (just let $S$ be the result of sorting the entries in each column of $U$ into increasing order). Consider this $S$. Then, $\widetilde{S} = \widetilde{U}$ (since $S$ is column-equivalent to $U$) and $\overline{e_U} = \pm \overline{e_S}$ (by Lemma 5.22.14, applied to $P = U$ and $Q = S$).

If the $n$-tableau $S$ is standard, then we have $\overline{e_S} \in \overline{\mathcal{Q}}^D$ (by the definition of $\overline{\mathcal{Q}}^D$, since $\overline{e_S}$ is a standard quasi-polytabloid), and thus $\overline{e_U} = \pm \underbrace{\overline{e_S}}_{\in \overline{\mathcal{Q}}^D} \in \overline{\mathcal{Q}}^D$ (since $\overline{\mathcal{Q}}^D$ is a $\mathbb{Z}$-module), which immediately completes our induction step (since our goal is to prove that $\overline{e_U} \in \overline{\mathcal{Q}}^D$).

Hence, for the rest of the induction step, we WLOG assume that $S$ is not standard.

Thus, Lemma 5.22.15 (applied to $S$ instead of $T$) shows that $\overline{e_S}$ can be written as a $\mathbb{Z}$-linear combination of quasi-polytabloids $\overline{e_P}$, where the subscripts $P$ are $n$-tableaux of shape $D$ satisfying $\widetilde{P} > \widetilde{S}$ (with respect to the column last letter order). In other words,

$$\overline{e_S} \in \operatorname{span}_{\mathbb{Z}} \left\{ \overline{e_P} \mid P \text{ is an } n\text{-tableau satisfying } \widetilde{P} > \widetilde{S} \right\}. \tag{414}$$

Now, let $P$ be an $n$-tableau satisfying $\widetilde{P} > \widetilde{S}$. Then, $\widetilde{P} > \widetilde{S} = \widetilde{U}$ and therefore $\operatorname{dep} \widetilde{P} < \operatorname{dep} \widetilde{U}$ (by (413), applied to $P$ and $U$ instead of $S$ and $T$). But our induction hypothesis says that (412) is already proved for all $n$-tableaux $T$ satisfying $\operatorname{dep} \widetilde{T} < \operatorname{dep} \widetilde{U}$. Applying this to $T = P$, we conclude that (412) is already proved for $T = P$ (since $\operatorname{dep} \widetilde{P} < \operatorname{dep} \widetilde{U}$). Hence, $\overline{e_P} \in \overline{\mathcal{Q}}^D$.

Forget that we fixed $P$. We thus have shown that $\overline{e_P} \in \overline{\mathcal{Q}}^D$ whenever $P$ is an $n$-tableau satisfying $\widetilde{P} > \widetilde{S}$. In other words,

$$\left\{ \overline{e_P} \mid P \text{ is an } n\text{-tableau satisfying } \widetilde{P} > \widetilde{S} \right\} \subseteq \overline{\mathcal{Q}}^D.$$

Now, (414) becomes

$$\overline{e_S} \in \operatorname{span}_{\mathbb{Z}} \underbrace{\left\{ \overline{e_P} \mid P \text{ is an } n\text{-tableau satisfying } \widetilde{P} > \widetilde{S} \right\}}_{\subseteq \overline{\mathcal{Q}}^D} \subseteq \operatorname{span}_{\mathbb{Z}} \left( \overline{\mathcal{Q}}^D \right) = \overline{\mathcal{Q}}^D$$

(since $\overline{\mathcal{Q}}^D$ is a $\mathbb{Z}$-module). Hence, $\overline{e_U} = \pm \underbrace{\overline{e_S}}_{\in \overline{\mathcal{Q}}^D} \in \overline{\mathcal{Q}}^D$ (since $\overline{\mathcal{Q}}^D$ is a $\mathbb{Z}$-module). But this is precisely what we intended to prove. Thus, we have proved (412) for $T = U$. This completes the induction step. Thus, (412) is proved for all $T$.

In other words, we have proved that each quasi-polytabloid $\overline{e_T}$ belongs to $\overline{\mathcal{Q}}^D$. In other words, each quasi-polytabloid $\overline{e_T}$ belongs to the $\mathbb{Z}$-linear span of all standard quasi-polytabloids (since $\overline{\mathcal{Q}}^D$ is the $\mathbb{Z}$-linear span of all standard quasi-polytabloids). In other words, each quasi-polytabloid $\overline{e_T}$ is a $\mathbb{Z}$-linear combination of standard quasi-polytabloids. This proves Lemma 5.22.16. □

Lemma 5.22.16 easily yields the following:

**Lemma 5.22.17.** The $\mathbf{k}$-module $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G})$ is spanned by the family $\left(\overline{e_T}\right)_{T \in \operatorname{SYT}(D)}$ of all standard quasi-polytabloids.

*Proof.* The $\mathbf{k}$-module $\mathcal{T}^D$ is spanned by the family $(e_T)_{T \in \operatorname{Tab}(D)}$ (since it is free with basis $(e_T)_{T \in \operatorname{Tab}(D)}$). Hence, its quotient module $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G})$ is spanned by the family $\left(\overline{e_T}\right)_{T \in \operatorname{Tab}(D)}$. In other words,

$$\mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G}) = \operatorname{span}_{\mathbf{k}} \left\{ \overline{e_T} \mid T \in \operatorname{Tab}(D) \right\}. \tag{415}$$

On the other hand, each $T \in \operatorname{Tab}(D)$ satisfies

$$\begin{aligned}
\overline{e_T} &\in \operatorname{span}_{\mathbb{Z}} \left\{ \overline{e_Q} \mid Q \text{ is a standard tableau of shape } D \right\} \\
&\qquad \left( \begin{array}{c} \text{since Lemma 5.22.16 shows that } \overline{e_T} \text{ is a } \mathbb{Z}\text{-linear combination of} \\ \text{the } \overline{e_Q}\text{, where } Q \text{ ranges over all standard tableaux of shape } D \end{array} \right) \\
&= \operatorname{span}_{\mathbb{Z}} \left\{ \overline{e_Q} \mid Q \in \operatorname{SYT}(D) \right\} \\
&\qquad (\text{since } \operatorname{SYT}(D) \text{ is the set of all standard tableaux of shape } D) \\
&\subseteq \operatorname{span}_{\mathbf{k}} \left\{ \overline{e_Q} \mid Q \in \operatorname{SYT}(D) \right\} \\
&\qquad (\text{since } \operatorname{span}_{\mathbb{Z}} A \subseteq \operatorname{span}_{\mathbf{k}} A \text{ for any subset } A \text{ of any } \mathbf{k}\text{-module}).
\end{aligned}$$

In other words,

$$\left\{ \overline{e_T} \mid T \in \operatorname{Tab}(D) \right\} \subseteq \operatorname{span}_{\mathbf{k}} \left\{ \overline{e_Q} \mid Q \in \operatorname{SYT}(D) \right\}.$$

Thus,

$$\begin{aligned}
&\operatorname{span}_{\mathbf{k}} \underbrace{\left\{ \overline{e_T} \mid T \in \operatorname{Tab}(D) \right\}}_{\subseteq \operatorname{span}_{\mathbf{k}} \left\{ \overline{e_Q} \mid Q \in \operatorname{SYT}(D) \right\}} \\
&\subseteq \operatorname{span}_{\mathbf{k}} \left( \operatorname{span}_{\mathbf{k}} \left\{ \overline{e_Q} \mid Q \in \operatorname{SYT}(D) \right\} \right) \\
&\subseteq \operatorname{span}_{\mathbf{k}} \left\{ \overline{e_Q} \mid Q \in \operatorname{SYT}(D) \right\} \\
&\qquad (\text{since } \operatorname{span}_{\mathbf{k}} \left\{ \overline{e_Q} \mid Q \in \operatorname{SYT}(D) \right\} \text{ is a } \mathbf{k}\text{-module}) \\
&= \operatorname{span}_{\mathbf{k}} \left\{ \overline{e_T} \mid T \in \operatorname{SYT}(D) \right\}.
\end{aligned}$$

Using (415), we can rewrite this as

$$\mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G}) \subseteq \operatorname{span}_{\mathbf{k}} \{\overline{e_T} \mid T \in \operatorname{SYT}(D)\}.$$

Combining this with the obvious inclusion

$$\operatorname{span}_{\mathbf{k}} \{\overline{e_T} \mid T \in \operatorname{SYT}(D)\} \subseteq \mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G}),$$

we conclude that

$$\mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G}) = \operatorname{span}_{\mathbf{k}} \{\overline{e_T} \mid T \in \operatorname{SYT}(D)\}.$$

In other words, the $\mathbf{k}$-module $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G})$ is spanned by the family $(\overline{e_T})_{T \in \operatorname{SYT}(D)}$. This proves Lemma 5.22.17. □

We can now prove Theorem 5.22.10 at last:

*Proof of Theorem 5.22.10.* Consider the morphism $\pi_D : \mathcal{T}^D \to \mathcal{S}^D$ from Proposition 5.22.3.

Each element of $\mathfrak{C}$ is a column relator (since $\mathfrak{C}$ is a set of column relators) and thus belongs to $\operatorname{Ker}(\pi_D)$ (by Proposition 5.22.6 **(a)**). In other words, $\mathfrak{C} \subseteq \operatorname{Ker}(\pi_D)$.

Each element of $\mathfrak{G}$ is a Garnir relator (since $\mathfrak{G}$ is a set of Garnir relators) and thus belongs to $\operatorname{Ker}(\pi_D)$ (by Proposition 5.22.6 **(b)**). In other words, $\mathfrak{G} \subseteq \operatorname{Ker}(\pi_D)$.

Combining $\mathfrak{C} \subseteq \operatorname{Ker}(\pi_D)$ and $\mathfrak{G} \subseteq \operatorname{Ker}(\pi_D)$, we obtain $\mathfrak{C} \cup \mathfrak{G} \subseteq \operatorname{Ker}(\pi_D)$. Hence,

$$\operatorname{span}_{\mathbf{k}} \underbrace{(\mathfrak{C} \cup \mathfrak{G})}_{\subseteq \operatorname{Ker}(\pi_D)} \subseteq \operatorname{span}_{\mathbf{k}} (\operatorname{Ker}(\pi_D)) \subseteq \operatorname{Ker}(\pi_D) \qquad (\text{since } \operatorname{Ker}(\pi_D) \text{ is a } \mathbf{k}\text{-module}).$$

Now, let us show the reverse inclusion. Indeed, let $v \in \operatorname{Ker}(\pi_D)$. We shall prove that $v \in \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G})$. For this purpose, we shall show that $\overline{v} = 0$ in the quotient $\mathbf{k}$-module $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G})$.

Lemma 5.22.17 shows that the $\mathbf{k}$-module $\mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G})$ is spanned by the family $(\overline{e_T})_{T \in \operatorname{SYT}(D)}$ of all standard quasi-polytabloids. Hence, the vector $\overline{v} \in \mathcal{T}^D / \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G})$ is a $\mathbf{k}$-linear combination of these standard quasi-polytabloids. In other words, we can write $\overline{v}$ as

$$\overline{v} = \sum_{T \in \operatorname{SYT}(D)} \alpha_T \overline{e_T} \tag{416}$$

for some scalars $\alpha_T \in \mathbf{k}$. Consider these $\alpha_T$. Hence,

$$\overline{\sum_{T \in \operatorname{SYT}(D)} \alpha_T e_T} = \sum_{T \in \operatorname{SYT}(D)} \alpha_T \overline{e_T} = \overline{v} \qquad (\text{by (416)}).$$

In other words,

$$\sum_{T \in \operatorname{SYT}(D)} \alpha_T e_T - v \in \operatorname{span}_{\mathbf{k}} (\mathfrak{C} \cup \mathfrak{G}) \subseteq \operatorname{Ker}(\pi_D).$$

Hence,

$$\pi_D \left( \sum_{T \in \operatorname{SYT}(D)} \alpha_T e_T - v \right) = 0. \tag{417}$$

But the map $\pi_D$ is $\mathbf{k}$-linear. Hence,

$$\pi_D\left(\sum_{T\in \mathrm{SYT}(D)} \alpha_T e_T - v\right) = \sum_{T\in \mathrm{SYT}(D)} \alpha_T \underbrace{\pi_D\left(e_T\right)}_{\substack{=\mathbf{e}_T \\ \text{(by the definition of } \pi_D\text{)}}} - \underbrace{\pi_D\left(v\right)}_{\substack{=0 \\ \text{(since } v\in \mathrm{Ker}(\pi_D)\text{)}}}$$
$$= \sum_{T\in \mathrm{SYT}(D)} \alpha_T \mathbf{e}_T.$$

Comparing this with (417), we obtain

$$\sum_{T\in \mathrm{SYT}(D)} \alpha_T \mathbf{e}_T = 0. \tag{418}$$

However, Theorem 5.7.4 shows that the standard polytabloids of shape $D$ are $\mathbf{k}$-linearly independent. In other words, the vectors $\mathbf{e}_T$, where $T$ ranges over all standard tableaux of shape $D$, are $\mathbf{k}$-linearly independent (since these vectors are the standard polytabloids of shape $D$). In other words, the vectors $\mathbf{e}_T$ with $T \in \mathrm{SYT}\left(D\right)$ are $\mathbf{k}$-linearly independent (since $\mathrm{SYT}\left(D\right)$ is the set of all standard tableaux of shape $D$). Hence, from (418), we conclude that all the coefficients $\alpha_T$ must be 0. In other words, $\alpha_T = 0$ for each $T \in \mathrm{SYT}\left(D\right)$. Hence, (416) becomes

$$\overline{v} = \sum_{T\in \mathrm{SYT}(D)} \underbrace{\alpha_T}_{=0} \overline{e_T} = 0.$$

In other words, $v \in \mathrm{span}_{\mathbf{k}}\left(\mathfrak{C} \cup \mathfrak{G}\right)$.

Forget that we fixed $v$. We thus have proved that $v \in \mathrm{span}_{\mathbf{k}}\left(\mathfrak{C} \cup \mathfrak{G}\right)$ for each $v \in \mathrm{Ker}\left(\pi_D\right)$. In other words, $\mathrm{Ker}\left(\pi_D\right) \subseteq \mathrm{span}_{\mathbf{k}}\left(\mathfrak{C} \cup \mathfrak{G}\right)$. Combining this with $\mathrm{span}_{\mathbf{k}}\left(\mathfrak{C} \cup \mathfrak{G}\right) \subseteq \mathrm{Ker}\left(\pi_D\right)$, we conclude that $\mathrm{Ker}\left(\pi_D\right) = \mathrm{span}_{\mathbf{k}}\left(\mathfrak{C} \cup \mathfrak{G}\right)$. This proves Theorem 5.22.10. □

*Proof of Theorem 5.22.7.* Define the two sets

$$\mathfrak{C} = \left\{\text{all column relators in } \mathcal{T}^D\right\} \qquad \text{and} \tag{419}$$
$$\mathfrak{G} = \left\{\text{all Garnir relators in } \mathcal{T}^D\right\}. \tag{420}$$

Then, we note the following:

- The set $\mathfrak{C}$ is a triangulating set of column relators.

  [*Proof:* Clearly, $\mathfrak{C}$ is a set of column relators. We must prove that this set $\mathfrak{C}$ is triangulating. In other words, we must prove that for any $n$-tableau $T \in \mathrm{Tab}\left(D\right)$ that is not column-standard, there exists a $u \in \mathcal{C}\left(T\right)$ such that the $n$-tableau $uT$ is column-standard and such that $e_T - \left(-1\right)^u e_{uT} \in \mathfrak{C}$ (because this is what "triangulating" means, by Definition 5.22.8 **(a)**).

  So let us do this. Let $T \in \mathrm{Tab}\left(D\right)$ be an $n$-tableau that is not column-standard. We must show that there exists a $u \in \mathcal{C}\left(T\right)$ such that the $n$-tableau $uT$ is column-standard and such that $e_T - \left(-1\right)^u e_{uT} \in \mathfrak{C}$.

  The $n$-column-tabloid $\widetilde{T}$ contains a unique column-standard $n$-tableau (since Lemma 5.22.13 says that each $n$-column-tabloid of shape $D$ contains a unique

column-standard $n$-tableau). Let $P$ be the latter column-standard $n$-tableau. Then, $P$ is contained in the $n$-column-tabloid $\widetilde{T}$. In other words, $P$ is column-equivalent to $T$. In other words, the $n$-tableaux $T$ and $P$ are column-equivalent.

Now, Proposition 5.3.14 **(b)** (applied to $S = P$) shows that the $n$-tableaux $T$ and $P$ are column-equivalent if and only if there exists some $w \in \mathcal{C}(T)$ such that $P = w \rightharpoonup T$. Hence, there exists some $w \in \mathcal{C}(T)$ such that $P = w \rightharpoonup T$ (since the $n$-tableaux $T$ and $P$ are column-equivalent). Consider this $w$.

Now, the definition of a column relator (Definition 5.22.4 **(a)**, applied to $u = w$) shows that $e_T - (-1)^w e_{wT}$ is a column relator (since $T \in \operatorname{Tab}(D)$ and $w \in \mathcal{C}(T)$). In other words, $e_T - (-1)^w e_{wT} \in \mathfrak{C}$ (since $\mathfrak{C}$ is the set of all column relators). Moreover, the tableau $wT = w \rightharpoonup T = P$ is column-standard. Hence, there exists a $u \in \mathcal{C}(T)$ such that the $n$-tableau $uT$ is column-standard and such that $e_T - (-1)^u e_{uT} \in \mathfrak{C}$ (namely, $u = w$). This is precisely what we needed to prove. Thus, we have shown that $\mathfrak{C}$ is a triangulating set of column relators.]

- The set $\mathfrak{G}$ is a triangulating set of Garnir relators.

  [*Proof:* Clearly, $\mathfrak{G}$ is a set of Garnir relators. We must prove that this set $\mathfrak{G}$ is triangulating. In other words, we must prove that for any column-standard $n$-tableau $T \in \operatorname{Tab}(D)$ that is not standard, there exists a triangulating Garnir pentuple $(T, j, k, Z, L)$ with first entry $T$ such that $\sum_{x \in L} (-1)^x e_{xT} \in \mathfrak{G}$ (because this is what "triangulating" means, by Definition 5.22.8 **(c)**).

  So let us do this. Let $T \in \operatorname{Tab}(D)$ be a column-standard $n$-tableau that is not standard. We must construct a triangulating Garnir pentuple $(T, j, k, Z, L)$ with first entry $T$ such that $\sum_{x \in L} (-1)^x e_{xT} \in \mathfrak{G}$.

  The tableau $T$ is column-standard but not standard. Hence, $T$ is not row-standard (since otherwise, $T$ would be both column-standard and row-standard, and therefore standard). Therefore, $T$ has two entries in the same row that are in the wrong order (i.e., the entry further left is $\geq$ to the entry further right). In other words, there exist two entries $T(i,j)$ and $T(i,k)$ of $T$ with $j < k$ and $T(i,j) \geq T(i,k)$. Consider these $i$, $j$ and $k$. Define two sets $Z$ and $L$ as in the above proof of Lemma 5.9.13. In that proof, we have shown that

  - the $j$ and $k$ are two distinct integers;
  - the set $Z$ is a subset of $\operatorname{Col}(j,T) \cup \operatorname{Col}(k,T)$ such that $|Z| > |D\langle j\rangle \cup D\langle k\rangle|$;
  - the set $L$ is a left transversal of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$.

  Therefore, $(T, j, k, Z, L)$ is a Garnir pentuple (by Definition 5.22.4 **(b)**). Moreover, each $x \in L \setminus \mathcal{C}(T)$ satisfies $\widetilde{xT} > \widetilde{T}$ in the column last letter order (as we showed during our above proof of Lemma 5.9.13). In other words, the Garnir pentuple $(T, j, k, Z, L)$ is triangulating (by Definition 5.22.8 **(b)**). Finally, $\sum_{x \in L} (-1)^x e_{xT}$ is a Garnir relator (by Definition 5.22.4 **(c)**), and thus belongs to $\mathfrak{G}$ (since $\mathfrak{G}$ is the set of all Garnir relators). Thus, we have constructed a triangulating Garnir pentuple $(T, j, k, Z, L)$ with first entry $T$ such that $\sum_{x \in L} (-1)^x e_{xT} \in \mathfrak{G}$. As we said above, this completes our proof that $\mathfrak{G}$ is a triangulating set of Garnir relators.]

Hence, we can apply Theorem 5.22.10 and conclude that

$$\begin{aligned}\operatorname{Ker}\left(\pi_D\right) &= \operatorname{span}_{\mathbf{k}}\left(\mathfrak{C} \cup \mathfrak{G}\right) \\ &= \operatorname{span}_{\mathbf{k}}\left(\left\{\text{all column relators in } \mathcal{T}^D\right\} \cup \left\{\text{all Garnir relators in } \mathcal{T}^D\right\}\right)\end{aligned}$$

(by (419) and (420)). Hence, Theorem 5.22.7 is proved. □

## 5.23. Linear equations defining the Specht module

The Specht module $\mathcal{S}^D$ corresponding to a diagram $D$ was defined as a span of certain elements of the (usually much larger) Young module $\mathcal{M}^D$ (see Definition 5.4.1 **(b)**). A natural question to ask is whether it can also be described as the solution set of a system of linear equations – i.e., as a kernel of a linear map. In this section, we shall present such a description, at least in the case when $D$ is a skew Young diagram. The linear equations in this description will be called the *coGarnir equations*, due to the fact that we derive them from the Garnir relations. They hold for all vectors in $\mathcal{S}^D$ for any diagram $D$, but they do not determine $\mathcal{S}^D$ in general; however, they do when $D$ is a skew Young diagram.

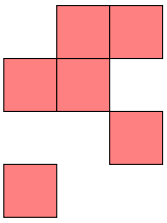
### 5.23.1. The coGarnir equations

To state the coGarnir equations, we need some preparation. First, we give analogues of Definition 5.8.1 and Definition 5.8.2:

**Definition 5.23.1.** Let $D$ be a diagram. Let $i \in \mathbb{Z}$. Then, $D\lfloor i \rfloor$ shall mean the set of all integers $j$ that satisfy $(i,j) \in D$.

For example, if

$$D = \{(1,2),\ (1,3),\ (2,1),\ (2,2),\ (3,3),\ (4,1)\} = $$



,

then $D\lfloor 1 \rfloor = \{2,3\}$ and $D\lfloor 2 \rfloor = \{1,2\}$ and $D\lfloor 3 \rfloor = \{3\}$ and $D\lfloor 4 \rfloor = \{1\}$ and $D\lfloor i \rfloor = \varnothing$ for all $i \notin [4]$. If $D$ is a skew Young diagram $Y\left(\lambda/\mu\right)$, then each set $D\lfloor i \rfloor$ is an integer interval (equalling $\{\mu_i + 1,\ \mu_i + 2,\ \ldots,\ \lambda_i\}$ when $i > 0$ and being empty otherwise), but of course this does not generally hold for bad-shape diagrams.

**Definition 5.23.2.** Let $T$ be an $n$-tableau (of any shape). Let $i \in \mathbb{Z}$. Then, $\operatorname{Row}(i, T)$ shall denote the set of all entries in the $i$-th row of $T$.

The following (somewhat ugly) notation will shorten some of our formulas:

**Definition 5.23.3.** Let $D$ be a diagram with $|D| = n$. Then, $\overline{\operatorname{Tab}}(D)$ shall denote the set $\{n\text{-tabloids of shape } D\}$. As we know from Definition 5.3.10, this set is a left $S_n$-set.

As we recall (see Definition 5.3.15), the Young module $\mathcal{M}^D$ is the permutation module corresponding to this left $S_n$-set $\overline{\operatorname{Tab}}(D)$. Hence, as a $\mathbf{k}$-module, it is the free $\mathbf{k}$-module $\mathbf{k}^{\left(\overline{\operatorname{Tab}}(D)\right)}$ and has a basis $\left(\overline{P}\right)_{\overline{P} \in \overline{\operatorname{Tab}}(D)}$. An element $\mathbf{a}$ of $\mathcal{M}^D$ can thus be written as a family $(a_{\overline{P}})_{\overline{P} \in \overline{\operatorname{Tab}}(D)}$ of scalars $a_{\overline{P}} \in \mathbf{k}$, or, equivalently, as a formal $\mathbf{k}$-linear combination $\sum_{\overline{P} \in \overline{\operatorname{Tab}}(D)} a_{\overline{P}} \overline{P}$ with coefficients $a_{\overline{P}} \in \mathbf{k}$.

We now define the coGarnir equations. As the name suggests, these are somewhat similar (in their construction) to the Garnir relations (Theorem 5.8.17); in particular, they too rely on left transversals of subgroups (Definition 5.8.9). However, they differ in a few other aspects (rows are used instead of columns, and permutations appear without sign). Here is the precise definition:

**Definition 5.23.4.** Let $D$ be a diagram with $|D| = n$. Let $T$ be an $n$-tableau of shape $D$. Let $j$ and $k$ be two distinct integers. Let $Z$ be a subset of $\operatorname{Row}(j, T) \cup \operatorname{Row}(k, T)$ such that $|Z| > |D\lfloor j \rfloor \cup D\lfloor k \rfloor|$. (Here we are using the notations introduced in Definition 5.23.1 and Definition 5.23.2.)

Define $S_{n,Z}$ as in Proposition 3.7.4. This set $S_{n,Z}$ is a subgroup of $S_n$ (by Proposition 3.7.4 **(a)**, applied to $X = Z$). Thus, both $\mathcal{R}(T)$ and $S_{n,Z}$ are subgroups of $S_n$. Hence, their intersection $S_{n,Z} \cap \mathcal{R}(T)$ is a subgroup of $S_{n,Z}$. Let $L$ be a left transversal of $S_{n,Z} \cap \mathcal{R}(T)$ in $S_{n,Z}$. Then:

**(a)** The set $L$ is called a *row-Garnir set* for $T$.

**(b)** Now let $\mathbf{a} = (a_{\overline{P}})_{\overline{P} \in \overline{\operatorname{Tab}}(D)} \in \mathcal{M}^D$ be an element of the Young module $\mathcal{M}^D$. Then, we say that $\mathbf{a}$ satisfies the $(T, L)$-*coGarnir equation* if we have

$$\sum_{x \in L} a_{\overline{xT}} = 0. \tag{421}$$

**Definition 5.23.5.** Let $D$ be a diagram with $|D| = n$. Let $\mathbf{a} \in \mathcal{M}^D$ be an element of the Young module $\mathcal{M}^D$. Then, we say that $\mathbf{a}$ satisfies *all coGarnir equations* if $\mathbf{a}$ satisfies the $(T, L)$-coGarnir equation for each $n$-tableau $T$ of shape $D$ and each row-Garnir set $L$ for $T$.

**Example 5.23.6.** Let $n = 7$ and $D = Y(3,3,1)$ and

$$T = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 & 6 \\ \hline 7 \\ \cline{1-1}\end{array} = 123\backslash\backslash 456\backslash\backslash 7$$

(using the shorthand notation from Convention 5.2.4). Set $j = 1$ and $k = 2$ and $Z = \{2,3,4,5\}$ (the set of entries highlighted in green in the tableau shown above). These $j$, $k$ and $Z$ satisfy all requirements of Definition 5.23.4.

Now, let us choose a left transversal $L$ of $S_{n,Z} \cap \mathcal{R}(T)$ in $S_{n,Z}$, for example

$$L = \{\mathrm{id},\ t_{2,4},\ t_{2,5},\ t_{3,4},\ t_{3,5},\ t_{2,4}t_{3,5}\}.$$

Then, $L$ is a row-Garnir set for $T$. Thus, an element $\mathbf{a} = (a_{\overline{P}})_{\overline{P} \in \overline{\mathrm{Tab}}(D)} \in \mathcal{M}^D$ satisfies the $(T,L)$-coGarnir equation if and only if

$$a_{\overline{\mathrm{id}\, T}} + a_{\overline{t_{2,4}T}} + a_{\overline{t_{2,5}T}} + a_{\overline{t_{3,4}T}} + a_{\overline{t_{3,5}T}} + a_{\overline{t_{2,4}t_{3,5}T}} = 0,$$

that is, if and only if

$$\begin{aligned} &a_{\overline{123\backslash\backslash 456\backslash\backslash 7}} + a_{\overline{143\backslash\backslash 256\backslash\backslash 7}} + a_{\overline{153\backslash\backslash 426\backslash\backslash 7}} + a_{\overline{124\backslash\backslash 356\backslash\backslash 7}} \\ &\quad + a_{\overline{125\backslash\backslash 436\backslash\backslash 7}} + a_{\overline{145\backslash\backslash 236\backslash\backslash 7}} = 0. \end{aligned}$$

Note that there are no minus signs here!

Of course, for $\mathbf{a}$ to satisfy all coGarnir equations, several more equalities need to hold.

**Exercise 5.23.1.** [2] Prove that the $(T,L)$-coGarnir equation (as defined in Definition 5.23.4) can be equivalently rewritten as

$$\sum_{\overline{P} \in \Omega} a_{\overline{P}} = 0,$$

where $\Omega$ is the orbit of $\overline{T} \in \overline{\mathrm{Tab}}(D)$ under the action of the subgroup $S_{n,Z}$ of $S_n$. Thus, this equation depends only on $T$ and $Z$, not on $L$.

We now state our main results:

**Theorem 5.23.7.** Let $D$ be a diagram with $|D| = n$. Let $\mathbf{a} \in \mathcal{S}^D$ be an element of the Specht module $\mathcal{S}^D$. Then, $\mathbf{a}$ (viewed as an element of the Young module $\mathcal{M}^D$) satisfies all coGarnir equations.

**Theorem 5.23.8.** Let $D$ be a skew Young diagram with $|D| = n$. Then, the Specht module $\mathcal{S}^D$ can be described as

$$\mathcal{S}^D = \left\{ \mathbf{a} \in \mathcal{M}^D \ \mid \ \mathbf{a} \text{ satisfies all coGarnir equations} \right\}.$$

(This is an identity, not just an isomorphism!)

We shall prove these two theorems in the three subsections that follow.

### 5.23.2. Transposing the Garnir relations

In preparation for proving Theorem 5.23.7, we shall transform the (original) Garnir relations (206). First, we translate (206) from the Specht module $\mathcal{S}^D$ to its left ideal avatar $\mathbf{k}\left[S_n\right] \cdot \mathbf{E}_T$:

**Corollary 5.23.9.** Let $D$ be a diagram. Let $T$ be an $n$-tableau of shape $D$. Let $j$ and $k$ be two distinct integers. Let $Z$ be a subset of $\operatorname{Col}(j,T) \cup \operatorname{Col}(k,T)$ such that $|Z| > |D\langle j\rangle \cup D\langle k\rangle|$. (Here we are using the notations introduced in Definition 5.8.1 and Definition 5.8.2.)

Define $S_{n,Z}$ as in Proposition 3.7.4. This set $S_{n,Z}$ is a subgroup of $S_n$ (by Proposition 3.7.4 **(a)**, applied to $X = Z$). Thus, both $\mathcal{C}(T)$ and $S_{n,Z}$ are subgroups of $S_n$. Hence, their intersection $S_{n,Z} \cap \mathcal{C}(T)$ is a subgroup of $S_{n,Z}$. Let $L$ be a left transversal of $S_{n,Z} \cap \mathcal{C}(T)$ in $S_{n,Z}$. Then, in $\mathbf{k}\left[S_n\right]$, we have

$$\sum_{x \in L} (-1)^x \, x \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} = 0.$$

*Proof.* We have $|D| = n$ (since $T$ is an $n$-tableau of shape $D$). Corollary 5.5.23 thus shows that there is a left $\mathbf{k}\left[S_n\right]$-module isomorphism

$$\overline{\alpha} : \mathbf{k}\left[S_n\right] \cdot \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \to \mathcal{S}^D$$

that sends $\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$ to $\mathbf{e}_T$. Consider this $\overline{\alpha}$.

Now, (206) yields

$$\sum_{x \in L} (-1)^x \, \mathbf{e}_{xT} = 0. \tag{422}$$

But $\underbrace{\sum_{x \in L} (-1)^x \, x \, \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}}_{\in \mathbf{k}\left[S_n\right]} \in \mathbf{k}\left[S_n\right] \cdot \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$, so that we can apply the

map $\overline{\alpha}$ to $\sum_{x \in L} (-1)^x \, x \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$. We thus obtain

$$\begin{aligned}
&\overline{\alpha}\left(\sum_{x \in L} (-1)^x \, x \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}\right) \\
&= \sum_{x \in L} (-1)^x \, x \quad \underbrace{\overline{\alpha}\left(\nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}\right)}_{\substack{= \mathbf{e}_T \\ \text{(since } \overline{\alpha} \text{ sends } \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \text{ to } \mathbf{e}_T\text{)}}} \qquad \left(\begin{array}{c}\text{since } \overline{\alpha} \text{ is a left } \mathbf{k}[S_n]\text{-module} \\ \text{morphism}\end{array}\right) \\
&= \sum_{x \in L} (-1)^x \quad \underbrace{x \mathbf{e}_T}_{\substack{= \mathbf{e}_{xT} \\ \text{(since Lemma 5.4.6 \textbf{(a)} (applied to } u = x\text{)} \\ \text{yields } \mathbf{e}_{xT} = x\mathbf{e}_T\text{)}}} = \sum_{x \in L} (-1)^x \, \mathbf{e}_{xT} = 0
\end{aligned}$$

(by (422)). In other words, $\sum_{x \in L} (-1)^x \, x \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} \in \operatorname{Ker} \overline{\alpha} = 0$ (since $\overline{\alpha}$ is an isomorphism and thus injective). Hence, $\sum_{x \in L} (-1)^x \, x \nabla^-_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} = 0$. This proves Corollary 5.23.9. □

Next, we transform Corollary 5.23.9 by reflecting the diagram $D$ and the tableau $T$ across the diagonal:

**Corollary 5.23.10.** Let $D$ be a diagram. Let $T$ be an $n$-tableau of shape $D$. Let $j$ and $k$ be two distinct integers. Let $Z$ be a subset of $\operatorname{Row}(j, T) \cup \operatorname{Row}(k, T)$ such that $|Z| > |D\lfloor j \rfloor \cup D\lfloor k \rfloor|$. (Here we are using the notations introduced in Definition 5.23.1 and Definition 5.23.2.)

Define $S_{n,Z}$ as in Proposition 3.7.4. This set $S_{n,Z}$ is a subgroup of $S_n$ (by Proposition 3.7.4 **(a)**, applied to $X = Z$). Thus, both $\mathcal{R}(T)$ and $S_{n,Z}$ are subgroups of $S_n$. Hence, their intersection $S_{n,Z} \cap \mathcal{R}(T)$ is a subgroup of $S_{n,Z}$. Let $L$ be a left transversal of $S_{n,Z} \cap \mathcal{R}(T)$ in $S_{n,Z}$. Then, in $\mathbf{k}[S_n]$, we have

$$\sum_{x \in L} x \nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T} = 0.$$

*Proof of Corollary 5.23.10.* We shall use the notations introduced in Definition 5.8.1 and Definition 5.8.2 as well.

Consider the map $\mathbf{r} : \mathbb{Z}^2 \to \mathbb{Z}^2$ from Theorem 5.1.10. This map $\mathbf{r}$ is an involution, so that it is invertible and its inverse is $\mathbf{r}^{-1} = \mathbf{r}$. For any cell $(i, j) \in \mathbb{Z}^2$, we have the following chain of logical equivalences:

$$\begin{aligned}
((i,j) \in \mathbf{r}(D)) &\iff \left(\mathbf{r}^{-1}(i,j) \in D\right) \\
&\iff ((j,i) \in D) \qquad (423)
\end{aligned}$$

(since $\underbrace{\mathbf{r}^{-1}}_{=\mathbf{r}}(i,j) = \mathbf{r}(i,j) = (j,i)$ (by the definition of $\mathbf{r}$)).

In Definition 5.15.10, we introduced an $n$-tableau $T\mathbf{r}$ of shape $\mathbf{r}(D)$. This $n$-tableau $T\mathbf{r}$ is obtained by reflecting $T$ across the northwest-to-southeast diagonal. This reflection turns the rows of $T$ into the columns of $T\mathbf{r}$. More specifically, it turns the $i$-th row of $T$ into the $i$-th column of $T\mathbf{r}$ for each $i \in \mathbb{Z}$. Thus, for each $i \in \mathbb{Z}$, the entries in the $i$-th row of $T$ are precisely the entries in the $i$-th column of $T\mathbf{r}$. Hence, for each $i \in \mathbb{Z}$, we have

$$\operatorname{Row}(i, T) = \operatorname{Col}(i, T\mathbf{r}) \tag{424}$$

(since $\operatorname{Row}(i, T)$ is the set of all entries in the $i$-th row of $T$, while $\operatorname{Col}(i, T\mathbf{r})$ is the set of all entries in the $i$-th column of $T\mathbf{r}$). Thus, in particular,

$$\underbrace{\operatorname{Row}(j, T)}_{\substack{=\operatorname{Col}(j, T\mathbf{r}) \\ \text{(by (424))}}} \cup \underbrace{\operatorname{Row}(k, T)}_{\substack{=\operatorname{Col}(k, T\mathbf{r}) \\ \text{(by (424))}}} = \operatorname{Col}(j, T\mathbf{r}) \cup \operatorname{Col}(k, T\mathbf{r}).$$

Hence, $Z$ is a subset of $\operatorname{Col}(j, T\mathbf{r}) \cup \operatorname{Col}(k, T\mathbf{r})$ (since $Z$ is a subset of $\operatorname{Row}(j, T) \cup \operatorname{Row}(k, T)$).

Moreover, each $p \in \mathbb{Z}$ satisfies

$$D\lfloor p \rfloor = (\mathbf{r}(D))\langle p \rangle \tag{425}$$

[146]. Hence, in particular, $D\lfloor j \rfloor = (\mathbf{r}(D))\langle j \rangle$ and $D\lfloor k \rfloor = (\mathbf{r}(D))\langle k \rangle$.

We assumed that $|Z| > |D\lfloor j \rfloor \cup D\lfloor k \rfloor|$. In view of $D\lfloor j \rfloor = (\mathbf{r}(D))\langle j \rangle$ and $D\lfloor k \rfloor = (\mathbf{r}(D))\langle k \rangle$, we can rewrite this as

$$|Z| > |(\mathbf{r}(D))\langle j \rangle \cup (\mathbf{r}(D))\langle k \rangle|.$$

From Proposition 5.15.12 **(e)**, we know that $\mathcal{R}(T) = \mathcal{C}(T\mathbf{r})$ and $\mathcal{C}(T) = \mathcal{R}(T\mathbf{r})$.

Recall that $L$ is a left transversal of $S_{n,Z} \cap \mathcal{R}(T)$ in $S_{n,Z}$. In other words, $L$ is a left transversal of $S_{n,Z} \cap \mathcal{C}(T\mathbf{r})$ in $S_{n,Z}$ (since $\mathcal{R}(T) = \mathcal{C}(T\mathbf{r})$). Hence, we

[146] *Proof.* Let $p \in \mathbb{Z}$. Then, $D\lfloor p \rfloor$ is defined as the set of all integers $j$ that satisfy $(p, j) \in D$. In other words,

$$D\lfloor p \rfloor = \{j \in \mathbb{Z} \mid (p, j) \in D\} = \{i \in \mathbb{Z} \mid (p, i) \in D\} \tag{426}$$

(here, we have renamed the index $j$ as $i$). On the other hand, $(\mathbf{r}(D))\langle p \rangle$ is defined as the set of all integers $i$ that satisfy $(i, p) \in \mathbf{r}(D)$ (by Definition 5.8.1). In other words,

$$(\mathbf{r}(D))\langle p \rangle = \{i \in \mathbb{Z} \mid (i, p) \in \mathbf{r}(D)\}. \tag{427}$$

However, if $i \in \mathbb{Z}$ is any integer, then the logical equivalence $((i, p) \in \mathbf{r}(D)) \iff ((p, i) \in D)$ holds (by (423), applied to $(i, p)$ instead of $(i, j)$). Thus, the right hand sides of the equalities (427) and (426) are equal (since the condition $(i, p) \in \mathbf{r}(D)$ in the former can be replaced by the equivalent condition $(p, i) \in D$ in the latter). Thus, the left hand sides of these two equalities are also equal. In other words,

$$(\mathbf{r}(D))\langle p \rangle = D\lfloor p \rfloor.$$

Thus, (425) is proved.

can apply Corollary 5.23.9 to $\mathbf{r}(D)$ and $T\mathbf{r}$ instead of $D$ and $T$. As a result, we obtain

$$\sum_{x\in L}(-1)^x\, x\nabla^-_{\mathrm{Col}(T\mathbf{r})}\nabla_{\mathrm{Row}(T\mathbf{r})}=0. \tag{428}$$

Now, recall the map $T_{\mathrm{sign}}:\mathbf{k}[S_n]\to\mathbf{k}[S_n]$ from Definition 3.11.2. (This map has nothing to do with the tableau $T$, despite the "$T$" in its name.) Theorem 3.11.5 **(a)** says that this map $T_{\mathrm{sign}}:\mathbf{k}[S_n]\to\mathbf{k}[S_n]$ is a $\mathbf{k}$-algebra automorphism. Hence, $T_{\mathrm{sign}}$ is invertible, thus injective.

Furthermore, $T_{\mathrm{sign}}$ is a $\mathbf{k}$-algebra morphism. Thus,

$$\begin{aligned}
&T_{\mathrm{sign}}\left(\sum_{x\in L}x\nabla_{\mathrm{Row}\,T}\nabla^-_{\mathrm{Col}\,T}\right)\\
&=\sum_{x\in L}\underbrace{T_{\mathrm{sign}}(x)}_{\substack{=(-1)^x x\\ \text{(by the definition of } T_{\mathrm{sign}},\\ \text{since } x\in L\subseteq S_n)}}\cdot\underbrace{T_{\mathrm{sign}}\left(\nabla_{\mathrm{Row}\,T}\right)}_{\substack{=\nabla^-_{\mathrm{Col}(T\mathbf{r})}\\ \text{(since (276) (applied to } P=T)\\ \text{yields } \nabla^-_{\mathrm{Col}(T\mathbf{r})}=T_{\mathrm{sign}}(\nabla_{\mathrm{Row}\,T}))}}\cdot\underbrace{T_{\mathrm{sign}}\left(\nabla^-_{\mathrm{Col}\,T}\right)}_{\substack{=\nabla_{\mathrm{Row}(T\mathbf{r})}\\ \text{(since (275) (applied to } P=T)\\ \text{yields } \nabla_{\mathrm{Row}(T\mathbf{r})}=T_{\mathrm{sign}}\left(\nabla^-_{\mathrm{Col}\,T}\right))}}\\
&=\sum_{x\in L}(-1)^x\, x\nabla^-_{\mathrm{Col}(T\mathbf{r})}\nabla_{\mathrm{Row}(T\mathbf{r})}=0\qquad\text{(by (428))}.
\end{aligned}$$

Thus, $\sum_{x\in L}x\nabla_{\mathrm{Row}\,T}\nabla^-_{\mathrm{Col}\,T}\in\mathrm{Ker}\left(T_{\mathrm{sign}}\right)=0$ (since $T_{\mathrm{sign}}$ is injective), so that $\sum_{x\in L}x\nabla_{\mathrm{Row}\,T}\nabla^-_{\mathrm{Col}\,T}=0$. This proves Corollary 5.23.10. □

We can restate Corollary 5.23.10 in a form particularly convenient for our needs:

**Corollary 5.23.11.** Let $D$ be a diagram. Let $T$ be an $n$-tableau of shape $D$. Let $L$ be a row-Garnir set for $T$. Define the Young symmetrizer $\mathbf{E}_T$ as in Definition 5.11.1 **(b)**. Then, in $\mathbf{k}[S_n]$, we have

$$\mathbf{E}_T\sum_{x\in L}x^{-1}=0.$$

*Proof.* Recall that $L$ is a row-Garnir set for $T$. According to Definition 5.23.4 **(a)**, this means that there exist two distinct integers $j$ and $k$ as well as a subset $Z$ of $\mathrm{Row}(j,T)\cup\mathrm{Row}(k,T)$ such that $|Z|>|D\lfloor j\rfloor\cup D\lfloor k\rfloor|$ and such that $L$ is a left transversal of $S_{n,Z}\cap\mathcal{R}(T)$ in $S_{n,Z}$. Consider these $j$, $k$ and $Z$. Then, Corollary 5.23.10 shows that

$$\sum_{x\in L}x\nabla_{\mathrm{Row}\,T}\nabla^-_{\mathrm{Col}\,T}=0. \tag{429}$$

However, Theorem 3.11.14 **(a)** says that the map $S:\mathbf{k}[S_n]\to\mathbf{k}[S_n]$ is a $\mathbf{k}$-algebra anti-automorphism. Applying this map $S$ to the equality (429), we

obtain

$$S\left(\sum_{x\in L} x\nabla_{\operatorname{Row} T}\nabla^{-}_{\operatorname{Col} T}\right) = S(0) = 0$$

(since $S$ is $\mathbf{k}$-linear). Hence,

$$\begin{aligned}
0 &= S\left(\sum_{x\in L} x\nabla_{\operatorname{Row} T}\nabla^{-}_{\operatorname{Col} T}\right) \\
&= \sum_{x\in L} \underbrace{S\left(\nabla^{-}_{\operatorname{Col} T}\right)}_{=\nabla^{-}_{\operatorname{Col} T}} \underbrace{S\left(\nabla_{\operatorname{Row} T}\right)}_{=\nabla_{\operatorname{Row} T}} \underbrace{S(x)}_{=x^{-1}} \\
&\qquad \begin{array}{c}\text{(by Proposition 5.5.10)}\end{array} \quad \begin{array}{c}\text{(by Proposition 5.5.10)}\end{array} \quad \begin{array}{c}\text{(by the definition of } S\text{,}\\ \text{since } x\in L\subseteq S_{n,Z}\subseteq S_n\text{)}\end{array} \\
&\qquad \left(\text{since } S \text{ is a } \mathbf{k}\text{-algebra anti-morphism}\right) \\
&= \sum_{x\in L} \underbrace{\nabla^{-}_{\operatorname{Col} T}\nabla_{\operatorname{Row} T}}_{=\mathbf{E}_T} x^{-1} = \sum_{x\in L} \mathbf{E}_T x^{-1} = \mathbf{E}_T \sum_{x\in L} x^{-1}. \\
&\qquad \begin{array}{c}\text{(by the definition of } \mathbf{E}_T\text{)}\end{array}
\end{aligned}$$

This proves Corollary 5.23.11. □

### 5.23.3. Proof of the coGarnir equations: Correctness

We are now ready to prove Theorem 5.23.7 (and, with it, the "$\subseteq$" part of Theorem 5.23.8):

*Proof of Theorem 5.23.7.* We have $\mathbf{a} \in \mathcal{S}^D \subseteq \mathcal{M}^D = \mathbf{k}^{\left(\overline{\operatorname{Tab}}(D)\right)}$. Thus, $\mathbf{a}$ is a family $(a_{\overline{P}})_{\overline{P}\in\overline{\operatorname{Tab}}(D)}$ of scalars $a_{\overline{P}} \in \mathbf{k}$. Consider these scalars $a_{\overline{P}}$. Thus, $\mathbf{a} = (a_{\overline{P}})_{\overline{P}\in\overline{\operatorname{Tab}}(D)}$.

We must show that $\mathbf{a}$ satisfies all coGarnir equations. In other words, we must show that $\mathbf{a}$ satisfies the $(T, L)$-coGarnir equation for each $n$-tableau $T$ of shape $D$ and each row-Garnir set $L$ for $T$ (because this is what "$\mathbf{a}$ satisfies all coGarnir equations" means, according to Definition 5.23.5).

So let $T$ be an $n$-tableau of shape $D$, and let $L$ be a row-Garnir set for $T$. Then, we must show that $\mathbf{a}$ satisfies the $(T, L)$-coGarnir equation. In other words, we must prove that

$$\sum_{x\in L} a_{\overline{xT}} = 0 \tag{430}$$

(because this is what "$\mathbf{a}$ satisfies the $(T, L)$-coGarnir equation" means, according to Definition 5.23.4 **(b)**, since $\mathbf{a} = (a_{\overline{P}})_{\overline{P}\in\overline{\operatorname{Tab}}(D)}$).

Recall that $L$ is a row-Garnir set for $T$. According to Definition 5.23.4 **(a)**, this means that there exist two distinct integers $j$ and $k$ as well as a subset $Z$ of $\operatorname{Row}(j, T) \cup \operatorname{Row}(k, T)$ such that $|Z| > |D\lfloor j\rfloor \cup D\lfloor k\rfloor|$ and such that $L$ is a left transversal of $S_{n,Z} \cap \mathcal{R}(T)$ in $S_{n,Z}$. Consider these $j$, $k$ and $Z$.

Consider the left $\mathbf{k}[S_n]$-module isomorphism

$$\alpha : \mathbf{k}[S_n]\cdot\nabla_{\operatorname{Row} T} \to \mathcal{M}^D$$

from Theorem 5.5.13 **(a)**. From Proposition 5.5.22 **(c)**, we obtain

$$\alpha^{-1}\left(\sum_{\substack{\overline{S}\text{ is an }n\text{-tabloid}\\ \text{of shape }D}} a_{\overline{S}}\overline{S}\right) = \sum_{u\in S_n} a_{u\overline{T}}u. \tag{431}$$

We can rewrite the summation sign "$\sum_{\substack{\overline{S}\text{ is an }n\text{-tabloid}\\ \text{of shape }D}}$" as "$\sum_{\overline{S}\in\overline{\mathrm{Tab}}(D)}$" (since we have $\overline{\mathrm{Tab}}(D) = \{n\text{-tabloids of shape } D\}$). Hence, the equality (431) rewrites as

$$\alpha^{-1}\left(\sum_{\overline{S}\in\overline{\mathrm{Tab}}(D)} a_{\overline{S}}\overline{S}\right) = \sum_{u\in S_n} a_{u\overline{T}}u. \tag{432}$$

But

$$\begin{aligned}\mathbf{a} &= \left(a_{\overline{P}}\right)_{\overline{P}\in\overline{\mathrm{Tab}}(D)} = \left(a_{\overline{S}}\right)_{\overline{S}\in\overline{\mathrm{Tab}}(D)}\\ &= \sum_{\overline{S}\in\overline{\mathrm{Tab}}(D)} a_{\overline{S}}\overline{S}\end{aligned} \tag{433}$$

(since $\mathcal{M}^D$ is the free $\mathbf{k}$-module $\mathbf{k}^{\left(\overline{\mathrm{Tab}}(D)\right)}$, and thus the family $\left(a_{\overline{S}}\right)_{\overline{S}\in\overline{\mathrm{Tab}}(D)}$ can be rewritten as the $\mathbf{k}$-linear combination $\sum_{\overline{S}\in\overline{\mathrm{Tab}}(D)} a_{\overline{S}}\overline{S}$ of its standard basis vectors $\overline{S}$). In view of this, we can rewrite (432) as

$$\alpha^{-1}\left(\mathbf{a}\right) = \sum_{u\in S_n} a_{u\overline{T}}u. \tag{434}$$

Now, define the Young symmetrizer $\mathbf{E}_T$ as in Definition 5.11.1 **(b)**. Theorem 5.5.13 **(b)** shows that the isomorphism $\alpha$ sends $\mathbf{k}\left[S_n\right]\cdot\nabla^-_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}$ to $\mathcal{S}^D$. Hence, its inverse $\alpha^{-1}$ sends $\mathcal{S}^D$ back to $\mathbf{k}\left[S_n\right]\cdot\nabla^-_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}$. In other words,

$$\alpha^{-1}\left(\mathcal{S}^D\right) = \mathbf{k}\left[S_n\right]\cdot \underbrace{\nabla^-_{\mathrm{Col}\,T}\nabla_{\mathrm{Row}\,T}}_{=\mathbf{E}_T} = \mathbf{k}\left[S_n\right]\cdot\mathbf{E}_T.$$

(by the definition of $\mathbf{E}_T$)

Hence, from $\mathbf{a}\in\mathcal{S}^D$, we obtain $\alpha^{-1}\left(\mathbf{a}\right)\in\alpha^{-1}\left(\mathcal{S}^D\right) = \mathbf{k}\left[S_n\right]\cdot\mathbf{E}_T$. In other words, $\alpha^{-1}\left(\mathbf{a}\right) = \mathbf{b}\mathbf{E}_T$ for some $\mathbf{b}\in\mathbf{k}\left[S_n\right]$. Consider this $\mathbf{b}$. Thus,

$$\underbrace{\alpha^{-1}\left(\mathbf{a}\right)}_{=\mathbf{b}\mathbf{E}_T}\cdot\sum_{x\in L}x^{-1} = \mathbf{b}\ \underbrace{\mathbf{E}_T\sum_{x\in L}x^{-1}}_{=0} = 0.$$

(by Corollary 5.23.11)

Comparing this with

$$\begin{aligned}
&\underbrace{\alpha^{-1}\left(\mathbf{a}\right)}_{=\sum\limits_{u\in S_n} a_{u\overline{T}}u} \cdot \sum_{x\in L} x^{-1} \\
&\text{(by (434))} \\
&= \sum_{u\in S_n} a_{u\overline{T}}u \cdot \sum_{x\in L} x^{-1} = \sum_{x\in L} \sum_{u\in S_n} a_{u\overline{T}}ux^{-1} \\
&= \underbrace{\sum_{x\in L}}_{=\sum\limits_{w\in S_n}} \underbrace{\sum_{w\in S_n}}_{\sum\limits_{x\in L}} a_{wx\overline{T}} \underbrace{wxx^{-1}}_{=w} \qquad \left(\begin{array}{c}\text{here, we substituted } wx \text{ for } u \\ \text{in the inner sum, since} \\ \text{the map } S_n \to S_n,\ w \mapsto wx \\ \text{(for a fixed } x \in L\text{) is a bijection}\end{array}\right) \\
&= \sum_{w\in S_n} \sum_{x\in L} a_{wx\overline{T}}w = \sum_{w\in S_n} \left(\sum_{x\in L} a_{wx\overline{T}}\right) w,
\end{aligned}$$

we obtain

$$\sum_{w\in S_n} \left(\sum_{x\in L} a_{wx\overline{T}}\right) w = 0.$$

Since the standard basis $(w)_{w\in S_n}$ of $\mathbf{k}\left[S_n\right]$ is $\mathbf{k}$-linearly independent (because it is a basis), this equality entails that all the coefficients $\sum\limits_{x\in L} a_{wx\overline{T}}$ must be 0. In other words,

$$\sum_{x\in L} a_{wx\overline{T}} = 0 \qquad \text{for each } w \in S_n.$$

Applying this to $w = 1$ (where 1 means the identity permutation $\mathrm{id} \in S_n$), we obtain

$$\sum_{x\in L} a_{1x\overline{T}} = 0. \tag{435}$$

But each $x \in L$ satisfies $1x\overline{T} = x\overline{T} = \overline{xT}$. Hence, we can rewrite (435) as $\sum\limits_{x\in L} a_{\overline{xT}} = 0$. This proves (430). As we explained, our proof of Theorem 5.23.7 is thus complete. □

### 5.23.4. Proof of the coGarnir equations: Sufficiency

We now approach the proof of Theorem 5.23.8. First, we define two subsets of $\mathcal{M}^D$:

**Definition 5.23.12.** Let $D$ be a diagram with $|D| = n$. Let $\mathrm{Tab}\left(D\right)$ denote the set $\{n\text{-tableaux of shape } D\}$. Define two subsets $\mathcal{C}^D$ and $\mathcal{G}^D$ of $\mathcal{M}^D$ by

$$\mathcal{C}^D := \mathrm{span}_{\mathbf{k}}\left\{\overline{T} \ \mid \ T \in \mathrm{Tab}\left(D\right) \text{ is row-standard but not standard}\right\}$$

and

$$\mathcal{G}^D := \left\{ \mathbf{a} \in \mathcal{M}^D \ \mid \ \mathbf{a} \text{ satisfies all coGarnir equations} \right\}.$$

In a moment, we will see that both $\mathcal{C}^D$ and $\mathcal{G}^D$ are $\mathbf{k}$-submodules of $\mathcal{M}^D$. In general, $\mathcal{C}^D$ is not a left $\mathbf{k}[S_n]$-submodule, however (i.e., not a subrepresentation). As for $\mathcal{G}^D$, we will eventually see that $\mathcal{G}^D = \mathcal{S}^D$ when $D$ is a skew Young diagram; thus, in this case, $\mathcal{G}^D$ is a left $\mathbf{k}[S_n]$-submodule of $\mathcal{M}^D$.

We begin with the easy parts:

**Proposition 5.23.13.** Let $D$ be a diagram with $|D| = n$. Then, both $\mathcal{C}^D$ and $\mathcal{G}^D$ are $\mathbf{k}$-submodules of $\mathcal{M}^D$.

*Proof.* Clearly, the set $\mathcal{C}^D$ is a $\mathbf{k}$-submodule of $\mathcal{M}^D$ (since it is defined as a $\mathbf{k}$-linear span). It thus remains to prove that $\mathcal{G}^D$ is a $\mathbf{k}$-submodule of $\mathcal{M}^D$.

Recall that $\mathcal{M}^D$ is the free $\mathbf{k}$-module $\mathbf{k}^{\left(\overline{\mathrm{Tab}}(D)\right)}$, so that each vector $\mathbf{a} \in \mathcal{M}^D$ can be written as a family $(a_{\overline{P}})_{\overline{P} \in \overline{\mathrm{Tab}}(D)}$, and the entries $a_{\overline{P}}$ of this family are the coordinates of $\mathbf{a}$ (with respect to the standard basis).

By its definition, $\mathcal{G}^D$ is the set of all vectors $\mathbf{a} \in \mathcal{M}^D$ that satisfy all coGarnir equations. However, for any given vector $\mathbf{a} = (a_{\overline{P}})_{\overline{P} \in \overline{\mathrm{Tab}}(D)} \in \mathcal{M}^D$, we have the following chain of equivalences:

$$\begin{aligned}
&(\mathbf{a} \text{ satisfies all coGarnir equations}) \\
&\iff (\mathbf{a} \text{ satisfies the } (T, L)\text{-coGarnir equation for each } n\text{-tableau } T \text{ of shape } D \\
&\qquad \text{and each row-Garnir set } L \text{ for } T) \qquad (\text{by Definition 5.23.5}) \\
&\iff (\text{the equation (421) holds for each } n\text{-tableau } T \text{ of shape } D \\
&\qquad \text{and each row-Garnir set } L \text{ for } T) \qquad (\text{by Definition 5.23.4 } \textbf{(b)}).
\end{aligned}$$

Thus, the statement "$\mathbf{a}$ satisfies all coGarnir equations" (for a given vector $\mathbf{a} \in \mathcal{M}^D$) is tantamount to a set of linear equations on the coordinates of $\mathbf{a}$ (namely, the equations (421) for all $n$-tableaux $T$ of shape $D$ and all row-Garnir sets $L$ for $T$). Therefore, $\mathcal{G}^D$ is the set of all vectors $\mathbf{a} \in \mathcal{M}^D$ that satisfy these linear equations (since $\mathcal{G}^D$ is the set of all vectors $\mathbf{a} \in \mathcal{M}^D$ that satisfy all coGarnir equations).

Thus we have shown that $\mathcal{G}^D$ is the set of solutions of a system of linear equations. Hence, $\mathcal{G}^D$ is a $\mathbf{k}$-submodule of $\mathcal{M}^D$ (since the set of solutions of a system of linear equations is always a $\mathbf{k}$-submodule). This completes the proof of Proposition 5.23.13. □

**Lemma 5.23.14.** Let $D$ be a diagram with $|D| = n$. Then, $\mathcal{S}^D \subseteq \mathcal{G}^D$.

*Proof.* Each element $\mathbf{b} \in \mathcal{S}^D$ satisfies all coGarnir equations (by Theorem 5.23.7, applied to $\mathbf{a} = \mathbf{b}$), and thus satisfies

$$\mathbf{b} \in \left\{ \mathbf{a} \in \mathcal{M}^D \ \mid \ \mathbf{a} \text{ satisfies all coGarnir equations} \right\} = \mathcal{G}^D$$

(by the definition of $\mathcal{G}^D$). In other words, $\mathcal{S}^D \subseteq \mathcal{G}^D$. This proves Lemma 5.23.14. $\square$

**Lemma 5.23.15.** Let $D$ be a diagram with $|D| = n$. Then, $\mathcal{M}^D = \mathcal{C}^D + \mathcal{S}^D$.

*Proof.* Note that $\mathcal{C}^D + \mathcal{S}^D$ is a $\mathbf{k}$-submodule of $\mathcal{M}^D$ (since $\mathcal{C}^D$ and $\mathcal{S}^D$ are $\mathbf{k}$-submodules), and thus is closed under $\mathbf{k}$-linear combinations. In other words, $\operatorname{span}_{\mathbf{k}}\left(\mathcal{C}^D + \mathcal{S}^D\right) \subseteq \mathcal{C}^D + \mathcal{S}^D$.

We shall now show that

$$\overline{T} \in \mathcal{C}^D + \mathcal{S}^D \qquad \text{for each } \overline{T} \in \overline{\operatorname{Tab}}\left(D\right). \tag{436}$$

*Proof of (436):* Recall that the set $\overline{\operatorname{Tab}}\left(D\right) = \{n\text{-tabloids of shape } D\}$ is equipped with a total order – namely, the Young last letter order (which we defined in Proposition 5.7.7). We shall prove (436) by strong induction on $\overline{T}$ with respect to this order:

*Induction step:* Fix some $\overline{P} \in \overline{\operatorname{Tab}}\left(D\right)$. Assume (as the induction hypothesis) that (436) holds for each $\overline{T} \in \overline{\operatorname{Tab}}\left(D\right)$ satisfying $\overline{T} < \overline{P}$ (where $<$ is the smaller relation of the Young last letter order, i.e., the relation introduced in Definition 5.7.5). Our goal is to show that (436) holds for $\overline{T} = \overline{P}$. In other words, our goal is to prove that $\overline{P} \in \mathcal{C}^D + \mathcal{S}^D$.

Clearly, $P$ is an $n$-tableau of shape $D$. In other words, $P \in \operatorname{Tab}\left(D\right)$. By Proposition 5.3.8, we can WLOG assume that $P$ is row-standard[147]. Assume this. We are in one of the following two cases:

*Case 1:* The tableau $P$ is not standard.

*Case 2:* The tableau $P$ is standard.

Let us first consider Case 1. In this case, the tableau $P$ is not standard. Thus, $P \in \operatorname{Tab}\left(D\right)$ is row-standard but not standard. Hence, the standard basis vector $\overline{P}$ of $\mathcal{M}^D$ can be written as $\overline{T}$ for some $n$-tableau $T \in \operatorname{Tab}\left(D\right)$ that is row-standard but not standard (namely, for $T = P$). In other words,

$$\begin{aligned}
\overline{P} &\in \left\{\overline{T} \mid T \in \operatorname{Tab}\left(D\right) \text{ is row-standard but not standard}\right\} \\
&\subseteq \operatorname{span}_{\mathbf{k}}\left\{\overline{T} \mid T \in \operatorname{Tab}\left(D\right) \text{ is row-standard but not standard}\right\} \\
&= \mathcal{C}^D \qquad \left(\text{by the definition of } \mathcal{C}^D\right) \\
&\subseteq \mathcal{C}^D + \mathcal{S}^D.
\end{aligned}$$

[147] Indeed, Proposition 5.3.8 shows that there is a bijection

$$\begin{aligned}
&\text{from } \{\text{row-standard } n\text{-tableaux of shape } D\} \\
&\text{to } \{n\text{-tabloids of shape } D\}
\end{aligned}$$

that sends each row-standard $n$-tableau $T$ to its tabloid $\overline{T}$. Hence, the preimage of the $n$-tabloid $\overline{P} \in \overline{\operatorname{Tab}}\left(D\right) = \{n\text{-tabloids of shape } D\}$ under this bijection is a row-standard $n$-tableau $T$ of shape $D$ satisfying $\overline{T} = \overline{P}$ (since it is a preimage of $\overline{P}$). By replacing $P$ with this row-standard $n$-tableau $T$, we can thus make $P$ row-standard without changing the $n$-tabloid $\overline{P}$.

Thus, $\overline{P} \in \mathcal{C}^D + \mathcal{S}^D$ is proved in Case 1.

Let us now consider Case 2. In this case, the tableau $P$ is standard. Hence, Lemma 5.7.9 **(b)** (applied to $T = P$) shows that

$$\mathbf{e}_P = \overline{P} + \left(\text{a linear combination of } n\text{-tabloids } \overline{S} \text{ with } \overline{S} < \overline{P}\right).$$

Hence,

$$\begin{aligned}\mathbf{e}_P - \overline{P} &= \left(\text{a linear combination of } n\text{-tabloids } \overline{S} \text{ with } \overline{S} < \overline{P}\right) \\ &= \left(\text{a linear combination of } \overline{S} \text{ for } \overline{S} \in \overline{\operatorname{Tab}}(D) \text{ satisfying } \overline{S} < \overline{P}\right).\end{aligned}$$

In other words,

$$\begin{aligned}\mathbf{e}_P - \overline{P} &\in \operatorname{span}_{\mathbf{k}}\left\{\overline{S} \mid \overline{S} \in \overline{\operatorname{Tab}}(D) \text{ satisfying } \overline{S} < \overline{P}\right\} \\ &= \operatorname{span}_{\mathbf{k}}\left\{\overline{T} \mid \overline{T} \in \overline{\operatorname{Tab}}(D) \text{ satisfying } \overline{T} < \overline{P}\right\}\end{aligned}$$

(here, we have renamed the index $\overline{S}$ as $\overline{T}$). But our induction hypothesis says that (436) holds for each $\overline{T} \in \overline{\operatorname{Tab}}(D)$ satisfying $\overline{T} < \overline{P}$. In other words, each $\overline{T} \in \overline{\operatorname{Tab}}(D)$ satisfying $\overline{T} < \overline{P}$ satisfies $\overline{T} \in \mathcal{C}^D + \mathcal{S}^D$. In other words,

$$\left\{\overline{T} \mid \overline{T} \in \overline{\operatorname{Tab}}(D) \text{ satisfying } \overline{T} < \overline{P}\right\} \subseteq \mathcal{C}^D + \mathcal{S}^D.$$

Hence, our above computation becomes

$$\begin{aligned}\mathbf{e}_P - \overline{P} &\in \operatorname{span}_{\mathbf{k}} \underbrace{\left\{\overline{T} \mid \overline{T} \in \overline{\operatorname{Tab}}(D) \text{ satisfying } \overline{T} < \overline{P}\right\}}_{\subseteq \mathcal{C}^D + \mathcal{S}^D} \\ &\subseteq \operatorname{span}_{\mathbf{k}}\left(\mathcal{C}^D + \mathcal{S}^D\right) \subseteq \mathcal{C}^D + \mathcal{S}^D.\end{aligned}$$

In other words, $\mathbf{e}_P - \overline{P} = \mathbf{c} + \mathbf{s}$ for some $\mathbf{c} \in \mathcal{C}^D$ and some $\mathbf{s} \in \mathcal{S}^D$. Consider these $\mathbf{c}$ and $\mathbf{s}$. Observe that $\mathbf{e}_P$ is a polytabloid and thus belongs to $\mathcal{S}^D$ (since $\mathcal{S}^D$ is defined as the span of all polytabloids $\mathbf{e}_T$ with $T \in \operatorname{Tab}(D)$). That is, $\mathbf{e}_P \in \mathcal{S}^D$. Now, solving the equation $\mathbf{e}_P - \overline{P} = \mathbf{c} + \mathbf{s}$ for $\overline{P}$, we find

$$\begin{aligned}\overline{P} = \mathbf{e}_P - (\mathbf{c}+\mathbf{s}) &= \underbrace{\mathbf{e}_P}_{\in \mathcal{S}^D} - \underbrace{\mathbf{s}}_{\in \mathcal{S}^D} - \underbrace{\mathbf{c}}_{\in \mathcal{C}^D} \in \underbrace{\mathcal{S}^D - \mathcal{S}^D}_{\subseteq \mathcal{S}^D} - \mathcal{C}^D \\ &\qquad \text{(since } \mathcal{S}^D \text{ is a } \mathbf{k}\text{-module)} \\ &\subseteq \mathcal{S}^D - \mathcal{C}^D = \underbrace{-\mathcal{C}^D}_{=\mathcal{C}^D} + \mathcal{S}^D = \mathcal{C}^D + \mathcal{S}^D. \\ &\qquad \text{(since } \mathcal{C}^D \text{ is a } \mathbf{k}\text{-module)}\end{aligned}$$

Thus, $\overline{P} \in \mathcal{C}^D + \mathcal{S}^D$ is proved in Case 2.

We have now proved $\overline{P} \in \mathcal{C}^D + \mathcal{S}^D$ in both Cases 1 and 2. Hence, $\overline{P} \in \mathcal{C}^D + \mathcal{S}^D$ always holds. This completes the induction step. Thus, (436) is proved. $\square$

Now, recall that the $\mathbf{k}$-module $\mathcal{M}^D$ has a basis $\left(\overline{T}\right)_{\overline{T} \in \overline{\operatorname{Tab}}(D)}$. Hence,

$$\mathcal{M}^D = \operatorname{span}_{\mathbf{k}} \underbrace{\left\{ \overline{T} \mid \overline{T} \in \overline{\operatorname{Tab}}(D) \right\}}_{\substack{\subseteq \mathcal{C}^D + \mathcal{S}^D \\ \text{(by (436))}}} \subseteq \operatorname{span}_{\mathbf{k}} \left( \mathcal{C}^D + \mathcal{S}^D \right) \subseteq \mathcal{C}^D + \mathcal{S}^D.$$

Combining this with the obvious inclusion $\mathcal{C}^D + \mathcal{S}^D \subseteq \mathcal{M}^D$, we obtain $\mathcal{M}^D = \mathcal{C}^D + \mathcal{S}^D$. This proves Lemma 5.23.15. □

We now come to some combinatorial lemmas. Our next lemma is an analogue of Lemma 5.8.7:

**Lemma 5.23.16.** Let $D$ be the skew Young diagram $Y(\lambda/\mu)$ of a skew partition $\lambda/\mu$. Let $T$ be an $n$-tableau of shape $D$. Let $j$ and $k$ be two integers with $j < k$. Let $i \in \mathbb{Z}$ be such that $(j,i) \in D$ and $(k,i) \in D$. Define two sets

$$X = \left\{ T\left(j,i'\right) \mid i' \in \mathbb{Z} \text{ with } i' \geq i \text{ and } \left(j,i'\right) \in D \right\} \qquad \text{and}$$
$$Y = \left\{ T\left(k,i'\right) \mid i' \in \mathbb{Z} \text{ with } i' \leq i \text{ and } \left(k,i'\right) \in D \right\}.$$

(That is, $X$ is the set of all entries in the $j$-th row of $T$ at the cell $(j,i)$ and further east, whereas $Y$ is the set of all entries in the $k$-th row of $T$ at the cell $(k,i)$ and further west.) Then:

**(a)** We have $X \subseteq \operatorname{Row}(j,T)$ and $Y \subseteq \operatorname{Row}(k,T)$ and $|X \cup Y| > |D\lfloor j \rfloor \cup D\lfloor k \rfloor|$.

**(b)** If the tableau $T$ is row-standard and satisfies $T(j,i) > T(k,i)$, then any element of $X$ is larger than any element of $Y$.

*Proof.* This is entirely analogous to Lemma 5.8.7; the only difference is that the horizontal and vertical directions are interchanged (i.e., rows become columns, north becomes west, each cell $(i,j)$ becomes $(j,i)$, etc.). □

Our next lemma is an analogue of Lemma 5.9.9:

**Lemma 5.23.17.** Let $T$ be an $n$-tableau (of any shape). Let $j$ and $k$ be two integers such that $j < k$. Let $X \subseteq \operatorname{Row}(j,T)$ and $Y \subseteq \operatorname{Row}(k,T)$ be two subsets such that each element of $X$ is larger than each element of $Y$. Let $w \in S_{n,X\cup Y} \setminus \mathcal{R}(T)$. Then, $\overline{wT} > \overline{T}$ (with respect to the Young last letter order).

*Proof.* This is entirely analogous to Lemma 5.9.9; the only difference is that the horizontal and vertical directions are interchanged (i.e., rows become columns, $\mathcal{C}(T)$ becomes $\mathcal{R}(T)$, each cell $(i,j)$ becomes $(j,i)$, etc.). (Of course, the same change needs to be made in Lemma 5.9.11 and its proof.) □

**Lemma 5.23.18.** Let $D$ be a diagram with $|D| = n$. Let $\mathbf{a} \in \mathcal{M}^D$. Write $\mathbf{a}$ in the form $\mathbf{a} = (a_{\overline{P}})_{\overline{P} \in \overline{\mathrm{Tab}}(D)}$ (this can be done, since each element of $\mathcal{M}^D$ can be written in this form). Assume that $\mathbf{a} \in \mathcal{C}^D$. Then,

$$a_{\overline{Q}} = 0 \qquad \text{for all standard tableaux } Q \in \mathrm{Tab}\left(D\right).$$

*Proof.* Recall that an element of $\mathcal{M}^D = \mathbf{k}^{\left(\overline{\mathrm{Tab}}(D)\right)}$ is just a family of scalars, indexed by the $n$-tabloids of shape $D$. These scalars are called the coordinates of this element. If $\mathbf{b} = (b_{\overline{P}})_{\overline{P} \in \overline{\mathrm{Tab}}(D)}$ is any element of $\mathcal{M}^D$, and if $\overline{Q} \in \overline{\mathrm{Tab}}\left(D\right)$ is any $n$-tabloid of shape $D$, then $\left[\overline{Q}\right] \mathbf{b}$ shall denote the $\overline{Q}$-coordinate of $\mathbf{b}$ (that is, the scalar $b_{\overline{Q}}$). Clearly,

$$\left[\overline{Q}\right] \overline{P} = 0 \qquad \text{for any two distinct } \overline{P}, \overline{Q} \in \overline{\mathrm{Tab}}\left(D\right). \tag{437}$$

But $\mathbf{a} \in \mathcal{C}^D = \mathrm{span}_{\mathbf{k}} \left\{ \overline{T} \mid T \in \mathrm{Tab}\left(D\right) \text{ is row-standard but not standard} \right\}$. Hence, we can write $\mathbf{a}$ as a $\mathbf{k}$-linear combination

$$\mathbf{a} = \sum_{\substack{T \in \mathrm{Tab}(D) \text{ is row-standard} \\ \text{but not standard}}} c_{\overline{T}} \overline{T} \tag{438}$$

for some scalars $c_{\overline{T}} \in \mathbf{k}$. Consider these scalars $c_{\overline{T}}$.

Now, fix a standard tableau $Q \in \mathrm{Tab}\left(D\right)$. Then, $Q$ is clearly row-standard. Hence, if some $n$-tableau $T \in \mathrm{Tab}\left(D\right)$ is row-standard but not standard, then $\overline{T} \neq \overline{Q}$ [148] and therefore

$$\left[\overline{Q}\right] \overline{T} = 0 \tag{439}$$

(by (437) (applied to $P = T$), since $\overline{T} \neq \overline{Q}$ shows that $\overline{T}$ and $\overline{Q}$ are distinct).

Now, from (438), we obtain

$$\left[\overline{Q}\right] \mathbf{a} = \left[\overline{Q}\right] \left( \sum_{\substack{T \in \mathrm{Tab}(D) \text{ is row-standard} \\ \text{but not standard}}} c_{\overline{T}} \overline{T} \right) = \sum_{\substack{T \in \mathrm{Tab}(D) \text{ is row-standard} \\ \text{but not standard}}} c_{\overline{T}} \underbrace{\left[\overline{Q}\right] \overline{T}}_{\substack{=0 \\ \text{(by (439))}}} = 0.$$

[148] *Proof.* Proposition 5.3.8 shows that there is a bijection

$$\begin{aligned} &\text{from } \{\text{row-standard } n\text{-tableaux of shape } D\} \\ &\text{to } \{n\text{-tabloids of shape } D\} \end{aligned}$$

that sends each row-standard $n$-tableau $T$ to its tabloid $\overline{T}$. Thus, in particular, this bijection is injective. In other words, if $T_1$ and $T_2$ are two distinct row-standard $n$-tableaux of shape $D$, then $\overline{T_1} \neq \overline{T_2}$.

Now, let $T \in \mathrm{Tab}\left(D\right)$ be an $n$-tableau that is row-standard but not standard. Then, $T \neq Q$ (since $Q$ is standard but $T$ is not). Therefore, $T$ and $Q$ are two distinct row-standard $n$-tableaux of shape $D$. Therefore, $\overline{T} \neq \overline{Q}$ (because if $T_1$ and $T_2$ are two distinct row-standard $n$-tableaux of shape $D$, then $\overline{T_1} \neq \overline{T_2}$).

But $\mathbf{a} = (a_{\overline{P}})_{\overline{P} \in \overline{\operatorname{Tab}}(D)}$. Hence, the definition of $\left[\overline{Q}\right] \mathbf{a}$ shows that $\left[\overline{Q}\right] \mathbf{a} = a_{\overline{Q}}$. Therefore, $a_{\overline{Q}} = \left[\overline{Q}\right] \mathbf{a} = 0$. This proves Lemma 5.23.18. □

**Lemma 5.23.19.** Let $D$ be a skew Young diagram with $|D| = n$. Then, $\mathcal{C}^D \cap \mathcal{G}^D = 0$.

*Proof.* Clearly, $\mathcal{C}^D \cap \mathcal{G}^D$ is a $\mathbf{k}$-submodule of $\mathcal{M}^D$ (since $\mathcal{C}^D$ and $\mathcal{G}^D$ are $\mathbf{k}$-submodules of $\mathcal{M}^D$, by Proposition 5.23.13). Thus, in order to prove that $\mathcal{C}^D \cap \mathcal{G}^D = 0$, it suffices to show that each $\mathbf{a} \in \mathcal{C}^D \cap \mathcal{G}^D$ satisfies $\mathbf{a} = 0$. This is what we will now do.

Let $\mathbf{a} \in \mathcal{C}^D \cap \mathcal{G}^D$. We must show that $\mathbf{a} = 0$.

Write $\mathbf{a} \in \mathcal{C}^D \cap \mathcal{G}^D \subseteq \mathcal{M}^D$ in the form $\mathbf{a} = (a_{\overline{P}})_{\overline{P} \in \overline{\operatorname{Tab}}(D)}$ (this can be done, since each element of $\mathcal{M}^D$ can be written in this form).

We have $\mathbf{a} \in \mathcal{C}^D \cap \mathcal{G}^D \subseteq \mathcal{G}^D$. In other words, $\mathbf{a}$ satisfies all coGarnir equations (by the definition of $\mathcal{G}^D$). Moreover, $\mathbf{a} \in \mathcal{C}^D \cap \mathcal{G}^D \subseteq \mathcal{C}^D$.

We must show that $\mathbf{a} = 0$.

Assume the contrary. Thus, $\mathbf{a} \neq 0$. In other words, $(a_{\overline{P}})_{\overline{P} \in \overline{\operatorname{Tab}}(D)} \neq 0$ (since $\mathbf{a} = (a_{\overline{P}})_{\overline{P} \in \overline{\operatorname{Tab}}(D)}$). In other words, there exists some $\overline{P} \in \overline{\operatorname{Tab}}(D)$ such that $a_{\overline{P}} \neq 0$. Consider the **largest** such $\overline{P}$ with respect to the Young last letter order (i.e., to the total order defined in Proposition 5.7.7). Thus, $a_{\overline{P}} \neq 0$, but every $n$-tabloid $\overline{Q} \in \overline{\operatorname{Tab}}(D)$ satisfying $\overline{Q} > \overline{P}$ satisfies

$$a_{\overline{Q}} = 0 \tag{440}$$

(since $\overline{P}$ is the **largest** $n$-tabloid such that $a_{\overline{P}} \neq 0$).

Clearly, $P$ is an $n$-tableau of shape $D$. In other words, $P \in \operatorname{Tab}(D)$. By Proposition 5.3.8, we can WLOG assume that $P$ is row-standard[149]. Assume this.

If the tableau $P$ was standard, then Lemma 5.23.18 (applied to $Q = P$) would yield $a_{\overline{P}} = 0$ (since $\mathbf{a} \in \mathcal{C}^D$), which would contradict $a_{\overline{P}} \neq 0$. Thus, $P$ cannot be standard. Hence, $P$ cannot be column-standard (since $P$ is row-standard).

In other words, there exist two adjacent cells $(j, i) \in D$ and $(j+1, i) \in D$ such that $P(j, i) \geq P(j+1, i)$ [150]. Consider these two cells. Since all entries of $P$

[149] Indeed, Proposition 5.3.8 shows that there is a bijection

$$\begin{aligned} &\text{from } \{\text{row-standard } n\text{-tableaux of shape } D\} \\ &\text{to } \{n\text{-tabloids of shape } D\} \end{aligned}$$

that sends each row-standard $n$-tableau $T$ to its tabloid $\overline{T}$. Hence, the preimage of the $n$-tabloid $\overline{P} \in \overline{\operatorname{Tab}}(D) = \{n\text{-tabloids of shape } D\}$ under this bijection is a row-standard $n$-tableau $T$ of shape $D$ satisfying $\overline{T} = \overline{P}$ (since it is a preimage of $\overline{P}$). By replacing $P$ with this row-standard $n$-tableau $T$, we can thus make $P$ row-standard without changing the $n$-tabloid $\overline{P}$.

[150] *Proof.* If every $(i, j) \in D$ satisfying $(i+1, j) \in D$ satisfied $P(i, j) < P(i+1, j)$, then $P$ would

are distinct (because $P$ is an $n$-tableau), we have $P(j,i) \neq P(j+1,i)$. Hence, from $P(j,i) \geq P(j+1,i)$, we obtain $P(j,i) > P(j+1,i)$.

Set $k := j+1$. Thus, $k = j+1 > j$, whence $j < k$. Thus, the integers $j$ and $k$ are distinct. Also, from $k = j+1$, we obtain $(k,i) = (j+1,i) \in D$. Finally, $P(j,i) > P(j+1,i) = P(k,i)$ (since $j+1 = k$).

Now, define the two sets

$$X = \left\{ P\left(j,i'\right) \mid i' \in \mathbb{Z} \text{ with } i' \geq i \text{ and } \left(j,i'\right) \in D \right\} \qquad \text{and}$$
$$Y = \left\{ P\left(k,i'\right) \mid i' \in \mathbb{Z} \text{ with } i' \leq i \text{ and } \left(k,i'\right) \in D \right\}.$$

(That is, $X$ is the set of all entries in the $j$-th row of $P$ at the cell $(j,i)$ and further east, whereas $Y$ is the set of all entries in the $k$-th row of $P$ at the cell $(k,i)$ and further west.)

Now, Lemma 5.23.16 **(a)** (applied to $T = P$) yields $X \subseteq \operatorname{Row}(j,P)$ and $Y \subseteq \operatorname{Row}(k,P)$ and $|X \cup Y| > |D\lfloor j \rfloor \cup D\lfloor k \rfloor|$ (since $j < k$). Furthermore, Lemma 5.23.16 **(b)** (applied to $T = P$) yields that any element of $X$ is larger than any element of $Y$ (since $P$ is row-standard and satisfies $P(j,i) > P(k,i)$). Hence, Lemma 5.23.17 (applied to $T = P$) shows that

$$\text{each } w \in S_{n, X \cup Y} \setminus \mathcal{R}(P) \text{ satisfies } \overline{wP} > \overline{P} \tag{441}$$

(with respect to the Young last letter order).

Now, set $Z := X \cup Y$. Then, $Z$ is a subset of $\operatorname{Row}(j,P) \cup \operatorname{Row}(k,P)$ (since $Z = \underbrace{X}_{\subseteq \operatorname{Row}(j,P)} \cup \underbrace{Y}_{\subseteq \operatorname{Row}(k,P)} \subseteq \operatorname{Row}(j,P) \cup \operatorname{Row}(k,P)$) and has size $|Z| = |X \cup Y| > |D\lfloor j \rfloor \cup D\lfloor k \rfloor|$.

Moreover, the set $S_{n,Z}$ (defined as in Proposition 3.7.4) is a subgroup of $S_n$ (by Proposition 3.7.4 **(a)**, applied to $Z$ instead of $X$). Thus, both $\mathcal{R}(T)$ and $S_{n,Z}$ are subgroups of $S_n$. Hence, their intersection $S_{n,Z} \cap \mathcal{R}(T)$ is a subgroup of $S_{n,Z}$. Pick any left transversal $L$ of $S_{n,Z} \cap \mathcal{R}(T)$ in $S_{n,Z}$. Then, $L$ is a row-Garnir set for $P$ (by Definition 5.23.4 **(a)**, applied to $T = P$). Hence, $\mathbf{a}$ satisfies the $(P,L)$-coGarnir equation (by Definition 5.23.5, since $\mathbf{a}$ satisfies all coGarnir equations). In other words, we have

$$\sum_{x \in L} a_{\overline{xP}} = 0 \tag{442}$$

(by Definition 5.23.4 **(b)**).

be column-standard (by Proposition 5.2.7 **(b)**, applied to $T = P$), which would contradict the fact that $P$ is not column-standard. Hence, not every $(i,j) \in D$ satisfying $(i+1,j) \in D$ satisfies $P(i,j) < P(i+1,j)$. In other words, there exists some $(i,j) \in D$ satisfying $(i+1,j) \in D$ but not $P(i,j) < P(i+1,j)$. In other words, there exists some $(i,j) \in D$ satisfying $(i+1,j) \in D$ and $P(i,j) \geq P(i+1,j)$. Renaming the cell $(i,j)$ as $(j,i)$ in this results, we can rewrite it as follows: There exists some $(j,i) \in D$ satisfying $(j+1,i) \in D$ and $P(j,i) \geq P(j+1,i)$. In other words, there exist two adjacent cells $(j,i) \in D$ and $(j+1,i) \in D$ such that $P(j,i) \geq P(j+1,i)$.

Recall that $L$ is a left transversal of $S_{n,Z} \cap \mathcal{R}(P)$ in $S_{n,Z}$. Hence, $L \subseteq S_{n,Z} = S_{n,X \cup Y}$ (since $Z = X \cup Y$). Next, we claim that

$$\text{each } w \in L \setminus \mathcal{R}(P) \text{ satisfies } \overline{wP} > \overline{P} \tag{443}$$

(with respect to the Young last letter order). Indeed, each $w \in L \setminus \mathcal{R}(P)$ satisfies $w \in \underbrace{L}_{\subseteq S_{n,X \cup Y}} \setminus \mathcal{R}(P) \subseteq S_{n,X \cup Y} \setminus \mathcal{R}(P)$ and thus $\overline{wP} > \overline{P}$ (by (441)); thus, (443) is proved.

Recall again that $L$ is a left transversal of $S_{n,Z} \cap \mathcal{R}(P)$ in $S_{n,Z}$. Hence, Lemma 5.8.16 **(a)** (applied to $G = S_n$ and $H = S_{n,Z}$ and $K = \mathcal{R}(P)$) shows that there is exactly one $u \in L$ that belongs to $\mathcal{R}(P)$ (since $L$ is a left transversal of $S_{n,Z} \cap \mathcal{R}(P)$ in $S_{n,Z}$). Consider this $u$. Thus, $u \in \mathcal{R}(P)$. Moreover, Lemma 5.8.16 **(b)** (applied to $G = S_n$ and $H = S_{n,Z}$ and $K = \mathcal{R}(P)$) shows that

$$L \setminus \{u\} = L \setminus \mathcal{R}(P). \tag{444}$$

Proposition 5.3.12 **(a)** (applied to $w = u$ and $T = P$) shows that the tableau $u \rightharpoonup P$ is row-equivalent to $P$ (since $u \in \mathcal{R}(P)$). In other words, $\overline{u \rightharpoonup P} = \overline{P}$ (since row-equivalent $n$-tableaux belong to the same $n$-tabloid). In other words, $\overline{uP} = \overline{P}$ (since $uP = u \rightharpoonup P$).

Now, from (442), we obtain

$$\begin{aligned} 0 &= \sum_{x \in L} a_{\overline{xP}} = \underbrace{a_{\overline{uP}}}_{\substack{=a_{\overline{P}} \\ (\text{since } \overline{uP} = \overline{P})}} + \underbrace{\sum_{\substack{x \in L; \\ x \neq u}} a_{\overline{xP}}}_{= \sum\limits_{x \in L \setminus \{u\}}} \qquad \left( \begin{array}{c} \text{here, we have split off the} \\ \text{addend for } x = u \text{ from the sum} \end{array} \right) \\ &= a_{\overline{P}} + \sum_{x \in L \setminus \{u\}} a_{\overline{xP}}. \end{aligned} \tag{445}$$

Now, $x \in L \setminus \{u\}$ be arbitrary. Thus, $x \in L \setminus \{u\} = L \setminus \mathcal{R}(P)$ (by (444)), so that (441) shows that $\overline{xP} > \overline{P}$. Thus, (440) (applied to $\overline{Q} = \overline{xP}$) yields $a_{\overline{xP}} = 0$.

Forget that we fixed $x$. We thus have shown that $a_{\overline{xP}} = 0$ for each $x \in L \setminus \{u\}$. Hence, $\sum\limits_{x \in L \setminus \{u\}} \underbrace{a_{\overline{xP}}}_{=0} = 0$. Thus, (445) becomes

$$0 = a_{\overline{P}} + \underbrace{\sum_{x \in L \setminus \{u\}} a_{\overline{xP}}}_{=0} = a_{\overline{P}} \neq 0.$$

This is clearly a contradiction. Thus, our assumption was wrong, and $\mathbf{a} = 0$ is proved. This completes our proof of Lemma 5.23.19. $\square$

**Theorem 5.23.20.** Let $D$ be a skew Young diagram with $|D| = n$. Then:

**(a)** We have $\mathcal{G}^D = \mathcal{S}^D$.

**(b)** We have $\mathcal{M}^D = \mathcal{S}^D \oplus \mathcal{C}^D$ (an internal direct sum) as $\mathbf{k}$-modules.

*Proof.* Proposition 5.23.13 shows that $\mathcal{C}^D$ and $\mathcal{G}^D$ are $\mathbf{k}$-submodules of $\mathcal{M}^D$.

**(a)** We have $\mathcal{S}^D \subseteq \mathcal{G}^D$ by Lemma 5.23.14.

Now, let $\mathbf{a} \in \mathcal{G}^D$. Then, $\mathbf{a} \in \mathcal{G}^D \subseteq \mathcal{M}^D = \mathcal{C}^D + \mathcal{S}^D$ by Lemma 5.23.15. Hence, $\mathbf{a} = \mathbf{c} + \mathbf{s}$ for some $\mathbf{c} \in \mathcal{C}^D$ and some $\mathbf{s} \in \mathcal{S}^D$. Consider these $\mathbf{c}$ and $\mathbf{s}$. Now, from $\mathbf{a} = \mathbf{c} + \mathbf{s}$, we obtain

$$\mathbf{c} = \underbrace{\mathbf{a}}_{\in \mathcal{G}^D} - \underbrace{\mathbf{s}}_{\in \mathcal{S}^D \subseteq \mathcal{G}^D} \in \mathcal{G}^D - \mathcal{G}^D \subseteq \mathcal{G}^D \qquad \left(\text{since } \mathcal{G}^D \text{ is a } \mathbf{k}\text{-module}\right).$$

Combining this with $\mathbf{c} \in \mathcal{C}^D$, we obtain $\mathbf{c} \in \mathcal{C}^D \cap \mathcal{G}^D = 0$ by Lemma 5.23.19. Hence, $\mathbf{c} = 0$. Thus, $\mathbf{a} = \underbrace{\mathbf{c}}_{=0} + \mathbf{s} = \mathbf{s} \in \mathcal{S}^D$.

Forget that we fixed $\mathbf{a}$. We thus have shown that $\mathbf{a} \in \mathcal{S}^D$ for each $\mathbf{a} \in \mathcal{G}^D$. In other words, $\mathcal{G}^D \subseteq \mathcal{S}^D$. Combining this with $\mathcal{S}^D \subseteq \mathcal{G}^D$, we obtain $\mathcal{G}^D = \mathcal{S}^D$. This proves Theorem 5.23.20 **(a)**.

**(b)** Lemma 5.23.19 yields $\mathcal{C}^D \cap \mathcal{G}^D = 0$. Since $\mathcal{G}^D = \mathcal{S}^D$ (by Theorem 5.23.20 **(a)**), we can rewrite this as $\mathcal{C}^D \cap \mathcal{S}^D = 0$. Hence, the sum $\mathcal{C}^D + \mathcal{S}^D$ is a direct sum – i.e., we have $\mathcal{C}^D + \mathcal{S}^D = \mathcal{C}^D \oplus \mathcal{S}^D$.

But Lemma 5.23.15 yields $\mathcal{M}^D = \mathcal{C}^D + \mathcal{S}^D = \mathcal{C}^D \oplus \mathcal{S}^D = \mathcal{S}^D \oplus \mathcal{C}^D$ as $\mathbf{k}$-modules. This proves Theorem 5.23.20 **(b)**. □

*Proof of Theorem 5.23.8.* Theorem 5.23.20 **(a)** yields $\mathcal{G}^D = \mathcal{S}^D$. Hence,

$$\mathcal{S}^D = \mathcal{G}^D = \left\{ \mathbf{a} \in \mathcal{M}^D \mid \mathbf{a} \text{ satisfies all coGarnir equations} \right\}.$$

This proves Theorem 5.23.8. □

### 5.23.5. Triangulating row-Garnir sets

Theorem 5.23.8 expresses the Specht module $\mathcal{S}^D$ for a skew Young diagram $D$ as the solution set of a system of linear equations – namely, of the coGarnir equations (421). However, this system is vastly redundant, in the sense that it only takes a (much smaller) subset to determine the same solution set. A class of such sufficient subsets shall now be revealed (in Theorem 5.23.26). These come from *triangulating* sets, which are defined as follows (note the similarity to Definition 5.22.8 **(b)** and **(c)**):

**Definition 5.23.21.** Let $D$ be a diagram with $|D| = n$.

**(a)** We let $\operatorname{RowGar}(D)$ be the set of all pairs $(T, L)$, where $T$ is an $n$-tableau of shape $D$ and where $L$ is a row-Garnir set for $T$.

**(b)** A subset $\mathbf{G}$ of $\operatorname{RowGar}(D)$ will be called *triangulating* if it has the following property: For each $n$-tableau $T$ of shape $D$ that is row-standard but not standard, there exists some row-Garnir set $L$ for $T$ such that $(T, L) \in \mathbf{G}$ and such that

$$\text{each } w \in L \setminus \mathcal{R}(T) \text{ satisfies } \overline{wT} > \overline{T} \tag{446}$$

(with respect to the Young last letter order).

**Example 5.23.22.** Let $n = 4$ and $D = Y(2,2)$. There are $4! = 24$ many $n$-tableaux of shape $D$, but only 6 of them are row-standard, and only 4 of these 6 fail to be standard. Namely, the latter 4 tableaux are

$$\begin{array}{|c|c|}\hline 1&4\\\hline 2&3\\\hline\end{array}\ ,\quad \begin{array}{|c|c|}\hline 2&3\\\hline 1&4\\\hline\end{array}\ ,\quad \begin{array}{|c|c|}\hline 2&4\\\hline 1&3\\\hline\end{array}\ ,\quad \begin{array}{|c|c|}\hline 3&4\\\hline 1&2\\\hline\end{array}\ .$$

Thus, a subset $\mathbf{G}$ of $\operatorname{RowGar}(D)$ is said to be triangulating if it has the property that for each of these 4 tableaux $T$, there exists some row-Garnir set $L$ for $T$ such that $(T, L) \in \mathbf{G}$ and such that (446) holds. This property can be ensured by picking one such $L$ for each $T$; here is one possibility (where the middle column shows the $Z$ from Definition 5.23.4):

| $T$ | $Z$ | $L$ |
|---|---|---|
| $\begin{array}{|c|c|}\hline 1&4\\\hline 2&3\\\hline\end{array}$ | $\{2,3,4\}$ | $\{\mathrm{id}, t_{2,4}, t_{3,4}\}$ |
| $\begin{array}{|c|c|}\hline 2&3\\\hline 1&4\\\hline\end{array}$ | $\{1,2,3\}$ | $\{\mathrm{id}, t_{2,1}, t_{3,1}\}$ |
| $\begin{array}{|c|c|}\hline 2&4\\\hline 1&3\\\hline\end{array}$ | $\{1,2,4\}$ | $\{\mathrm{id}, t_{2,1}, t_{4,1}\}$ |
| $\begin{array}{|c|c|}\hline 3&4\\\hline 1&2\\\hline\end{array}$ | $\{1,3,4\}$ | $\{\mathrm{id}, t_{3,1}, t_{4,1}\}$ |

We have constructed these $L$'s as follows:

- We have picked some $(j, i) \in D$ such that $(j+1, i) \in D$ and $T(j, i) > T(j+1, i)$. (Such a $(j, i)$ exists, since $T$ is not column-standard; we have seen this in our above proof of Lemma 5.23.19, where the tableau $P$

played the role that $T$ is now playing. Sometimes there are two choices for this $(j, i)$; in such cases, either choice is fine, but we chose the lexicographically smaller one.)

- We then defined the two sets

$$X = \left\{ T\left(j, i'\right) \mid i' \in \mathbb{Z} \text{ with } i' \geq i \text{ and } \left(j, i'\right) \in D \right\} \qquad \text{and}$$
$$Y = \left\{ T\left(k, i'\right) \mid i' \in \mathbb{Z} \text{ with } i' \leq i \text{ and } \left(k, i'\right) \in D \right\}$$

(just as we did in the proof of Lemma 5.23.19).

- We set $Z = X \cup Y$ (and we highlighted the elements of $Z$ in green in the above table).
- We chose a left transversal $L$ of $S_{n,Z} \cap \mathcal{R}(T)$ in $S_{n,Z}$. (Here again, there are several choices, but we picked the simplest one, consisting of the identity and transpositions.)

The $L$'s defined in this way always satisfy (446), as we have shown in the proof of Lemma 5.23.19 (see the equality (443), and keep in mind that $T$ is now playing the role of $P$). Thus, the subset

$$\mathbf{G} := \left\{ \left( \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 3 \\ \hline \end{array}\ , \ \{\mathrm{id}, t_{2,4}, t_{3,4}\} \right), \quad \left( \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & 4 \\ \hline \end{array}\ , \ \{\mathrm{id}, t_{2,1}, t_{3,1}\} \right), \right.$$
$$\left( \begin{array}{|c|c|}\hline 2 & 4 \\ \hline 1 & 3 \\ \hline \end{array}\ , \ \{\mathrm{id}, t_{2,1}, t_{4,1}\} \right), \quad \left. \left( \begin{array}{|c|c|}\hline 3 & 4 \\ \hline 1 & 2 \\ \hline \end{array}\ , \ \{\mathrm{id}, t_{3,1}, t_{4,1}\} \right) \right\}$$

of RowGar $(D)$ is triangulating.

**Proposition 5.23.23.** Let $D$ be a diagram with $|D| = n$. Then, the set RowGar $(D)$ is triangulating.

*Proof idea.* We must show that for each $n$-tableau $T$ of shape $D$ that is row-standard but not standard, there exists some row-Garnir set $L$ for $T$ such that $(T, L) \in \mathbf{G}$ and such that

$$\text{each } w \in L \setminus \mathcal{R}(T) \text{ satisfies } \overline{wT} > \overline{T}$$

(with respect to the Young last letter order). But this was (implicitly) shown in our proof of Lemma 5.23.19 (except that $T$ was called $P$ in that proof). Thus, Proposition 5.23.23 is proved. □

**Definition 5.23.24.** Let $D$ be a diagram with $|D| = n$. Let $\mathbf{G}$ be a subset of $\operatorname{RowGar}(D)$.

Let $\mathbf{a} \in \mathcal{M}^D$ be an element of the Young module $\mathcal{M}^D$. Then, we say that $\mathbf{a}$ satisfies *all* $\mathbf{G}$*-coGarnir equations* if $\mathbf{a}$ satisfies the $(T, L)$-coGarnir equation for each $(T, L) \in \mathbf{G}$.

**Example 5.23.25.** Let $n = 4$ and $D = Y(2,2)$. Let $\mathbf{G}$ be the triangulating subset of $\operatorname{RowGar}(D)$ constructed in Example 5.23.22. Let $\mathbf{a} = (a_{\overline{P}})_{\overline{P} \in \overline{\operatorname{Tab}}(D)} \in \mathcal{M}^D$ be an element of $\mathcal{M}^D$. Then, $\mathbf{a}$ satisfies all $\mathbf{G}$-coGarnir equations if and only if it satisfies the four equations

$$a_{\overline{14\backslash\backslash 23}} + a_{\overline{12\backslash\backslash 43}} + a_{\overline{13\backslash\backslash 24}} = 0,$$
$$a_{\overline{23\backslash\backslash 14}} + a_{\overline{13\backslash\backslash 24}} + a_{\overline{21\backslash\backslash 34}} = 0,$$
$$a_{\overline{24\backslash\backslash 13}} + a_{\overline{14\backslash\backslash 23}} + a_{\overline{21\backslash\backslash 43}} = 0,$$
$$a_{\overline{34\backslash\backslash 12}} + a_{\overline{14\backslash\backslash 32}} + a_{\overline{31\backslash\backslash 42}} = 0.$$

It should be kept in mind that the subscripts here are $n$-tabloids, not $n$-tableaux, so several of them are equal (e.g., we have $\overline{31\backslash\backslash 42} = \overline{13\backslash\backslash 24}$). Thus, these are four linear equations on the altogether six coordinates of $\mathbf{a}$. The solution set of these four equations turns out to be precisely the Specht module $\mathcal{S}^D$. This is no coincidence, as the following theorem shows:

**Theorem 5.23.26.** Let $D$ be a skew Young diagram with $|D| = n$. Let $\mathbf{G}$ be a triangulating subset of $\operatorname{RowGar}(D)$. Then, the Specht module $\mathcal{S}^D$ can be described as

$$\mathcal{S}^D = \left\{ \mathbf{a} \in \mathcal{M}^D \ \mid \ \mathbf{a} \text{ satisfies all } \mathbf{G}\text{-coGarnir equations} \right\}.$$

(This is an identity, not just an isomorphism!)

*Proof.* This can be proved in the same way as Theorem 5.23.8, once we make a few changes to Subsection 5.23.4:

- The definition of $\mathcal{G}^D$ (in Definition 5.23.12) must be replaced by

$$\mathcal{G}^D := \left\{ \mathbf{a} \in \mathcal{M}^D \ \mid \ \mathbf{a} \text{ satisfies all } \mathbf{G}\text{-coGarnir equations} \right\}.$$

  (That is, instead of requiring $\mathbf{a}$ to satisfy all coGarnir equations, we only require $\mathbf{a}$ to satisfy all $\mathbf{G}$-coGarnir equations.)

- In the proof of Proposition 5.23.13, we must replace "all coGarnir equations" by "all $\mathbf{G}$-coGarnir equations", and we must replace "for each $n$-tableau $T$ of shape $D$ and each row-Garnir set $L$ for $T$" by "for each $(T, L) \in \mathbf{G}$".

- A similar change needs to be done in the proof of Lemma 5.23.14.
- In the proof of Lemma 5.23.19, we need to pick the row-Garnir set $L$ strategically in order to ensure that $(P, L) \in \mathbf{G}$ (because this is needed to guarantee that $\mathbf{a}$ satisfies the $(P, L)$-coGarnir equation). Fortunately, the triangulating property of $\mathbf{G}$ makes this easy: Indeed, since $\mathbf{G}$ is triangulating, we know (from Definition 5.23.21 **(b)**) that for each $n$-tableau $T$ of shape $D$ that is row-standard but not standard, there exists some row-Garnir set $L$ for $T$ such that $(T, L) \in \mathbf{G}$ and such that

$$\text{each } w \in L \setminus \mathcal{R}(T) \text{ satisfies } \overline{wT} > \overline{T}$$

(with respect to the Young last letter order). We can apply this fact to $T = P$ (since $P$ is an $n$-tableau of shape $D$ that is row-standard but not standard), and thus conclude that there exists some row-Garnir set $L$ for $P$ such that $(P, L) \in \mathbf{G}$ and such that

$$\text{each } w \in L \setminus \mathcal{R}(P) \text{ satisfies } \overline{wP} > \overline{P}$$

(with respect to the Young last letter order). Consider this $L$. Since $\mathbf{a}$ satisfies all $\mathbf{G}$-coGarnir equations (because $\mathbf{a} \in \mathcal{G}^D$), we thus see that $\mathbf{a}$ satisfies the $(P, L)$-coGarnir equation (since $(P, L) \in \mathbf{G}$). In other words,

$$\sum_{x \in L} a_{\overline{xP}} = 0 \qquad \left(\text{by Definition 5.23.4 } \mathbf{(b)}\right).$$

Now, recall that $L$ is a left transversal of $S_{n,Z} \cap \mathcal{R}(P)$ in $S_{n,Z}$ for some subset $Z$ of $[n]$ (because $L$ is a row-Garnir set for $P$). Consider this $Z$. Now, proceed as in the original proof of Lemma 5.23.19, starting with the sentence "Recall again that $L$ is a left transversal of $S_{n,Z} \cap \mathcal{R}(P)$ in $S_{n,Z}$". This results in a contradiction, just as in the original proof, and so the proof is again complete.

Altogether, this modified version of Subsection 5.23.4 culminates in a proof of Theorem 5.23.8 but with

$$\left\{\mathbf{a} \in \mathcal{M}^D \mid \mathbf{a} \text{ satisfies all coGarnir equations}\right\}$$

replaced by

$$\left\{\mathbf{a} \in \mathcal{M}^D \mid \mathbf{a} \text{ satisfies all } \mathbf{G}\text{-coGarnir equations}\right\}.$$

But this is precisely the claim of Theorem 5.23.26. □

## 5.24. Appendix: Young's proof of the hook length formula

In this section, we shall prove Theorem 5.2.8 (the hook length formula). The proof I will present is neither the simplest nor the most elementary; but it has the distinction of being the historically first proof, as it comes from a 1902 paper by Young himself [Young77, *On Quantitative Substitutional Analysis (Second Paper)*, §4–§5]. Arguably, Young did not know Theorem 5.2.8 in the form we stated it in, but rather was proving something that only later turned out to be equivalent to Theorem 5.2.10, which is itself an equivalent version of Theorem 5.2.8. Thus, the proof I shall give in this section is a somewhat liberal modification of Young's, although all the main ideas have been preserved. The proof is long and rather intricate, and bears more semblance to a proof in a modern research paper than to a proof of a classical result in a standard textbook.[151] In my presentation, the proof gains additional length, as numerous steps that Young considered obvious enough to omit (often for lack of rigorous language to explain them in) have been expanded into lemmas and auxiliary claims with detailed proofs. In some cases, I was unable to understand Young's original arguments and had to replace them by longer but clearer justification (Proposition 5.24.1, which loosely corresponds to [Young77, *On Quantitative Substitutional Analysis (Second Paper)*, §4], is one such case).[152]

### 5.24.1. A consequence of the Garnir relations

At the core of our proof will be an identity that relates a row symmetrizer $\nabla_{\operatorname{Row} T}$, a column antisymmetrizer $\nabla^{-}_{\operatorname{Col} T}$ and several transpositions $t_{u,v}$:

**Proposition 5.24.1.** Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $\lambda$. Let $j$ and $k$ be two positive integers with $j > k$. Let $q \in \left[\lambda_j\right]$. Then,

$$\left(1 + \sum_{s=1}^{\lambda_k} t_{T(k,s),\ T(j,q)}\right) \nabla_{\operatorname{Row} T} \nabla^{-}_{\operatorname{Col} T} = 0.$$

[151] Garsia refers to it as a "gruesome brute force proof" in [GarEge20, Remark 1.5]. This is somewhat unfair, as it falsely suggests a straightforward approach that merely requires a lot of work. But the gruesomeness is hard to argue against!

[152] Another exposition of Young's proof appears in [RotGui71, Chapter II], which stays closer to the original argument but avoids formalizing it completely as "the notation would get hopelessly complicated".

**Example 5.24.2.** Let $n = 14$ and $\lambda = (5,4,3,2)$. Let $j = 3$ and $k = 2$ and $q = 2$. Let $T$ be the following $n$-tableau of shape $\lambda$:

$$T = \begin{array}{|c|c|c|c|c|}\hline 10 & 11 & 12 & 13 & 14 \\ \hline 6 & 7 & 8 & 9 \\ \cline{1-4} 3 & 4 & 5 \\ \cline{1-3} 1 & 2 \\ \cline{1-2}\end{array} .$$

Here, we have highlighted the cell $(j,q) = (3,2)$ green and the cells of the form $(k,s)$ (for $s \in [\lambda_k]$) yellow. Proposition 5.24.1 yields that

$$(1 + t_{6,4} + t_{7,4} + t_{8,4} + t_{9,4})\, \nabla_{\operatorname{Row} T} \nabla^{-}_{\operatorname{Col} T} = 0.$$

There are several ways to prove Proposition 5.24.1. Rota (in [RotGui71, Chapter II, Prop 2]) derives it by an artful application of Theorem 5.10.8.[153] We shall instead prove Proposition 5.24.1 by identifying it as a particular case of Corollary 5.23.10 (which, as we recall, is a consequence of the Garnir relations), obtained by choosing a specific left transversal $L$ (consisting of the identity and some transpositions). This is made possible by the following simple lemma:

**Lemma 5.24.3.** Let $X$ be a subset of $[n]$. Let $x \in X$. Consider the set $S_{n,X}$ defined in Proposition 3.7.4, and the set $S_{n,X\setminus\{x\}}$ defined similarly. Then:

**(a)** We have

$$\{w \in S_{n,X} \mid w(x) = x\} = S_{n,X\setminus\{x\}}. \tag{447}$$

**(b)** The set $S_{n,X}$ is a subgroup of $S_n$, whereas the set $S_{n,X\setminus\{x\}}$ is a subgroup of $S_{n,X}$.

**(c)** For each $u \in S_{n,X}$, we have $u(x) \in X$.

**(d)** The set $\{\mathrm{id}\} \cup \{t_{y,x} \mid y \in X \setminus \{x\}\}$ is a left transversal of $S_{n,X\setminus\{x\}}$ in $S_{n,X}$.

**(e)** For each $y \in X$, fix a permutation $\sigma_y \in S_{n,X}$ that sends $x$ to $y$. Then, the set $\{\sigma_y \mid y \in X\}$ is a left transversal of $S_{n,X\setminus\{x\}}$ in $S_{n,X}$.

*Proof.* This was already implicit in the proof of Lemma 3.7.10 (except that our $\sigma_y$ now correspond to the $\sigma_y^{-1}$ in Lemma 3.7.10), but let us give a detailed proof nevertheless.

We know that $S_{n,X}$ is a subgroup of $S_n$ (by Proposition 3.7.4 **(a)**). The same argument (applied to $X \setminus \{x\}$ instead of $X$) shows that $S_{n,X\setminus\{x\}}$ is a subgroup of $S_n$.

**(a)** The equality (447) is precisely the equality (37) that we showed in our proof of Lemma 3.7.10 above. Thus, Lemma 5.24.3 **(a)** is proved.

[153] This might be the argument Young intended to make in [Young77, *On Quantitative Substitutional Analysis (Second Paper)*, §4].

**(b)** We have already shown that $S_{n,X}$ is a subgroup of $S_n$. From (447), we obtain

$$S_{n,X\setminus\{x\}} = \{w \in S_{n,X} \mid w(x) = x\} \subseteq S_{n,X}.$$

Thus, $S_{n,X\setminus\{x\}}$ is a subgroup of $S_{n,X}$ (since both $S_{n,X\setminus\{x\}}$ and $S_{n,X}$ are subgroups of $S_n$). This completes the proof of Lemma 5.24.3 **(b)**.

**(c)** Let $u \in S_{n,X}$. Thus, $u \in S_n$ and $u(i) = i$ for all $i \notin X$ (by the definition of $S_{n,X}$).

We must show that $u(x) \in X$. Assume the contrary. Thus, $u(x) \notin X$. Set $z := u(x)$. Thus, $z = u(x) \notin X$. But recall that $u(i) = i$ for all $i \notin X$. Applying this to $i = z$, we find $u(z) = z$ (since $z \notin X$). Hence, $u(z) = z = u(x)$. Hence, $z = x$ (since $u$ is a permutation and thus injective), so that $z = x \in X$. This contradicts $z \notin X$. This contradiction shows that our assumption was false. Hence, $u(x) \in X$ must hold. This proves Lemma 5.24.3 **(c)**.

**(e)** Let $L := \left\{\sigma_y \mid y \in X\right\}$. This is clearly a subset of $S_{n,X}$.

We know (from Lemma 5.24.3 **(b)**) that $S_{n,X\setminus\{x\}}$ is a subgroup of $S_{n,X}$. Let $uS_{n,X\setminus\{x\}}$ be any left coset of $S_{n,X\setminus\{x\}}$ in $S_{n,X}$. We shall show that $uS_{n,X\setminus\{x\}}$ contains exactly one element of $L$.

Indeed, $u \in S_{n,X}$.

Let $z := u(x)$. Then, $z = u(x) \in X$ (by Lemma 5.24.3 **(c)**). Hence, the permutation $\sigma_z \in S_{n,X}$ is well-defined, and belongs to $L$ (since $L = \left\{\sigma_y \mid y \in X\right\}$). Thus, $\sigma_z \in L$. Moreover, the definition of $\sigma_z$ shows that $\sigma_z$ sends $x$ to $z$. In other words, $\sigma_z(x) = z$. From $\sigma_z \in S_{n,X}$ and $u \in S_{n,X}$, we obtain $u^{-1}\sigma_z \in S_{n,X}$ (since $S_{n,X}$ is a group). Moreover,

$$\left(u^{-1}\sigma_z\right)(x) = u^{-1}\left(\underbrace{\sigma_z(x)}_{=z=u(x)}\right) = u^{-1}(u(x)) = x.$$

Thus, $u^{-1}\sigma_z$ is a $w \in S_{n,X}$ that satisfies $w(x) = x$ (since $u^{-1}\sigma_z \in S_{n,X}$). In other words,

$$u^{-1}\sigma_z \in \{w \in S_{n,X} \mid w(x) = x\} = S_{n,X\setminus\{x\}} \qquad \text{(by (447))}.$$

In other words, $\sigma_z \in uS_{n,X\setminus\{x\}}$. Combining this with $\sigma_z \in L$, we obtain $\sigma_z \in \left(uS_{n,X\setminus\{x\}}\right) \cap L$. Thus,

$$\{\sigma_z\} \subseteq \left(uS_{n,X\setminus\{x\}}\right) \cap L. \tag{448}$$

Let us now prove that $\left(uS_{n,X\setminus\{x\}}\right) \cap L = \{\sigma_z\}$. Indeed, let $g \in \left(uS_{n,X\setminus\{x\}}\right) \cap L$. Thus, $g \in uS_{n,X\setminus\{x\}}$ and $g \in L$. From $g \in L = \left\{\sigma_y \mid y \in X\right\}$, we see that $g = \sigma_y$ for some $y \in X$. Consider this $y$. From $g \in uS_{n,X\setminus\{x\}}$, we see that $g = uv$ for some $v \in S_{n,X\setminus\{x\}}$. Consider this $v$. We have $v \in S_{n,X\setminus\{x\}} = \{w \in S_{n,X} \mid w(x) = x\}$ (by (447)), so that $v(x) = x$. Now, $\underbrace{g}_{=uv}(x) = (uv)(x) = u\left(\underbrace{v(x)}_{=x}\right) = u(x) = z$. However, from $g = \sigma_y$, we obtain $g(x) = \sigma_y(x) = y$ (since $\sigma_y$ sends $x$ to $y$). Comparing this with $g(x) = z$, we find $y = z$. Now, $g = \sigma_y = \sigma_z$ (since $y = z$), so that $g \in \{\sigma_z\}$.

Forget that we fixed $g$. We thus have shown that $g \in \{\sigma_z\}$ for each $g \in \left(uS_{n,X\setminus\{x\}}\right) \cap L$. In other words, $\left(uS_{n,X\setminus\{x\}}\right) \cap L \subseteq \{\sigma_z\}$. Combining this with (448), we obtain

$$\left(uS_{n,X\setminus\{x\}}\right) \cap L = \{\sigma_z\}.$$

Hence, $\left|\left(uS_{n,X\setminus\{x\}}\right) \cap L\right| = |\{\sigma_z\}| = 1$. In other words, the intersection $\left(uS_{n,X\setminus\{x\}}\right) \cap L$ contains exactly one element. In other words, the set $uS_{n,X\setminus\{x\}}$ contains exactly one element of $L$.

Forget that we fixed $uS_{n,X\setminus\{x\}}$. We thus have shown that each left coset $uS_{n,X\setminus\{x\}}$ contains exactly one element of $L$. In other words, $L$ is a left transversal of $S_{n,X\setminus\{x\}}$ in $S_{n,X}$ (by the definition of a left transversal). In other words, $\left\{\sigma_y \mid y \in X\right\}$ is a left transversal of $S_{n,X\setminus\{x\}}$ in $S_{n,X}$ (since $L = \left\{\sigma_y \mid y \in X\right\}$). This proves Lemma 5.24.3 **(e)**.

**(d)** The transposition $t_{y,x}$ is well-defined for each $y \in X \setminus \{x\}$, and belongs to $S_{n,X}$ (since both $y$ and $x$ belong to $X$). Let us furthermore define a permutation $t_{x,x} \in S_{n,X}$ by $t_{x,x} := \mathrm{id}$. Then, for each $y \in X$, the permutation $t_{y,x}$ is well-defined and belongs to $S_{n,X}$. Moreover, this permutation $t_{y,x}$ sends $x$ to $y$ (by its definition when $y \neq x$, and for obvious reasons when $y = x$). Hence, Lemma 5.24.3 **(e)** (applied to $\sigma_y = t_{y,x}$) shows that the set $\left\{t_{y,x} \mid y \in X\right\}$ is a left transversal of $S_{n,X\setminus\{x\}}$ in $S_{n,X}$. Since

$$\begin{aligned}\left\{t_{y,x} \mid y \in X\right\} &= \left\{\underbrace{t_{x,x}}_{=\mathrm{id}}\right\} \cup \left\{t_{y,x} \mid y \in X \setminus \{x\}\right\} \\ &\qquad \left(\begin{array}{c}\text{here, we have split off the element for } y = x, \\ \text{since } x \in X\end{array}\right) \\ &= \{\mathrm{id}\} \cup \left\{t_{y,x} \mid y \in X \setminus \{x\}\right\},\end{aligned}$$

we can rewrite this as follows: The set $\{\mathrm{id}\} \cup \left\{t_{y,x} \mid y \in X \setminus \{x\}\right\}$ is a left transversal of $S_{n,X\setminus\{x\}}$ in $S_{n,X}$. This proves Lemma 5.24.3 **(d)**. □

We will also use two nearly trivial facts about the rows of a Young diagram $Y(\lambda)$. Recall Definition 5.23.1 and Definition 5.23.2.

**Lemma 5.24.4.** Let $\lambda$ be a partition of $n$. Let $D = Y(\lambda)$. Let $i$ be a positive integer. Then, $D\lfloor i \rfloor = [\lambda_i]$.

*Proof.* We have $D = Y(\lambda)$. Thus, the cells in the $i$-th row of $D$ are $(i,1), (i,2), \ldots, (i,\lambda_i)$ (by the definition of $Y(\lambda)$). In other words, the integers $j$ that satisfy $(i,j) \in D$ are $1, 2, \ldots, \lambda_i$.

But $D\lfloor i \rfloor$ was defined as the set of all integers $j$ that satisfy $(i,j) \in D$. Hence, $D\lfloor i \rfloor = \{1, 2, \ldots, \lambda_i\}$ (since the integers $j$ that satisfy $(i,j) \in D$ are $1, 2, \ldots, \lambda_i$). Thus, $D\lfloor i \rfloor = \{1, 2, \ldots, \lambda_i\} = [\lambda_i]$. This proves Lemma 5.24.4. □

**Proposition 5.24.5.** Let $D$ be a diagram. Let $T$ be an $n$-tableau of shape $D$. Let $j \in \mathbb{Z}$. Then, $|\operatorname{Row}(j,T)| = |D\lfloor j \rfloor|$.

*Proof.* This is the analogue of Proposition 5.8.3 for rows instead of columns. The proof is analogous. □

Two more lemmas will be useful. Both lemmas are easy, and I recommend proving them yourself:

**Lemma 5.24.6.** Let $D$ be a diagram. Let $T$ be an $n$-tableau of shape $D$. Let $k \in \mathbb{Z}$. Then, $S_{n,\operatorname{Row}(k,T)} \subseteq \mathcal{R}(T)$. (For the definition of $S_{n,\operatorname{Row}(k,T)}$, see Proposition 3.7.4.)

*Proof.* Let $u \in S_{n,\operatorname{Row}(k,T)}$. Thus,

$$u \in S_{n,\operatorname{Row}(k,T)} = \{w \in S_n \mid w(i) = i \text{ for all } i \notin \operatorname{Row}(k,T)\}$$

(by the definition of $S_{n,\operatorname{Row}(k,T)}$). In other words, we have $u \in S_n$ and

$$u(i) = i \text{ for all } i \notin \operatorname{Row}(k,T). \tag{449}$$

Now, let $i \in [n]$. We shall show that the number $u(i)$ lies in the same row of $T$ as $i$ does. If $u(i) = i$, then this is obviously true. Thus, for the rest of this proof, we WLOG assume that $u(i) \neq i$. Thus, $i \in \operatorname{Row}(k,T)$ (because otherwise, we would have $i \notin \operatorname{Row}(k,T)$ and thus $u(i) = i$ (by (449)), which would contradict $u(i) \neq i$). In other words, $i$ lies in the $k$-th row of $T$. Furthermore, Lemma 5.24.3 **(c)** (applied to $X = \operatorname{Row}(k,T)$ and $x = i$) yields $u(i) \in \operatorname{Row}(k,T)$ (since $i \in \operatorname{Row}(k,T)$ and $u \in S_{n,\operatorname{Row}(k,T)}$). In other words, $u(i)$ lies in the $k$-th row of $T$. Thus, we have shown that both $u(i)$ and $i$ lie in the $k$-th row of $T$. Therefore, the number $u(i)$ lies in the same row of $T$ as $i$ does.

Forget that we fixed $i$. We thus have shown that for each $i \in [n]$, the number $u(i)$ lies in the same row of $T$ as $i$ does. In other words, the permutation $u$ is horizontal for $T$ (by the definition of "horizontal"). In other words, $u \in \mathcal{R}(T)$.

Forget that we fixed $u$. We thus have proved that $u \in \mathcal{R}(T)$ for each $u \in S_{n,\operatorname{Row}(k,T)}$. In other words, $S_{n,\operatorname{Row}(k,T)} \subseteq \mathcal{R}(T)$. This proves Lemma 5.24.6. □

**Lemma 5.24.7.** Let $D$ be a diagram. Let $T$ be an $n$-tableau of shape $D$. Let $j$ and $k$ be two distinct integers. Let $z \in \operatorname{Row}(j,T)$ be arbitrary. Let $Z$ be the set $\{z\} \cup \operatorname{Row}(k,T)$. Then, $|Z| = |D\lfloor k \rfloor| + 1$ and $Z \setminus \{z\} = \operatorname{Row}(k,T)$ and $S_{n,Z} \cap \mathcal{R}(T) = S_{n,Z\setminus\{z\}}$.

*Proof.* Proposition 5.24.5 (applied to $k$ instead of $j$) yields $|\operatorname{Row}(k,T)| = |D\lfloor k \rfloor|$.

We have $z \in \operatorname{Row}(j,T)$. In other words, $z$ is an entry in the $j$-th row of $T$ (since $\operatorname{Row}(j,T)$ is defined as the set of all entries in the $j$-th row of $T$).

But $T$ is an $n$-tableau, and thus is injective. Hence, all entries of $T$ are distinct. In particular, an entry in the $j$-th row of $T$ cannot also be found in the $k$-th row of $T$ (since $j$ and $k$ are distinct). In other words, an entry in the $j$-th row of $T$ cannot belong to $\operatorname{Row}(k,T)$ (since $\operatorname{Row}(k,T)$ is defined as the set of all entries in the $k$-th row of $T$).

Hence, in particular, $z$ cannot belong to $\operatorname{Row}(k,T)$ (since $z$ is an entry in the $j$-th row of $T$). In other words, $z \notin \operatorname{Row}(k,T)$. Hence, $|\operatorname{Row}(k,T) \cup \{z\}| = |\operatorname{Row}(k,T)| + 1$.

The definition of $Z$ shows that

$$Z = \{z\} \cup \operatorname{Row}(k,T) = \operatorname{Row}(k,T) \cup \{z\}.$$

Hence,

$$|Z| = |\operatorname{Row}(k,T) \cup \{z\}| = \underbrace{|\operatorname{Row}(k,T)|}_{=|D\lfloor k \rfloor|} + 1 = |D\lfloor k \rfloor| + 1.$$

From $Z = \operatorname{Row}(k,T) \cup \{z\}$, we obtain

$$Z \setminus \{z\} = (\operatorname{Row}(k,T) \cup \{z\}) \setminus \{z\} = \operatorname{Row}(k,T)$$

(since $z \notin \operatorname{Row}(k,T)$) and

$$Z \setminus \operatorname{Row}(k,T) = (\operatorname{Row}(k,T) \cup \{z\}) \setminus \operatorname{Row}(k,T) = \{z\}$$

(again since $z \notin \operatorname{Row}(k,T)$). Also, $z \in \{z\} \subseteq \{z\} \cup \operatorname{Row}(k,T) = Z$. Hence, Lemma 5.24.3 **(a)** (applied to $X = Z$ and $x = z$) shows that

$$\{w \in S_{n,Z} \mid w(z) = z\} = S_{n,Z\setminus\{z\}}. \tag{450}$$

Hence, $S_{n,Z\setminus\{z\}} = \{w \in S_{n,Z} \mid w(z) = z\} \subseteq S_{n,Z}$.

It remains to show that $S_{n,Z} \cap \mathcal{R}(T) = S_{n,Z\setminus\{z\}}$. For this purpose, we shall prove that $S_{n,Z} \cap \mathcal{R}(T) \subseteq S_{n,Z\setminus\{z\}}$ and $S_{n,Z\setminus\{z\}} \subseteq S_{n,Z} \cap \mathcal{R}(T)$.

Let $u \in S_{n,Z} \cap \mathcal{R}(T)$ be arbitrary. Thus, $u \in S_{n,Z}$ and $u \in \mathcal{R}(T)$. We have $z \in Z$. Thus, from $u \in S_{n,Z}$, we obtain $u(z) \in Z$ (by Lemma 5.24.3 **(c)**, applied to $X = Z$ and $x = z$). However, from $u \in \mathcal{R}(T)$, we see that the permutation $u$ is horizontal for $T$. Thus, for each $i \in [n]$, the number $u(i)$ lies in the same row of $T$ as $i$ does (by the definition of a horizontal permutation). Applying this to $i = z$, we see that the number $u(z)$ lies in the same row of $T$ as $z$ does. Since $z$ lies in the $j$-th row of $T$, we thus conclude that $u(z)$ lies in the $j$-th row of $T$ as well. In other words, $u(z)$ is an entry in the $j$-th row of $T$. Hence, $u(z)$ cannot belong to $\operatorname{Row}(k,T)$ (since an entry in the $j$-th row of $T$ cannot belong to $\operatorname{Row}(k,T)$). In other words, $u(z) \notin \operatorname{Row}(k,T)$.

Combining $u(z) \in Z$ with $u(z) \notin \operatorname{Row}(k,T)$, we find

$$u(z) \in Z \setminus \operatorname{Row}(k,T) = \{z\}.$$

In other words, $u(z) = z$. Hence, $u \in S_{n,Z}$ and $u(z) = z$. In other words,

$$u \in \{w \in S_{n,Z} \mid w(z) = z\} = S_{n,Z\setminus\{z\}} \qquad \text{(by (450))}.$$

Forget that we fixed $u$. We thus have shown that $u \in S_{n,Z\setminus\{z\}}$ for each $u \in S_{n,Z} \cap \mathcal{R}(T)$. In other words,

$$S_{n,Z} \cap \mathcal{R}(T) \subseteq S_{n,Z\setminus\{z\}}. \tag{451}$$

Let us now prove the reverse inclusion. We have

$$S_{n,Z\setminus\{z\}} = \underbrace{S_{n,Z\setminus\{z\}}}_{\subseteq S_{n,Z}} \cap \underbrace{S_{n,Z\setminus\{z\}}}_{\substack{=S_{n,\operatorname{Row}(k,T)} \\ \text{(since } Z\setminus\{z\}=\operatorname{Row}(k,T)\text{)}}} \subseteq S_{n,Z} \cap \underbrace{S_{n,\operatorname{Row}(k,T)}}_{\substack{\subseteq \mathcal{R}(T) \\ \text{(by Lemma 5.24.6)}}} \subseteq S_{n,Z} \cap \mathcal{R}(T).$$

Combining this with (451), we obtain $S_{n,Z} \cap \mathcal{R}(T) = S_{n,Z\setminus\{z\}}$. This completes the proof of Lemma 5.24.7 (since $|Z| = |D\lfloor k\rfloor| + 1$ and $Z \setminus \{z\} = \operatorname{Row}(k,T)$ have already been proved). $\square$

We are now ready to prove Proposition 5.24.1:

*Proof of Proposition 5.24.1.* Let $D := Y(\lambda)$.

We know that $T$ is an $n$-tableau of shape $\lambda$. In other words, $T$ is an $n$-tableau of shape $Y(\lambda)$. In other words, $T$ is an $n$-tableau of shape $D$ (since $D = Y(\lambda)$). Hence, $T$ is injective (since any $n$-tableau is injective). In other words, all entries of $T$ are distinct.

Let $z := T(j,q)$. Thus, $z$ is an entry in the $j$-th row of $T$ (since $(j,q)$ is a cell in the $j$-th row). In other words, $z \in \operatorname{Row}(j,T)$. Thus, $\{z\} \subseteq \operatorname{Row}(j,T)$.

Let $Z$ be the set $\{z\} \cup \operatorname{Row}(k,T)$. Thus,

$$Z = \underbrace{\{z\}}_{\subseteq \operatorname{Row}(j,T)} \cup \operatorname{Row}(k,T) \subseteq \operatorname{Row}(j,T) \cup \operatorname{Row}(k,T).$$

Thus, $Z$ is a subset of $\operatorname{Row}(j,T) \cup \operatorname{Row}(k,T)$.

The numbers $j$ and $k$ are distinct (since $j > k$). Hence, Lemma 5.24.7 yields that

$$|Z| = |D\lfloor k\rfloor| + 1 \qquad \text{and} \qquad Z \setminus \{z\} = \operatorname{Row}(k,T)$$
$$\text{and} \qquad S_{n,Z} \cap \mathcal{R}(T) = S_{n,Z\setminus\{z\}}.$$

Also, obviously, $z \in \{z\} \subseteq \{z\} \cup \operatorname{Row}(k,T) = Z$.

Lemma 5.24.4 yields $D\lfloor j\rfloor = \left[\lambda_j\right]$ and $D\lfloor k\rfloor = [\lambda_k]$. But $\lambda$ is a partition, thus a weakly decreasing sequence. Hence, $\lambda_j \leq \lambda_k$ (since $j > k$). Hence, $\left[\lambda_j\right] \subseteq [\lambda_k]$. In other words, $D\lfloor j\rfloor \subseteq D\lfloor k\rfloor$ (since $D\lfloor j\rfloor = \left[\lambda_j\right]$ and $D\lfloor k\rfloor = [\lambda_k]$), so that

$$D\lfloor j\rfloor \cup D\lfloor k\rfloor = D\lfloor k\rfloor.$$

However,

$$|Z| = |D\lfloor k\rfloor| + 1 > \left|\underbrace{D\lfloor k\rfloor}_{=D\lfloor j\rfloor \cup D\lfloor k\rfloor}\right| = |D\lfloor j\rfloor \cup D\lfloor k\rfloor|.$$

Lemma 5.24.3 **(d)** (applied to $X = Z$ and $x = z$) shows that the set $\{\mathrm{id}\} \cup \left\{t_{y,z} \mid y \in Z \setminus \{z\}\right\}$ is a left transversal of $S_{n,Z\setminus\{z\}}$ in $S_{n,Z}$. In other words, the set $\{\mathrm{id}\} \cup \left\{t_{y,z} \mid y \in Z \setminus \{z\}\right\}$ is a left transversal of $S_{n,Z} \cap \mathcal{R}(T)$ in $S_{n,Z}$ (since $S_{n,Z} \cap \mathcal{R}(T) = S_{n,Z\setminus\{z\}}$). Hence, Corollary 5.23.10 (applied to $L = \{\mathrm{id}\} \cup \left\{t_{y,z} \mid y \in Z \setminus \{z\}\right\}$) shows that

$$\sum_{x \in \{\mathrm{id}\} \cup \left\{t_{y,z} \mid y \in Z\setminus\{z\}\right\}} x \nabla_{\operatorname{Row} T} \nabla^{-}_{\operatorname{Col} T} = 0. \tag{452}$$

But $D = Y(\lambda)$. Hence, the cells in the $k$-th row of $D$ are $(k,1), (k,2), \ldots, (k,\lambda_k)$. Thus, the entries in the $k$-th row of $T$ are $T(k,1), T(k,2), \ldots, T(k,\lambda_k)$. Now, recall that

$$Z \setminus \{z\} = \operatorname{Row}(k,T) = \{T(k,1), T(k,2), \ldots, T(k,\lambda_k)\}$$

(since the entries in the $k$-th row of $T$ are $T(k,1), T(k,2), \ldots, T(k,\lambda_k)$). Hence,

$$\begin{aligned}\left\{t_{y,z} \mid y \in Z \setminus \{z\}\right\} &= \left\{t_{y,z} \mid y \in \{T(k,1), T(k,2), \ldots, T(k,\lambda_k)\}\right\} \\ &= \left\{t_{T(k,1),\,z}, t_{T(k,2),\,z}, \ldots, t_{T(k,\lambda_k),\,z}\right\} \\ &= \left\{t_{T(k,1),\,T(j,q)}, t_{T(k,2),\,T(j,q)}, \ldots, t_{T(k,\lambda_k),\,T(j,q)}\right\}\end{aligned}$$

(since $z = T(j,q)$). Thus,

$$\begin{aligned}&\{\mathrm{id}\} \cup \left\{t_{y,z} \mid y \in Z \setminus \{z\}\right\} \\ &= \{\mathrm{id}\} \cup \left\{t_{T(k,1),\,T(j,q)}, t_{T(k,2),\,T(j,q)}, \ldots, t_{T(k,\lambda_k),\,T(j,q)}\right\} \\ &= \left\{\mathrm{id}, t_{T(k,1),\,T(j,q)}, t_{T(k,2),\,T(j,q)}, \ldots, t_{T(k,\lambda_k),\,T(j,q)}\right\}. \qquad (453)\end{aligned}$$

Moreover, all the $\lambda_k + 1$ permutations $\mathrm{id}, t_{T(k,1),\,T(j,q)}, t_{T(k,2),\,T(j,q)}, \ldots, t_{T(k,\lambda_k),\,T(j,q)}$ are distinct[154]. Combining this with (453), we see that the set $\{\mathrm{id}\} \cup \left\{t_{y,z} \mid y \in Z \setminus \{z\}\right\}$ consists of the $\lambda_k + 1$ distinct elements $\mathrm{id}, t_{T(k,1),\,T(j,q)}, t_{T(k,2),\,T(j,q)}, \ldots, t_{T(k,\lambda_k),\,T(j,q)}$. Hence, the sum of the elements of this set is

$$\begin{aligned}\sum_{x \in \{\mathrm{id}\} \cup \left\{t_{y,z} \mid y \in Z \setminus \{z\}\right\}} x &= \mathrm{id} + \underbrace{t_{T(k,1),\,T(j,q)} + t_{T(k,2),\,T(j,q)} + \cdots + t_{T(k,\lambda_k),\,T(j,q)}}_{= \sum_{s=1}^{\lambda_k} t_{T(k,s),\,T(j,q)}} \\ &= \mathrm{id} + \sum_{s=1}^{\lambda_k} t_{T(k,s),\,T(j,q)}.\end{aligned}$$

In other words,

$$\mathrm{id} + \sum_{s=1}^{\lambda_k} t_{T(k,s),\,T(j,q)} = \sum_{x \in \{\mathrm{id}\} \cup \left\{t_{y,z} \mid y \in Z \setminus \{z\}\right\}} x.$$

[154] *Proof.* The $\lambda_k$ entries $T(k,1), T(k,2), \ldots, T(k,\lambda_k)$ of $T$ are distinct (since all entries of $T$ are distinct). Hence, the $\lambda_k$ transpositions $t_{T(k,1),\,T(j,q)}, t_{T(k,2),\,T(j,q)}, \ldots, t_{T(k,\lambda_k),\,T(j,q)}$ are distinct (since they send $T(j,q)$ to the distinct entries $T(k,1), T(k,2), \ldots, T(k,\lambda_k)$). These $\lambda_k$ transpositions are furthermore all distinct from the identity permutation id (since any transposition is distinct from id). Hence, altogether, all the $\lambda_k + 1$ permutations $\mathrm{id}, t_{T(k,1),\,T(j,q)}, t_{T(k,2),\,T(j,q)}, \ldots, t_{T(k,\lambda_k),\,T(j,q)}$ are distinct.

Hence,

$$\left(1+\sum_{s=1}^{\lambda_k} t_{T(k,s),\,T(j,q)}\right)\nabla_{\operatorname{Row}T}\nabla^{-}_{\operatorname{Col}T} = \left(\sum_{x\in\{\mathrm{id}\}\cup\left\{t_{y,z}\;\mid\;y\in Z\setminus\{z\}\right\}} x\right)\nabla_{\operatorname{Row}T}\nabla^{-}_{\operatorname{Col}T}$$
$$= \sum_{x\in\{\mathrm{id}\}\cup\left\{t_{y,z}\;\mid\;y\in Z\setminus\{z\}\right\}} x\nabla_{\operatorname{Row}T}\nabla^{-}_{\operatorname{Col}T} = 0$$

(by (452)). This proves Proposition 5.24.1. □

### 5.24.2. Arms, legs and hooks

Next, we shall take a closer look at the hooks $H_\lambda(c)$ introduced in Theorem 5.2.8. We will subdivide them into a "long arm" and a "short leg":

**Proposition 5.24.8.** Let $\lambda$ be a partition of $n$. Let $c=(i,j)$ be a cell in $Y(\lambda)$. Define the *long arm* of $c$ in $\lambda$ to be the set

$$\operatorname{Arm}^{+}_{\lambda}(c) := \{c\}\cup\{\text{all cells in } Y(\lambda) \text{ that lie east of } c \text{ (in the same row as } c)\}.$$

Define the *short leg* of $c$ in $\lambda$ to be the set

$$\operatorname{Leg}^{-}_{\lambda}(c) := \{\text{all cells in } Y(\lambda) \text{ that lie south of } c \text{ (in the same column as } c)\}.$$

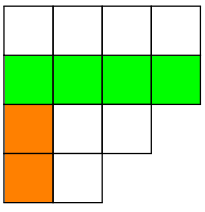
Also, define the *hook* $H_\lambda(c)$ of $c$ in $\lambda$ as in Theorem 5.2.8. Then:

**(a)** We have $\left|\operatorname{Arm}^{+}_{\lambda}(c)\right| = \lambda_i - j + 1$.

**(b)** We have $\left|\operatorname{Leg}^{-}_{\lambda}(c)\right| = \lambda^t_j - i$.

**(c)** We have $\left|\operatorname{Arm}^{+}_{\lambda}(c)\right| + \left|\operatorname{Leg}^{-}_{\lambda}(c)\right| = |H_\lambda(c)|$.

**Example 5.24.9.** Let $n=14$ and $\lambda=(5,4,3,2)$. Let $c$ be the cell $(2,1)\in Y(\lambda)$. Then, the long arm of $c$ consists of the green cells in the following picture, whereas the short leg consists of the orange cells:



.

In other words,

$$\operatorname{Arm}^{+}_{\lambda}(c) = \{(2,1),\,(2,2),\,(2,3),\,(2,4)\} \qquad \text{and}$$
$$\operatorname{Leg}^{-}_{\lambda}(c) = \{(1,3),\,(1,4)\}.$$

The picture in the above example should explain the words "arm" and "leg". The words "long" and "short" refer to the respective presence and absence of $c$. Of course, we could just as well define the "short arm" and the "long leg", but we will not need these.

The proof of Proposition 5.24.8 presents only typographical difficulties (due to the long equations):

*Proof of Proposition 5.24.8.* **(a)** The cells in the $i$-th row of $Y(\lambda)$ are $(i,1)$, $(i,2)$, $\ldots$, $(i,\lambda_i)$ (by the definition of $Y(\lambda)$). The definition of $\operatorname{Arm}_\lambda^+(c)$ yields

$$\begin{aligned}
\operatorname{Arm}_\lambda^+(c) &= \{c\} \cup \{\text{all cells in } Y(\lambda) \text{ that lie east of } c \text{ (in the same row as } c)\} \\
&= \{(i,j)\} \cup \underbrace{\{\text{all cells in } Y(\lambda) \text{ that lie east of } (i,j) \text{ (in the same row as } (i,j))\}}_{\substack{=\{\text{all cells in } Y(\lambda) \text{ that lie east of } (i,j) \text{ (in the } i\text{-th row)}\} \\ =\{\text{all cells in the } i\text{-th row of } Y(\lambda) \text{ that lie east of } (i,j)\} \\ =\{(i,j+1),\ (i,j+2),\ \ldots,\ (i,\lambda_i)\} \\ \text{(since the cells in the } i\text{-th row of } Y(\lambda) \text{ are } (i,1),\ (i,2),\ \ldots,\ (i,\lambda_i))}} \\
&\qquad\qquad (\text{since } c = (i,j)) \\
&= \{(i,j)\} \cup \{(i,j+1),\ (i,j+2),\ \ldots,\ (i,\lambda_i)\} \\
&= \{(i,j),\ (i,j+1),\ (i,j+2),\ \ldots,\ (i,\lambda_i)\}.
\end{aligned}$$

Hence,

$$\left|\operatorname{Arm}_\lambda^+(c)\right| = \left|\{(i,j),\ (i,j+1),\ (i,j+2),\ \ldots,\ (i,\lambda_i)\}\right| = \lambda_i - j + 1$$

(since the $\lambda_i - j + 1$ cells $(i,j)$, $(i,j+1)$, $(i,j+2)$, $\ldots$, $(i,\lambda_i)$ are obviously distinct). This proves Proposition 5.24.8 **(a)**.

**(b)** The cells in the $j$-th column of $Y(\lambda)$ are $(1,j)$, $(2,j)$, $\ldots$, $\left(\lambda_j^t, j\right)$ (by Theorem 5.1.10 **(e)**). The definition of $\operatorname{Leg}_\lambda^-(c)$ yields

$$\begin{aligned}
\operatorname{Leg}_\lambda^-(c) &= \{\text{all cells in } Y(\lambda) \text{ that lie south of } c \text{ (in the same column as } c)\} \\
&= \{\text{all cells in } Y(\lambda) \text{ that lie south of } (i,j) \text{ (in the same column as } (i,j))\} \\
&\qquad\qquad (\text{since } c = (i,j)) \\
&= \{\text{all cells in } Y(\lambda) \text{ that lie south of } (i,j) \text{ (in the } j\text{-th column)}\} \\
&= \{\text{all cells in the } j\text{-th column of } Y(\lambda) \text{ that lie south of } (i,j)\} \\
&= \left\{(i+1,j),\ (i+2,j),\ \ldots,\ \left(\lambda_j^t, j\right)\right\}
\end{aligned}$$

(since the cells in the $j$-th column of $Y(\lambda)$ are $(1,j)$, $(2,j)$, $\ldots$, $\left(\lambda_j^t, j\right)$). Hence,

$$\left|\operatorname{Leg}_\lambda^-(c)\right| = \left|\left\{(i+1,j),\ (i+2,j),\ \ldots,\ \left(\lambda_j^t, j\right)\right\}\right| = \lambda_j^t - i$$

(since the $\lambda_j^t - i$ cells $(i+1,j)$, $(i+2,j)$, $\ldots$, $\left(\lambda_j^t, j\right)$ are obviously distinct). This proves Proposition 5.24.8 **(b)**.

**(c)** The set $\operatorname{Arm}_\lambda^+(c)$ consists entirely of cells that lie in the same row as $c$ (by its definition), whereas the set $\operatorname{Leg}_\lambda^-(c)$ consists entirely of cells that lie in rows further

south than $c$ (again by its definition). Thus, there is no cell that belongs to both of these two sets. In other words, the sets $\operatorname{Arm}_\lambda^+(c)$ and $\operatorname{Leg}_\lambda^-(c)$ are disjoint.

However, the definition of $H_\lambda(c)$ in Theorem 5.2.8 says that

$$\begin{aligned} H_\lambda(c) &= \underbrace{\{c\} \cup \{\text{all cells in } Y(\lambda) \text{ that lie east of } c \text{ (in the same row as } c)\}}_{\substack{=\operatorname{Arm}_\lambda^+(c) \\ \text{(by the definition of } \operatorname{Arm}_\lambda^+(c))}} \\ &\qquad \cup \underbrace{\{\text{all cells in } Y(\lambda) \text{ that lie south of } c \text{ (in the same column as } c)\}}_{\substack{=\operatorname{Leg}_\lambda^-(c) \\ \text{(by the definition of } \operatorname{Leg}_\lambda^-(c))}} \\ &= \operatorname{Arm}_\lambda^+(c) \cup \operatorname{Leg}_\lambda^-(c). \end{aligned}$$

Thus,

$$|H_\lambda(c)| = \left|\operatorname{Arm}_\lambda^+(c) \cup \operatorname{Leg}_\lambda^-(c)\right| = \left|\operatorname{Arm}_\lambda^+(c)\right| + \left|\operatorname{Leg}_\lambda^-(c)\right|$$

(since the sets $\operatorname{Arm}_\lambda^+(c)$ and $\operatorname{Leg}_\lambda^-(c)$ are disjoint). This proves Proposition 5.24.8 **(c)**. □

The following property of long arms will also be useful:

**Proposition 5.24.10.** Let $\lambda$ be a partition of $n$. For any cell $c \in Y(\lambda)$, define the long arm $\operatorname{Arm}_\lambda^+(c)$ as in Proposition 5.24.8. Then:

**(a)** If $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ for some $\ell \in \mathbb{N}$, then

$$\prod_{c \in Y(\lambda)} \left|\operatorname{Arm}_\lambda^+(c)\right| = \prod_{i=1}^{\ell} \lambda_i!.$$

**(b)** If $T$ is any $n$-tableau of shape $\lambda$, then

$$\prod_{c \in Y(\lambda)} \left|\operatorname{Arm}_\lambda^+(c)\right| = |\mathcal{R}(T)|.$$

*Proof.* **(a)** Assume that $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ for some $\ell \in \mathbb{N}$. Then, the elements of the Young diagram $Y(\lambda)$ are precisely the pairs $(i,j)$ with $i \in [\ell]$ and $j \in [\lambda_i]$. Hence, the product sign $\prod\limits_{(i,j) \in Y(\lambda)}$ can be rewritten as $\prod\limits_{i \in [\ell]} \prod\limits_{j \in [\lambda_i]}$.

However,

$$\begin{aligned} \prod_{c \in Y(\lambda)} \left|\operatorname{Arm}_\lambda^+(c)\right| &= \prod_{(i,j) \in Y(\lambda)} \underbrace{\left|\operatorname{Arm}_\lambda^+(i,j)\right|}_{\substack{=\lambda_i - j + 1 \\ \text{(by Proposition 5.24.8 (a),} \\ \text{applied to } c=(i,j))}} \qquad \left(\begin{array}{c}\text{here, we have renamed} \\ \text{the index } c \text{ as } (i,j)\end{array}\right) \\ &= \prod_{(i,j) \in Y(\lambda)} (\lambda_i - j + 1) \\ &= \prod_{i \in [\ell]} \prod_{j \in [\lambda_i]} (\lambda_i - j + 1) \end{aligned} \tag{454}$$

(since the product sign $\prod_{(i,j)\in Y(\lambda)}$ can be rewritten as $\prod_{i\in[\ell]}\prod_{j\in[\lambda_i]}$). However, for each $i\in[\ell]$, we have

$$\underbrace{\prod_{j\in[\lambda_i]}}_{=\prod_{j=1}^{\lambda_i}}\underbrace{(\lambda_i-j+1)}_{=\lambda_i+1-j}=\prod_{j=1}^{\lambda_i}(\lambda_i+1-j)=\prod_{k=1}^{\lambda_i}k \qquad \left(\begin{array}{c}\text{here, we have substituted } k\\ \text{for } \lambda_i+1-j \text{ in the product}\end{array}\right)$$

$$=1\cdot 2\cdot\cdots\cdot\lambda_i=\lambda_i!.$$

Thus, we can rewrite (454) as

$$\prod_{c\in Y(\lambda)}\left|\mathrm{Arm}_\lambda^+(c)\right|=\underbrace{\prod_{i\in[\ell]}}_{=\prod_{i=1}^{\ell}}\lambda_i!=\prod_{i=1}^{\ell}\lambda_i!.$$

This proves Proposition 5.24.10 **(a)**.

**(b)** Let $D$ be the diagram $Y(\lambda)$. Let $T$ be any $n$-tableau of shape $\lambda$. Thus, $T$ is an $n$-tableau of shape $Y(\lambda)=D$.

For any $i\in\mathbb{Z}$, let $\mathrm{Row}(i,T)$ denote the set of all entries in the $i$-th row of $T$. Then, for each positive integer $i$, we have

$$\begin{aligned}|\mathrm{Row}(i,T)| &= |D\lfloor i\rfloor| && \text{(by Proposition 5.24.5, applied to } j=i)\\ &= |[\lambda_i]| && \text{(since Lemma 5.24.4 yields } D\lfloor i\rfloor=[\lambda_i])\\ &= \lambda_i. && (455)\end{aligned}$$

Let $\ell$ be the length of $\lambda$. Thus, $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_\ell)$. Hence, all cells of the Young diagram $Y(\lambda)$ lie in rows $1,2,\ldots,\ell$. Therefore, Proposition 5.5.7 **(a)** (applied to $D=Y(\lambda)$) yields

$$|\mathcal{R}(T)|=\prod_{i=1}^{\ell}\underbrace{|\mathrm{Row}(i,T)|}_{\substack{=\lambda_i\\ \text{(by (455))}}}!=\prod_{i=1}^{\ell}\lambda_i!=\prod_{c\in Y(\lambda)}\left|\mathrm{Arm}_\lambda^+(c)\right|$$

(by Proposition 5.24.10 **(a)**). This proves Proposition 5.24.10 **(b)**. □

### 5.24.3. A few last lemmas

To simplify the proof of Theorem 5.2.8, we outsource a few more arguments to lemmas. The following two lemmas are simple properties of row symmetrizers:

**Lemma 5.24.11.** Let $T$ be an $n$-tableau of any shape $D$. Let $p,q,r$ be three distinct elements of $[n]$. Assume that $p$ and $r$ lie in the same row of $T$. Then,

$$\left(1-t_{p,q}\right)t_{q,r}\nabla_{\mathrm{Row}\,T}=0.$$

*Proof.* We know that $p$ and $r$ lie in the same row of $T$. Thus, Proposition 5.5.8 **(c)** (applied to $i = p$ and $j = r$) shows that

$$t_{p,r} \nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T} t_{p,r} = \nabla_{\operatorname{Row} T}.$$

On the other hand, we have $p \notin \{q, r\}$ (since $p, q, r$ are distinct) and thus $t_{q,r}(p) = p$. Applying (5) to $\sigma = t_{q,r}$ and $i = p$ and $j = r$, we find

$$t_{q,r} \circ t_{p,r} \circ t_{q,r}^{-1} = t_{t_{q,r}(p), t_{q,r}(r)} = t_{p,q} \qquad \left(\text{since } t_{q,r}(p) = p \text{ and } t_{q,r}(r) = q\right).$$

Thus, $t_{p,q} = t_{q,r} \circ t_{p,r} \circ t_{q,r}^{-1} = t_{q,r} t_{p,r} t_{q,r}^{-1}$. Multiplying this equality by $t_{q,r}$ from the right, we obtain

$$t_{p,q} t_{q,r} = t_{q,r} t_{p,r} \underbrace{t_{q,r}^{-1} t_{q,r}}_{=1} = t_{q,r} t_{p,r}.$$

Now,

$$\left(1 - t_{p,q}\right) t_{q,r} = t_{q,r} - \underbrace{t_{p,q} t_{q,r}}_{= t_{q,r} t_{p,r}} = t_{q,r} - t_{q,r} t_{p,r} = t_{q,r}\left(1 - t_{p,r}\right).$$

Hence,

$$\underbrace{\left(1 - t_{p,q}\right) t_{q,r}}_{= t_{q,r}\left(1 - t_{p,r}\right)} \nabla_{\operatorname{Row} T} = t_{q,r} \underbrace{\left(1 - t_{p,r}\right) \nabla_{\operatorname{Row} T}}_{\substack{= \nabla_{\operatorname{Row} T} - t_{p,r} \nabla_{\operatorname{Row} T} \\ = 0 \\ \left(\text{since } t_{p,r} \nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T}\right)}} = 0.$$

This proves Lemma 5.24.11. □

**Lemma 5.24.12.** Let $T$ be an $n$-tableau of any shape $D$. Let $p, q, r$ be three distinct elements of $[n]$. Assume that $p$ and $r$ lie in the same row of $T$. Let $\mathbf{a} \in \mathcal{A}$ be an element that commutes with $t_{p,r}$. Then,

$$\nabla_{\operatorname{Row} T} \mathbf{a} t_{r,q} \nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T} \mathbf{a} t_{p,q} \nabla_{\operatorname{Row} T}.$$

*Proof.* As in the proof of Lemma 5.24.11, we can see that

$$t_{p,r} \nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T} t_{p,r} = \nabla_{\operatorname{Row} T}. \tag{456}$$

Moreover, (6) (applied to $k = 3$ and $(i_1, i_2, \ldots, i_k) = (p, r, q)$) yields $\operatorname{cyc}_{p,r,q} = t_{p,r} \circ t_{r,q}$. Meanwhile, (6) (applied to $k = 3$ and $(i_1, i_2, \ldots, i_k) = (q, p, r)$) yields $\operatorname{cyc}_{q,p,r} = t_{q,p} \circ t_{p,r}$. But we have $\operatorname{cyc}_{p,r,q} = \operatorname{cyc}_{q,p,r}$. Thus,

$$t_{p,r} \circ t_{r,q} = \operatorname{cyc}_{p,r,q} = \operatorname{cyc}_{q,p,r} = t_{q,p} \circ t_{p,r}.$$

Now,

$$\underbrace{\nabla_{\operatorname{Row} T}}_{\substack{=\nabla_{\operatorname{Row} T} t_{p,r} \\ \text{(by (456))}}} \mathbf{a} t_{r,q} \nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T} \underbrace{t_{p,r}\mathbf{a}}_{\substack{=\mathbf{a}t_{p,r} \\ \text{(since } \mathbf{a} \text{ commutes} \\ \text{with } t_{p,r}\text{)}}} t_{r,q} \nabla_{\operatorname{Row} T}$$

$$= \nabla_{\operatorname{Row} T} \mathbf{a} \underbrace{t_{p,r} t_{r,q}}_{\substack{=t_{p,r} \circ t_{r,q} \\ =t_{q,p} \circ t_{p,r} \\ =t_{q,p} t_{p,r}}} \nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T} \mathbf{a} \underbrace{t_{q,p}}_{=t_{p,q}} \underbrace{t_{p,r} \nabla_{\operatorname{Row} T}}_{\substack{=\nabla_{\operatorname{Row} T} \\ \text{(by (456))}}}$$

$$= \nabla_{\operatorname{Row} T} \mathbf{a} t_{p,q} \nabla_{\operatorname{Row} T}.$$

This proves Lemma 5.24.12. □

The following lemma is again an easy consequence of things proved before. Its main purpose is to isolate a simple argument from the proof of the hook length formula:

**Lemma 5.24.13.** Let $\mathbf{k} = \mathbb{Q}$. Let $\lambda$ be a partition of $n$. Let $T$ be an $n$-tableau of shape $Y(\lambda)$. Set $P := \nabla_{\operatorname{Row} T}$ and $N := \nabla^{-}_{\operatorname{Col} T}$. Let $\kappa \in \mathbb{Q}$ be such that $(PN)^2 = \kappa \cdot PN$. Then, the # of standard tableaux of shape $Y(\lambda)$ is $\dfrac{n!}{\kappa}$.

*Proof.* Multiplying the equalities $N = \nabla^{-}_{\operatorname{Col} T}$ and $P = \nabla_{\operatorname{Row} T}$, we find

$$NP = \nabla^{-}_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} = \mathbf{E}_T \qquad \text{(by Definition 5.11.1 \textbf{(b)})}.$$

Lemma 5.11.13 yields $[1]\,\mathbf{E}_T = 1$ (see Definition 5.11.6 for the meaning of $[1]\,\mathbf{E}_T$).

Let $f^\lambda$ be the # of standard tableaux of shape $Y(\lambda)$. Then, Theorem 5.11.3 yields

$$\mathbf{E}_T^2 = \frac{n!}{f^\lambda} \mathbf{E}_T. \tag{457}$$

Recall the antipode $S$ of $\mathbf{k}[S_n]$ (as defined in Definition 3.11.12). Then, Proposition 5.5.10 says that $S(\nabla_{\operatorname{Row} T}) = \nabla_{\operatorname{Row} T}$ and $S\left(\nabla^{-}_{\operatorname{Col} T}\right) = \nabla^{-}_{\operatorname{Col} T}$. In view of $P = \nabla_{\operatorname{Row} T}$ and $N = \nabla^{-}_{\operatorname{Col} T}$, we can rewrite these two equalities as $S(P) = P$ and $S(N) = N$.

We have $PNPN = (PN)^2 = \kappa \cdot PN$. Applying the map $S$ to both sides of this equality, we find

$$S(PNPN) = S(\kappa \cdot PN) = \kappa \cdot S(PN) \tag{458}$$

(since the map $S$ is $\mathbf{k}$-linear).

But $S$ is a $\mathbf{k}$-algebra anti-automorphism (by Theorem 3.11.14 **(a)**). Thus,

$$S(PN) = \underbrace{S(N)}_{=N} \cdot \underbrace{S(P)}_{=P} = NP = \mathbf{E}_T$$

and

$$S(PNPN) = \underbrace{S(N)}_{=N} \cdot \underbrace{S(P)}_{=P} \cdot \underbrace{S(N)}_{=N} \cdot \underbrace{S(P)}_{=P} = \underbrace{NP}_{=\mathbf{E}_T} \underbrace{NP}_{=\mathbf{E}_T} = \mathbf{E}_T \mathbf{E}_T = \mathbf{E}_T^2 = \frac{n!}{f^\lambda} \mathbf{E}_T$$

(by (457)). Using these two equalities, we can rewrite (458) as

$$\frac{n!}{f^\lambda} \mathbf{E}_T = \kappa \cdot \mathbf{E}_T.$$

Thus,

$$[1] \left( \frac{n!}{f^\lambda} \mathbf{E}_T \right) = [1] (\kappa \cdot \mathbf{E}_T) = \kappa \cdot \underbrace{[1] \mathbf{E}_T}_{=1} = \kappa.$$

Hence,

$$\kappa = [1] \left( \frac{n!}{f^\lambda} \mathbf{E}_T \right) = \frac{n!}{f^\lambda} \underbrace{[1] \mathbf{E}_T}_{=1} = \frac{n!}{f^\lambda}.$$

Hence, $f^\lambda = \dfrac{n!}{\kappa}$. In other words, the # of standard tableaux of shape $Y(\lambda)$ is $\dfrac{n!}{\kappa}$ (since $f^\lambda$ is the # of standard tableaux of shape $Y(\lambda)$). This proves Lemma 5.24.13. □

### 5.24.4. Proof of the hook length formula

We now have all the tools in place to prove the hook length formula:

*Proof of Theorem 5.2.8.* Let $\mathbf{k} = \mathbb{Q}$.

We have $|Y(\lambda)| = |\lambda| = n$. In other words, the diagram $Y(\lambda)$ consists of $n$ cells. Let us list these $n$ cells row by row, from the bottommost row to the topmost row, where the cells of each row are listed from left to right:

$$\begin{aligned} &(\ell, 1), (\ell, 2), \ldots, (\ell, \lambda_\ell), \\ &(\ell - 1, 1), (\ell - 1, 2), \ldots, (\ell - 1, \lambda_{\ell-1}), \\ &\ldots, \\ &(1, 1), (1, 2), \ldots, (1, \lambda_1) \end{aligned} \tag{459}$$

(where $\ell$ is the length of $\lambda$). Now, we define an $n$-tableau $T$ of shape $Y(\lambda)$ by filling these $n$ cells (in the order in which we just listed them) with the entries $1, 2, \ldots, n$ (respectively).

**Example 5.24.14.** If $n = 14$ and $\lambda = (5, 4, 3, 2)$, then

$$T = \begin{array}{|c|c|c|c|c|} \hline 10 & 11 & 12 & 13 & 14 \\ \hline 6 & 7 & 8 & 9 \\ \cline{1-4} 3 & 4 & 5 \\ \cline{1-3} 1 & 2 \\ \cline{1-2} \end{array} .$$

Thus we have defined an $n$-tableau $T$ of shape $Y(\lambda)$, that is, an $n$-tableau $T$ of shape $\lambda$. In particular, all entries of $T$ are distinct (since $T$ is an $n$-tableau, thus injective).

Let $P = \nabla_{\operatorname{Row} T}$ and $N = \nabla^-_{\operatorname{Col} T}$. Then, Corollary 5.5.6 **(a)** says that $\left(\nabla_{\operatorname{Row} T}\right)^2 = |\mathcal{R}(T)| \cdot \nabla_{\operatorname{Row} T}$. In other words, $P^2 = |\mathcal{R}(T)| \cdot P$ (since $P = \nabla_{\operatorname{Row} T}$).

For any $j \in \mathbb{Z}$, let $\operatorname{Col}(j, T)$ denote the set of all entries in the $j$-th column of $T$. Choose some $k \in \mathbb{N}$ such that all cells of $Y(\lambda)$ lie in columns $1, 2, \ldots, k$. (Such a $k$ clearly exists, since $Y(\lambda)$ is a Young diagram; in fact we can take $k = \lambda_1$.)

Proposition 5.5.3 **(b)** (applied to $D = Y(\lambda)$) shows that

$$\nabla^-_{\operatorname{Col} T} = \prod_{j=1}^{k} \nabla^-_{\operatorname{Col}(j,T)}. \tag{460}$$

Here, we are not specifying the order of the factors in the product, since all the factors commute.

Now, for each $m \in \{0, 1, \ldots, n\}$, we define an element $N_m \in \mathcal{A}$ by

$$N_m := \prod_{j=1}^{k} \nabla^-_{[m] \cap \operatorname{Col}(j,T)}. \tag{461}$$

Here, again, we are not specifying the order of the factors in the product, since all the factors commute[155].

[155] *Proof.* Let $m \in \{0, 1, \ldots, n\}$. We must prove that all factors in the product on the right hand side of (461) commute. In other words, we must prove that the elements $\nabla^-_{[m] \cap \operatorname{Col}(j,T)}$ for all $j \in [k]$ commute. In other words, we must prove that $\nabla^-_{[m] \cap \operatorname{Col}(u,T)}$ commutes with $\nabla^-_{[m] \cap \operatorname{Col}(v,T)}$ for any two elements $u, v \in [k]$. Let us prove this.

Let $u, v \in [k]$ be two elements. We must prove that $\nabla^-_{[m] \cap \operatorname{Col}(u,T)}$ commutes with $\nabla^-_{[m] \cap \operatorname{Col}(v,T)}$. If $u = v$, then this is obvious. Thus, we WLOG assume that $u \neq v$.

Recall that all entries of $T$ are distinct. Thus, in particular, an entry in the $u$-th column of $T$ cannot also appear in the $v$-th column of $T$ (since $u \neq v$). In other words, a number cannot belong to $\operatorname{Col}(u, T)$ and to $\operatorname{Col}(v, T)$ at the same time. In other words, $\operatorname{Col}(u, T) \cap \operatorname{Col}(v, T) = \varnothing$. Hence,

$$\begin{aligned}([m] \cap \operatorname{Col}(u, T)) \cap ([m] \cap \operatorname{Col}(v, T)) &= [m] \cap [m] \cap \underbrace{\operatorname{Col}(u, T) \cap \operatorname{Col}(v, T)}_{=\varnothing} \\ &= [m] \cap [m] \cap \varnothing = \varnothing.\end{aligned}$$

Thus, the two subsets $[m] \cap \operatorname{Col}(u, T)$ and $[m] \cap \operatorname{Col}(v, T)$ of $[n]$ are disjoint. Hence, Propo-

**Example 5.24.15.** If $n$, $\lambda$ and $T$ are as in Example 5.24.14, then we can pick $k = 5$ (since all cells of $Y(\lambda)$ lie in columns $1, 2, 3, 4, 5$) and obtain

$$\begin{aligned} N_7 &= \prod_{j=1}^{k} \nabla^-_{[7]\cap \mathrm{Col}(j,T)} \\ &= \underbrace{\nabla^-_{[7]\cap \mathrm{Col}(1,T)}}_{\substack{=\nabla^-_{[7]\cap\{10,6,3,1\}} \\ =\nabla^-_{\{6,3,1\}}}} \underbrace{\nabla^-_{[7]\cap \mathrm{Col}(2,T)}}_{\substack{=\nabla^-_{[7]\cap\{11,7,4,2\}} \\ =\nabla^-_{\{7,4,2\}}}} \underbrace{\nabla^-_{[7]\cap \mathrm{Col}(3,T)}}_{\substack{=\nabla^-_{[7]\cap\{12,8,5\}} \\ =\nabla^-_{\{5\}}=1}} \underbrace{\nabla^-_{[7]\cap \mathrm{Col}(4,T)}}_{\substack{=\nabla^-_{[7]\cap\{13,9\}} \\ =\nabla^-_{\varnothing}=1}} \underbrace{\nabla^-_{[7]\cap \mathrm{Col}(5,T)}}_{\substack{=\nabla^-_{[7]\cap\{14\}} \\ =\nabla^-_{\varnothing}=1}} \\ &= \nabla^-_{\{6,3,1\}} \nabla^-_{\{7,4,2\}}. \end{aligned}$$

Note that each $j \in [k]$ satisfies $[n] \cap \mathrm{Col}(j,T) = \mathrm{Col}(j,T)$ (since $\mathrm{Col}(j,T) \subseteq [n]$) and thus

$$\nabla^-_{[n]\cap \mathrm{Col}(j,T)} = \nabla^-_{\mathrm{Col}(j,T)}. \tag{462}$$

The definition of $N_n$ yields

$$\begin{aligned} N_n &= \prod_{j=1}^{k} \underbrace{\nabla^-_{[n]\cap \mathrm{Col}(j,T)}}_{\substack{=\nabla^-_{\mathrm{Col}(j,T)} \\ \text{(by (462))}}} = \prod_{j=1}^{k} \nabla^-_{\mathrm{Col}(j,T)} = \nabla^-_{\mathrm{Col}\,T} \qquad \text{(by (460))} \\ &= N \qquad \left(\text{since } N = \nabla^-_{\mathrm{Col}\,T}\right). \end{aligned} \tag{463}$$

Furthermore, each $j \in [k]$ satisfies $\underbrace{[0]}_{=\varnothing} \cap \mathrm{Col}(j,T) = \varnothing \cap \mathrm{Col}(j,T) = \varnothing$ and thus

$$\nabla^-_{[0]\cap \mathrm{Col}(j,T)} = \nabla^-_{\varnothing} = 1 \tag{464}$$

(by Example 3.7.2 **(a)**). The definition of $N_0$ yields

$$N_0 = \prod_{j=1}^{k} \underbrace{\nabla^-_{[0]\cap \mathrm{Col}(j,T)}}_{\substack{=1 \\ \text{(by (464))}}} = \prod_{j=1}^{k} 1 = 1. \tag{465}$$

---

sition 3.7.12 **(b)** (applied to $X = [m] \cap \mathrm{Col}(u,T)$ and $Y = [m] \cap \mathrm{Col}(v,T)$) yields that

$$\begin{aligned} \nabla_{[m]\cap \mathrm{Col}(u,T)} \nabla_{[m]\cap \mathrm{Col}(v,T)} &= \nabla_{[m]\cap \mathrm{Col}(v,T)} \nabla_{[m]\cap \mathrm{Col}(u,T)}; \\ \nabla_{[m]\cap \mathrm{Col}(u,T)} \nabla^-_{[m]\cap \mathrm{Col}(v,T)} &= \nabla^-_{[m]\cap \mathrm{Col}(v,T)} \nabla_{[m]\cap \mathrm{Col}(u,T)}; \\ \nabla^-_{[m]\cap \mathrm{Col}(u,T)} \nabla_{[m]\cap \mathrm{Col}(v,T)} &= \nabla_{[m]\cap \mathrm{Col}(v,T)} \nabla^-_{[m]\cap \mathrm{Col}(u,T)}; \\ \nabla^-_{[m]\cap \mathrm{Col}(u,T)} \nabla^-_{[m]\cap \mathrm{Col}(v,T)} &= \nabla^-_{[m]\cap \mathrm{Col}(v,T)} \nabla^-_{[m]\cap \mathrm{Col}(u,T)}. \end{aligned}$$

The last of these four equalities shows that $\nabla^-_{[m]\cap \mathrm{Col}(u,T)}$ commutes with $\nabla^-_{[m]\cap \mathrm{Col}(v,T)}$. This completes our proof.

**Remark 5.24.16.** Here is another way to view $N_m$: Let $m \in \{0,1,\ldots,n\}$. Let $T_m$ be the $m$-tableau that is obtained from $T$ by removing the cells filled with $m+1, m+2, \ldots, n$ (thus leaving only the cells filled with $1,2,\ldots,m$ in the tableau). For instance, if $n$, $\lambda$ and $T$ are as in Example 5.24.14, then

$$T_7 = \begin{array}{|c|c|c|}\hline 6 & 7 & \multicolumn{1}{c}{} \\ \hline 3 & 4 & 5 \\ \hline 1 & 2 & \multicolumn{1}{c}{} \\ \cline{1-2}\end{array}\ .$$

Now, $N_m$ is the column antisymmetrizer $\nabla^-_{\mathrm{Col}(T_m)} \in \mathbf{k}[S_m]$, mapped into $\mathbf{k}[S_n]$ using the default embedding $(I_{m\to n})_* : \mathbf{k}[S_m] \to \mathbf{k}[S_n]$ (see Corollary 3.12.13).

We shall use the notations $\mathrm{Arm}^+_\lambda(c)$ and $\mathrm{Leg}^-_\lambda(c)$ as defined in Proposition 5.24.8.

Furthermore, for each $m \in \{0,1,\ldots,n\}$, we define a rational number[156]

$$a_m := \prod_{c \in T^{-1}([m])} \frac{|H_\lambda(c)|}{\left|\mathrm{Arm}^+_\lambda(c)\right|}. \tag{466}$$

Thus,

$$\begin{aligned} a_0 &= \prod_{c \in T^{-1}([0])} \frac{|H_\lambda(c)|}{\left|\mathrm{Arm}^+_\lambda(c)\right|} = (\text{empty product}) \qquad \left(\text{since } T^{-1}([0]) = \varnothing\right) \\ &= 1 \end{aligned}$$

and

$$\begin{aligned} a_n &= \prod_{c \in T^{-1}([n])} \frac{|H_\lambda(c)|}{\left|\mathrm{Arm}^+_\lambda(c)\right|} = \prod_{c \in Y(\lambda)} \frac{|H_\lambda(c)|}{\left|\mathrm{Arm}^+_\lambda(c)\right|} \qquad \left(\text{since } T^{-1}([n]) = Y(\lambda)\right) \\ &= \frac{\prod\limits_{c \in Y(\lambda)} |H_\lambda(c)|}{\prod\limits_{c \in Y(\lambda)} \left|\mathrm{Arm}^+_\lambda(c)\right|} = \frac{\prod\limits_{c \in Y(\lambda)} |H_\lambda(c)|}{|\mathcal{R}(T)|} \end{aligned}$$

(since Proposition 5.24.10 **(b)** yields $\prod\limits_{c \in Y(\lambda)} \left|\mathrm{Arm}^+_\lambda(c)\right| = |\mathcal{R}(T)|$).

Now, I claim that each $m \in \{0,1,\ldots,n\}$ satisfies

$$PN_mPN = a_m P^2 N. \tag{467}$$

Before we come to the proof of this claim, let me explain why it yields Theorem 5.2.8. Indeed, assume (for a moment) that (467) is proved. Then, we can

[156] As usual, $T^{-1}([m])$ denotes the set of all $c \in Y(\lambda)$ that satisfy $T(c) \in [m]$.

apply (467) to $m = n$. As a result, we obtain $PN_nPN = a_nP^2N$. Comparing this with $P \underbrace{N_n}_{=N} PN = PNPN = (PN)^2$, we obtain

$$(PN)^2 = \underbrace{\frac{\prod_{c\in Y(\lambda)} |H_\lambda(c)|}{|\mathcal{R}(T)|}}_{=a_n} \underbrace{P^2}_{=|\mathcal{R}(T)|\cdot P} N = \underbrace{\frac{\prod_{c\in Y(\lambda)} |H_\lambda(c)|}{|\mathcal{R}(T)|} \cdot |\mathcal{R}(T)|}_{=\prod_{c\in Y(\lambda)} |H_\lambda(c)|} \cdot PN$$
$$= \left( \prod_{c\in Y(\lambda)} |H_\lambda(c)| \right) \cdot PN.$$

Hence, Lemma 5.24.13 (applied to $\kappa = \prod_{c\in Y(\lambda)} |H_\lambda(c)|$) yields that the # of standard tableaux of shape $Y(\lambda)$ is $\dfrac{n!}{\prod_{c\in Y(\lambda)} |H_\lambda(c)|}$. Thus, Theorem 5.2.8 is proved, provided that we can show (467).

It thus remains to prove (467). We shall do so by induction on $m$:

*Base case:* Recall that $a_0 = 1$ and $N_0 = 1$. Thus, $P \underbrace{N_0}_{=1} PN = PPN = P^2N$ and $\underbrace{a_0}_{=1} P^2N = P^2N$. Comparing these two equalities, we obtain $PN_0PN = a_0P^2N$. In other words, (467) holds for $m = 0$. This completes the base case.

*Induction step:* Let $m \in [n]$. Assume (as the induction hypothesis) that (467) holds for $m-1$ instead of $m$. We must prove that (467) also holds for $m$.

Let $c = (i, j)$ be the cell of $T$ that contains the entry $m$. Thus, $T(c) = m$. In other words, $T(i, j) = m$ (since $c = (i, j)$).

**Example 5.24.17.** For instance, let $n$, $\lambda$ and $T$ be as in Example 5.24.14, and let $m = 7$. Then, we have

$$T = \begin{array}{|c|c|c|c|c|} \hline 10 & 11 & 12 & 13 & 14 \\ \hline 6 & 7 & 8 & 9 \\ \cline{1-4} 3 & 4 & 5 \\ \cline{1-3} 1 & 2 \\ \cline{1-2} \end{array},$$

where the cells with entries $1, 2, \ldots, m-1$ are colored green, while the cell with entry $m$ is colored red. The element $N_{m-1}$ is entirely determined by the entries in green cells (see Remark 5.24.16, which writes $N_{m-1}$ as the column antisymmetrizer $\nabla^-_{\mathrm{Col}(T_{m-1})}$ of the green subtableau $T_{m-1}$), whereas the element $N_m$ is determined by the entries in green and red cells.

Theorem 5.1.10 **(e)** shows that the cells in the $j$-th column of $Y(\lambda)$ are

$$(1,j),\ (2,j),\ \ldots,\ \left(\lambda_j^t, j\right).$$

Hence, the entries in the $j$-th column of $T$ are $T(1,j),\ T(2,j),\ \ldots,\ T\left(\lambda_j^t, j\right)$. In other words,

$$\operatorname{Col}(j,T) = \left\{T(1,j),\ T(2,j),\ \ldots,\ T\left(\lambda_j^t, j\right)\right\} \tag{468}$$

(since $\operatorname{Col}(j,T)$ is defined as the set of all entries in the $j$-th column of $T$). Thus, $T(\ell, j)$ is well-defined for each $\ell \in \left[\lambda_j^t\right]$.

We note that $j \in [k]$ [157]. Moreover, $m$ is an entry in the $j$-th column of $T$ (since $m = T(i,j)$). In other words, $m \in \operatorname{Col}(j,T)$.

Now we shall show a sequence of claims. The first three claims are obvious from a quick look at the construction of $T$ and the fact that $T(i,j) = m$, but we give detailed proofs just for the sake of completeness.

*Claim 1:* Let $v$ be a positive integer such that $v \neq j$. Then,

$$[m] \cap \operatorname{Col}(v,T) = [m-1] \cap \operatorname{Col}(v,T).$$

*Proof of Claim 1.* Recall that all entries of $T$ are distinct. Hence, an entry in the $j$-th column of $T$ cannot also be found in the $v$-th column of $T$ (since $v \neq j$). In particular, $m$ cannot also be found in the $v$-th column of $T$ (since $m$ is an entry in the $j$-th column of $T$). In other words, $m \notin \operatorname{Col}(v,T)$. But $[m] = [m-1] \cup \{m\}$, and thus

$$\begin{aligned}
\underbrace{[m]}_{=[m-1]\cup\{m\}} \cap \operatorname{Col}(v,T) &= ([m-1] \cup \{m\}) \cap \operatorname{Col}(v,T) \\
&= ([m-1] \cap \operatorname{Col}(v,T)) \cup \underbrace{(\{m\} \cap \operatorname{Col}(v,T))}_{\substack{=\varnothing \\ (\text{since } m \notin \operatorname{Col}(v,T))}} \\
&= ([m-1] \cap \operatorname{Col}(v,T)) \cup \varnothing = [m-1] \cap \operatorname{Col}(v,T).
\end{aligned}$$

This proves Claim 1. □

*Claim 2:* We have $m \in [m] \cap \operatorname{Col}(j,T)$ and

$$[m-1] \cap \operatorname{Col}(j,T) = ([m] \cap \operatorname{Col}(j,T)) \setminus \{m\}.$$

[157] *Proof.* We know that $(i,j)$ is a cell of $Y(\lambda)$ (by its definition). Thus, $(i,j)$ lies in one of the columns $1,2,\ldots,k$ (since all cells of $Y(\lambda)$ lie in columns $1,2,\ldots,k$). In other words, column $j$ is one of the columns $1,2,\ldots,k$ (since $(i,j)$ lies in column $j$). In other words, $j \in \{1,2,\ldots,k\} = [k]$.

*Proof of Claim 2.* Combining $m \in [m]$ with $m \in \operatorname{Col}(j,T)$, we find $m \in [m] \cap \operatorname{Col}(j,T)$. Furthermore, $[m-1] = [m] \setminus \{m\}$, so that

$$\underbrace{[m-1]}_{=[m]\setminus\{m\}} \cap \operatorname{Col}(j,T) = ([m] \setminus \{m\}) \cap \operatorname{Col}(j,T) = ([m] \cap \operatorname{Col}(j,T)) \setminus \{m\}.$$

Thus, Claim 2 is proved. □

*Claim 3:* We have

$$[m-1] \cap \operatorname{Col}(j,T) = \left\{ T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t, j\right) \right\}.$$

*Proof of Claim 3.* In our construction of the tableau $T$, we filled the cells of $Y(\lambda)$ with the entries $1,2,\ldots,n$ in a specific order – namely in the order in which these cells appeared in the list (459). But the list (459) was listing the cells of $Y(\lambda)$ row by row, from bottom to top. Thus, if a cell $c \in Y(\lambda)$ lies further south than another cell $d \in Y(\lambda)$, then $c$ appears before $d$ in the list (459), and therefore the cell $c$ is filled with a smaller entry in the tableau $T$ than the cell $d$ (since the cells are filled with $1,2,\ldots,n$ in the order in which they appear in the list); in other words, we have $T(c) < T(d)$ in this case. Hence, in particular, the entries of $T$ strictly decrease down each column. In particular, we thus have

$$T(1,j) > T(2,j) > \cdots > T\left(\lambda_j^t, j\right).$$

In other words, if $p$ and $q$ are two numbers in $\left[\lambda_j^t\right]$ satisfying $p < q$, then

$$T(p,j) > T(q,j). \tag{469}$$

Thus,

$$\left\{ T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t, j\right) \right\} \subseteq [m-1]$$

[158] and

$$\{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\} \cap [m-1] = \varnothing$$

[158] *Proof.* Let $g \in \left\{ T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t, j\right) \right\}$. Thus, $g = T(s,j)$ for some $s \in \left\{ i+1, i+2, \ldots, \lambda_j^t \right\}$. Consider this $s$.

From $s \in \left\{ i+1, i+2, \ldots, \lambda_j^t \right\}$, we obtain $s \geq i+1 > i$, so that $i < s$. Therefore, $T(i,j) > T(s,j)$ (by (469), applied to $p = i$ and $q = s$). Hence, $T(s,j) < T(i,j) = m$. Hence, $g = T(s,j) < m$, so that $g \leq m-1$ (since $g$ and $m$ are integers). Therefore, $g \in [m-1]$.

Forget that we fixed $g$. We thus have shown that $g \in [m-1]$ for each $g \in \left\{ T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t, j\right) \right\}$. In other words, $\left\{ T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t, j\right) \right\} \subseteq [m-1]$.

[159]. However, (468) shows that

$$\begin{aligned}\operatorname{Col}(j,T) &= \left\{T(1,j),\ T(2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\}\\ &= \left\{T(1,j),\ T(2,j),\ \ldots,\ T(i,j),\ T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\}\\ &= \{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\} \cup \left\{T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\}.\end{aligned}$$

Therefore,

$$\begin{aligned}&[m-1] \cap \underbrace{\operatorname{Col}(j,T)}_{=\{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\} \cup \left\{T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\}}\\ &= [m-1] \cap \left(\{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\} \cup \left\{T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\}\right)\\ &= \underbrace{\left([m-1] \cap \{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\}\right)}_{=\{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\} \cap [m-1] = \varnothing}\\ &\qquad \cup \underbrace{\left([m-1] \cap \left\{T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\}\right)}_{\substack{=\left\{T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\}\\ \left(\text{since } \left\{T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\} \subseteq [m-1]\right)}}\\ &= \varnothing \cup \left\{T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\}\\ &= \left\{T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\}.\end{aligned}$$

This proves Claim 3. □

*Claim 4:* We have

$$N_{m-1} \sum_{\ell=i+1}^{\lambda_j^t} t_{m,\ T(\ell,j)} = N_{m-1} - N_m. \tag{470}$$

*Proof of Claim 4.* The equality (461) yields

$$N_m = \prod_{v=1}^{k} \nabla^-_{[m] \cap \operatorname{Col}(v,T)}$$

[159] *Proof.* Let $g \in \{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\} \cap [m-1]$. We shall derive a contradiction.

We have $g \in \{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\} \cap [m-1] \subseteq \{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\}$. In other words, $g = T(s,j)$ for some $s \in \{1,2,\ldots,i\}$. Consider this $s$. Thus, $s \leq i$ (since $s \in \{1,2,\ldots,i\}$). On the other hand, $g \in \{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\} \cap [m-1] \subseteq [m-1]$ and therefore $g \leq m-1 < m$. Hence, $T(s,j) = g < m = T(i,j)$. Therefore, $T(s,j) \neq T(i,j)$, so that $s \neq i$. Combining this with $s \leq i$, we obtain $s < i$. Therefore, $T(s,j) > T(i,j)$ (by (469), applied to $p = s$ and $q = i$). But this contradicts $T(s,j) < T(i,j)$.

Forget that we fixed $g$. We thus have found a contradiction for each $g \in \{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\} \cap [m-1]$. Hence, there exists no such $g$. In other words, $\{T(1,j),\ T(2,j),\ \ldots,\ T(i,j)\} \cap [m-1] = \varnothing$.

(here, we have renamed the index $j$ in (461) as $v$, since the letter $j$ already has another meaning in our context). As we know, the order of the factors in this product is immaterial, since they all commute. Thus,

$$N_m = \prod_{v=1}^{k} \nabla^-_{[m]\cap \mathrm{Col}(v,T)} = \prod_{v\in[k]} \nabla^-_{[m]\cap \mathrm{Col}(v,T)} \tag{471}$$

$$= \left( \prod_{\substack{v\in[k];\\ v\neq j}} \nabla^-_{[m]\cap \mathrm{Col}(v,T)} \right) \nabla^-_{[m]\cap \mathrm{Col}(j,T)} \tag{472}$$

(here, we have split off the factor for $v = j$ from our product, since we know that $j \in [k]$). The same argument can be applied to $m-1$ instead of $m$, and thus we find

$$N_{m-1} = \left( \prod_{\substack{v\in[k];\\ v\neq j}} \nabla^-_{[m-1]\cap \mathrm{Col}(v,T)} \right) \nabla^-_{[m-1]\cap \mathrm{Col}(j,T)}. \tag{473}$$

However, for each $v \in [k]$ satisfying $v \neq j$, we have $[m] \cap \mathrm{Col}(v,T) = [m-1] \cap \mathrm{Col}(v,T)$ (by Claim 1) and therefore

$$\nabla^-_{[m]\cap \mathrm{Col}(v,T)} = \nabla^-_{[m-1]\cap \mathrm{Col}(v,T)}. \tag{474}$$

Moreover, Claim 2 shows that

$$[m-1] \cap \mathrm{Col}(j,T) = ([m] \cap \mathrm{Col}(j,T)) \setminus \{m\}.$$

But Claim 2 also shows that $m \in [m] \cap \mathrm{Col}(j,T)$. Thus, Lemma 3.11.6 (applied to $X = [m] \cap \mathrm{Col}(j,T)$ and $x = m$) yields

$$\nabla^-_{[m]\cap \mathrm{Col}(j,T)} = \nabla^-_{([m]\cap \mathrm{Col}(j,T))\setminus\{m\}} \left( 1 - \sum_{y\in([m]\cap \mathrm{Col}(j,T))\setminus\{m\}} t_{y,m} \right).$$

In view of $[m-1] \cap \mathrm{Col}(j,T) = ([m] \cap \mathrm{Col}(j,T)) \setminus \{m\}$, we can rewrite this as

$$\nabla^-_{[m]\cap \mathrm{Col}(j,T)} = \nabla^-_{[m-1]\cap \mathrm{Col}(j,T)} \left( 1 - \sum_{y\in[m-1]\cap \mathrm{Col}(j,T)} t_{y,m} \right). \tag{475}$$

Claim 3 shows that $[m-1] \cap \mathrm{Col}(j,T) = \left\{ T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda^t_j, j\right) \right\}$.

Hence,

$$\begin{aligned}
&\sum_{y\in[m-1]\cap \mathrm{Col}(j,T)} t_{y,m} \\
&= \sum_{y\in\left\{T(i+1,j),\ T(i+2,j),\ \ldots,\ T\left(\lambda_j^t,j\right)\right\}} t_{y,m} \\
&= t_{T(i+1,j),\ m} + t_{T(i+2,j),\ m} + \cdots + t_{T\left(\lambda_j^t,j\right),\ m} \\
&\qquad \left( \begin{array}{c} \text{since the elements } T\left(i+1,j\right),\ T\left(i+2,j\right),\ \ldots,\ T\left(\lambda_j^t,j\right) \\ \text{are distinct (because all entries of } T \text{ are distinct)} \end{array} \right) \\
&= \sum_{\ell=i+1}^{\lambda_j^t} \underbrace{t_{T(\ell,j),\ m}}_{\substack{=t_{m,\ T(\ell,j)} \\ \text{(since } t_{p,q}=t_{q,p} \text{ for any } p\neq q\text{)}}} = \sum_{\ell=i+1}^{\lambda_j^t} t_{m,\ T(\ell,j)}.
\end{aligned}$$

Thus, we can rewrite (475) as

$$\nabla^-_{[m]\cap\mathrm{Col}(j,T)} = \nabla^-_{[m-1]\cap\mathrm{Col}(j,T)} \left(1 - \sum_{\ell=i+1}^{\lambda_j^t} t_{m,\ T(\ell,j)}\right). \tag{476}$$

Now, (472) becomes

$$\begin{aligned}
N_m &= \left( \prod_{\substack{v\in[k];\\ v\neq j}} \underbrace{\nabla^-_{[m]\cap\mathrm{Col}(v,T)}}_{\substack{=\nabla^-_{[m-1]\cap\mathrm{Col}(v,T)} \\ \text{(by (474))}}} \right) \underbrace{\nabla^-_{[m]\cap\mathrm{Col}(j,T)}}_{=\nabla^-_{[m-1]\cap\mathrm{Col}(j,T)} \left(1-\sum_{\ell=i+1}^{\lambda_j^t} t_{m,\ T(\ell,j)}\right)} \\
&\qquad \text{(by (476))} \\
&= \underbrace{\left( \prod_{\substack{v\in[k];\\ v\neq j}} \nabla^-_{[m-1]\cap\mathrm{Col}(v,T)} \right) \nabla^-_{[m-1]\cap\mathrm{Col}(j,T)}}_{\substack{=N_{m-1} \\ \text{(by (473))}}} \left(1 - \sum_{\ell=i+1}^{\lambda_j^t} t_{m,\ T(\ell,j)}\right) \\
&= N_{m-1}\left(1 - \sum_{\ell=i+1}^{\lambda_j^t} t_{m,\ T(\ell,j)}\right) = N_{m-1} - N_{m-1} \sum_{\ell=i+1}^{\lambda_j^t} t_{m,\ T(\ell,j)}.
\end{aligned}$$

Subtracting $N_{m-1}$ from both sides of this equality, we find

$$N_m - N_{m-1} = -N_{m-1} \sum_{\ell=i+1}^{\lambda_j^t} t_{m,\ T(\ell,j)}.$$

Thus,

$$N_{m-1} \sum_{\ell=i+1}^{\lambda_j^t} t_{m,\ T(\ell,j)} = -\left(N_m - N_{m-1}\right) = N_{m-1} - N_m.$$

This proves Claim 4. □

*Claim 5:* Let $\ell \in \left\{i+1, i+2, \ldots, \lambda_j^t\right\}$. Then,

$$\left(1 + \sum_{s=1}^{\lambda_i} t_{T(i,s),\ T(\ell,j)}\right) PN = 0.$$

*Proof of Claim 5.* We have $\ell \in \left\{i+1, i+2, \ldots, \lambda_j^t\right\}$, so that $\ell \geq i+1 > i$ and $\ell \leq \lambda_j^t$. The cell $(\ell, j)$ is one of the cells $(1,j)$, $(2,j)$, $\ldots$, $\left(\lambda_j^t, j\right)$ (since $\ell \geq i+1 \geq 1$ and $\ell \leq \lambda_j^t$), and therefore is a cell in the $j$-th column of $Y(\lambda)$ (since the cells in the $j$-th column of $Y(\lambda)$ are $(1,j)$, $(2,j)$, $\ldots$, $\left(\lambda_j^t, j\right)$). Thus, in particular, $(\ell, j) \in Y(\lambda)$. In other words, $j \in [\lambda_\ell]$ (by the definition of $Y(\lambda)$). Thus, Proposition 5.24.1 (applied to $\ell$, $i$ and $j$ instead of $j$, $k$ and $q$) yields

$$\left(1 + \sum_{s=1}^{\lambda_i} t_{T(i,s),\ T(\ell,j)}\right) \nabla_{\operatorname{Row} T} \nabla^-_{\operatorname{Col} T} = 0.$$

This rewrites as

$$\left(1 + \sum_{s=1}^{\lambda_i} t_{T(i,s),\ T(\ell,j)}\right) PN = 0$$

(since $P = \nabla_{\operatorname{Row} T}$ and $N = \nabla^-_{\operatorname{Col} T}$). Thus, Claim 5 is proved. □

*Claim 6:* Let $p, q \in [n]$ be two distinct numbers that lie in the same column of $T$ and satisfy $p < m$ and $q < m$. Then, $N_{m-1} = N' \cdot \left(1 - t_{p,q}\right)$ for some $N' \in \mathcal{A}$.

*Proof of Claim 6.* We have assumed that $p$ and $q$ lie in the same column of $T$. In other words, $p$ and $q$ lie in the $s$-th column of $T$ for some $s \in [k]$ (since all cells of $Y(\lambda)$ lie in columns $1, 2, \ldots, k$). Consider this $s$. Set $X := [m-1] \cap \operatorname{Col}(s, T)$ and

$$\nabla_X^{\text{even}} := \sum_{\substack{w \in S_{n,X}; \\ (-1)^w = 1}} w$$

(where $S_{n,X}$ is defined according to Proposition 3.7.4).

We have $p \in \operatorname{Col}(s, T)$ (since $p$ lies in the $s$-th column of $T$). Moreover, from $p < m$, we obtain $p \leq m-1$ (since $p$ and $m$ are integers), so that $p \in [m-1]$ (since $p \in [n]$ entails $p \geq 1$). Combining this with $p \in \operatorname{Col}(s, T)$, we obtain

$p \in [m-1] \cap \operatorname{Col}(s,T) = X$. Similarly, $q \in X$. Hence, $p$ and $q$ are two distinct elements of $X$. Therefore, Proposition 3.7.4 **(d)** (applied to $p$ and $q$ instead of $i$ and $j$) yields

$$\nabla_X = \nabla_X^{\text{even}} \cdot \left(1 + t_{p,q}\right) = \left(1 + t_{p,q}\right) \cdot \nabla_X^{\text{even}}$$

and

$$\nabla_X^- = \nabla_X^{\text{even}} \cdot \left(1 - t_{p,q}\right) = \left(1 - t_{p,q}\right) \cdot \nabla_X^{\text{even}}.$$

However, the equality (471) (which we showed during our proof of Claim 4) shows that

$$N_m = \prod_{v \in [k]} \nabla^-_{[m] \cap \operatorname{Col}(v,T)}.$$

The same argument (applied to $m-1$ instead of $m$) yields

$$N_{m-1} = \prod_{v \in [k]} \nabla^-_{[m-1] \cap \operatorname{Col}(v,T)} = \left( \prod_{\substack{v \in [k]; \\ v \neq s}} \nabla^-_{[m-1] \cap \operatorname{Col}(v,T)} \right) \underbrace{\nabla^-_{[m-1] \cap \operatorname{Col}(s,T)}}_{\substack{= \nabla_X^- \\ \text{(since } [m-1] \cap \operatorname{Col}(s,T) = X\text{)}}}$$

$$\left( \begin{array}{c} \text{here, we split off the factor for } v = s \\ \text{from the product, since } s \in [k] \end{array} \right)$$

$$= \left( \prod_{\substack{v \in [k]; \\ v \neq s}} \nabla^-_{[m-1] \cap \operatorname{Col}(v,T)} \right) \underbrace{\nabla_X^-}_{= \nabla_X^{\text{even}} \cdot \left(1 - t_{p,q}\right)}$$

$$= \left( \prod_{\substack{v \in [k]; \\ v \neq s}} \nabla^-_{[m-1] \cap \operatorname{Col}(v,T)} \right) \nabla_X^{\text{even}} \cdot \left(1 - t_{p,q}\right).$$

Therefore, $N_{m-1} = N' \cdot \left(1 - t_{p,q}\right)$ for some $N' \in \mathcal{A}$ (namely, for $N' = \left( \prod_{\substack{v \in [k]; \\ v \neq s}} \nabla^-_{[m-1] \cap \operatorname{Col}(v,T)} \right) \nabla_X^{\text{even}}$). This proves Claim 6. □

*Claim 7:* Let $p, q \in [n]$ be two distinct numbers that lie in the same row of $T$. Then, $t_{p,q} P = P t_{p,q} = P$.

*Proof of Claim 7.* Proposition 5.5.8 **(c)** (applied to $Y(\lambda)$, $p$ and $q$ instead of $D$, $i$ and $j$) shows that $t_{i,j} \nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T} t_{i,j} = \nabla_{\operatorname{Row} T}$. In view of $P = \nabla_{\operatorname{Row} T}$, we can rewrite this as $t_{p,q} P = P t_{p,q} = P$. Thus, Claim 7 is proved. □

*Claim 8:* Let $\ell \in \left\{ i+1, i+2, \ldots, \lambda'_j \right\}$. Let $s \in [j-1]$. Then,

$$N_{m-1} t_{T(i,s),\ T(\ell,j)} P = 0.$$

*Proof of Claim 8.* We have $s \in [j-1]$, so that $s \leq j-1 < j \leq \lambda_i$ (since $(i,j) \in Y(\lambda)$). Thus, $(i,s) \in Y(\lambda)$. Therefore, the entry $T(i,s)$ of $T$ is well-defined. Furthermore, $(\ell,j) \in Y(\lambda)$ (as we saw in the proof of Claim 5). Hence, the entry $T(\ell,j)$ of $T$ is well-defined. Moreover, from $(\ell,j) \in Y(\lambda)$, we obtain $j \leq \lambda_\ell$. Altogether, $s < j \leq \lambda_\ell$, so that $(\ell,s) \in Y(\lambda)$. Thus, the entry $T(\ell,s)$ of $T$ is well-defined.

We have $\ell \in \left\{i+1, i+2, \ldots, \lambda_j^t\right\}$, thus $\ell \geq i+1 > i$ and therefore $\ell \neq i$. Also, $s \neq j$ (since $s < j$).

Set $p := T(\ell,s)$ and $q := T(i,s)$ and $r := T(\ell,j)$. The two cells $(\ell,s)$ and $(\ell,j)$ are distinct (since $s \neq j$), and are both distinct from $(i,s)$ (since $\ell \neq i$). In other words, the three cells $(\ell,s)$ and $(i,s)$ and $(\ell,j)$ are distinct. Hence, the corresponding entries $T(\ell,s)$ and $T(i,s)$ and $T(\ell,j)$ are also distinct (since all entries of $T$ are distinct). In other words, $p$ and $q$ and $r$ are distinct (since $p = T(\ell,s)$ and $q = T(i,s)$ and $r = T(\ell,j)$).

The numbers $p$ and $q$ both lie in the $s$-th column of $T$ (since $p = T(\ell,s)$ and $q = T(i,s)$). Thus, they lie in the same column of $T$.

The numbers $p$ and $r$ both lie in the $\ell$-th row of $T$ (since $p = T(\ell,s)$ and $r = T(\ell,j)$). Thus, they lie in the same row of $T$. Hence, Lemma 5.24.11 (applied to $D = Y(\lambda)$) yields

$$\left(1 - t_{p,q}\right) t_{q,r} \nabla_{\operatorname{Row} T} = 0.$$

Moreover, the construction of $T$ easily shows that $p < m$ [160] and $q < m$ [161]. Moreover, the two numbers $p$ and $q$ are distinct (as we saw above) and lie in the same column of $T$ (as we also saw above). Hence, Claim 6 yields that $N_{m-1} = N' \cdot \left(1 - t_{p,q}\right)$ for some $N' \in \mathcal{A}$. Consider this $N'$.

Now,

$$\underbrace{N_{m-1}}_{=N' \cdot \left(1-t_{p,q}\right)} \underbrace{t_{T(i,s),\, T(\ell,j)}}_{\substack{=t_{q,r} \\ \text{(since } T(i,s)=q \\ \text{and } T(\ell,j)=r\text{)}}} \underbrace{P}_{=\nabla_{\operatorname{Row} T}} = N' \cdot \underbrace{\left(1 - t_{p,q}\right) t_{q,r} \nabla_{\operatorname{Row} T}}_{=0} = 0.$$

[160] *Proof.* We have $\ell > i$. Thus, the cell $(\ell,s)$ lies further south than the cell $(i,j)$.

However, during our proof of Claim 3, we saw that if a cell $c \in Y(\lambda)$ lies further south than another cell $d \in Y(\lambda)$, then $T(c) < T(d)$. We can apply this to $c = (\ell,s)$ and $d = (i,j)$, and thus conclude that $T(\ell,s) < T(i,j)$ (since the cell $(\ell,s)$ lies further south than the cell $(i,j)$). In other words, $p < m$ (since $p = T(\ell,s)$ and $m = T(i,j)$).

[161] *Proof.* The cell $(i,s)$ lies in the same row as the cell $(i,j)$ but further west (since $s < j$).

In our construction of the tableau $T$, we filled the cells of $Y(\lambda)$ with the entries $1, 2, \ldots, n$ in a specific order – namely in the order in which these cells appeared in the list (459). But the list (459) was listing the cells of $Y(\lambda)$ row by row, where the cells of each row were listed from left to right. Thus, if a cell $c \in Y(\lambda)$ lies in the same row as another cell $d \in Y(\lambda)$ but further west, then $c$ appears before $d$ in the list (459), and therefore the cell $c$ is filled with a smaller entry in the tableau $T$ than the cell $d$ (since the cells are filled with $1, 2, \ldots, n$ in the order in which they appear in the list); in other words, we have $T(c) < T(d)$ in this case. Applying this to $c = (i,s)$ and $d = (i,j)$, we conclude that $T(i,s) < T(i,j)$ (since the cell $(i,s)$ lies in the same row as the cell $(i,j)$ but further west). In other words, $q < m$ (since $q = T(i,s)$ and $m = T(i,j)$).

This proves Claim 8. □

*Claim 9:* Let $p, q \in [n]$ be two distinct numbers that satisfy $p \geq m$ and $q \geq m$. Then, $t_{p,q} N_{m-1} = N_{m-1} t_{p,q}$.

*Proof of Claim 9.* The equality (461) yields

$$N_m = \prod_{v=1}^{k} \nabla^-_{[m] \cap \mathrm{Col}(v,T)}$$

(here, we have renamed the index $j$ in (461) as $v$, since the letter $j$ already has another meaning in our context). The same argument can be applied to $m-1$ instead of $m$, and thus we find

$$\begin{aligned} N_{m-1} &= \prod_{v=1}^{k} \nabla^-_{[m-1] \cap \mathrm{Col}(v,T)} \\ &= \nabla^-_{[m-1] \cap \mathrm{Col}(1,T)} \nabla^-_{[m-1] \cap \mathrm{Col}(2,T)} \cdots \nabla^-_{[m-1] \cap \mathrm{Col}(k,T)}. \end{aligned}$$

However, it is easy to see (using Proposition 3.7.12 **(d)**) that the transposition $t_{p,q}$ commutes with the element $\nabla^-_{[m-1] \cap \mathrm{Col}(v,T)}$ for each $v \in [k]$ [162]. In other words, $t_{p,q}$ commutes with each of the elements $\nabla^-_{[m-1] \cap \mathrm{Col}(1,T)}, \nabla^-_{[m-1] \cap \mathrm{Col}(2,T)}, \ldots, \nabla^-_{[m-1] \cap \mathrm{Col}(k,T)}$.

However, recall a well-known fact from algebra:

*Fact:* Let $A$ be a ring. Let $a \in A$ and $b_1, b_2, \ldots, b_k \in A$ be some elements. If $a$ commutes with each of the elements $b_1, b_2, \ldots, b_k$, then $a$ must also commute with their product $b_1 b_2 \cdots b_k$. [163]

[162] *Proof.* Let $v \in [k]$. We must prove that $t_{p,q}$ commutes with $\nabla^-_{[m-1] \cap \mathrm{Col}(v,T)}$.

Let $X := [m-1] \cap \mathrm{Col}(v,T)$ and $Y := \{m, m+1, \ldots, n\}$. Then, both $X$ and $Y$ are subsets of $[n]$. Moreover,

$$\underbrace{X}_{\substack{=[m-1]\cap\mathrm{Col}(v,T) \\ \subseteq [m-1]=\{1,2,\ldots,m-1\}}} \cap \underbrace{Y}_{=\{m,m+1,\ldots,n\}} \subseteq \{1,2,\ldots,m-1\} \cap \{m, m+1, \ldots, n\} = \varnothing$$

(since the sets $\{1,2,\ldots,m-1\}$ and $\{m, m+1, \ldots, n\}$ are obviously disjoint). Thus, $X \cap Y = \varnothing$. In other words, the sets $X$ and $Y$ are disjoint. Furthermore, from $p \geq m$ and $p \leq n$ (since $p \in [n]$), we obtain $p \in \{m, m+1, \ldots, n\} = Y$. Similarly, $q \in Y$. Thus, $p$ and $q$ are two distinct elements of $Y$. Hence, Proposition 3.7.12 **(d)** shows that $\nabla^-_X t_{p,q} = t_{p,q} \nabla^-_X$. In other words, $t_{p,q}$ commutes with $\nabla^-_X$. In other words, $t_{p,q}$ commutes with $\nabla^-_{[m-1] \cap \mathrm{Col}(v,T)}$ (since $X = [m-1] \cap \mathrm{Col}(v,T)$). This completes our proof.

[163] The easiest way to prove this is just by observing that

$$\begin{aligned} a(b_1 b_2 \cdots b_k) &= \underbrace{ab_1}_{=b_1 a} b_2 b_3 \cdots b_k = b_1 \underbrace{ab_2}_{=b_2 a} b_3 \cdots b_k = b_1 b_2 \underbrace{ab_3}_{=b_3 a} \cdots b_k \\ &= b_1 b_2 b_3 a \cdots b_k = \cdots = b_1 b_2 b_3 \cdots b_k a = (b_1 b_2 \cdots b_k) a. \end{aligned}$$

This can be easily formalized as an induction on $k$.

Applying this fact to $A = \mathcal{A}$ and $a = t_{p,q}$ and $b_v = \nabla^-_{[m-1]\cap \mathrm{Col}(v,T)}$, we conclude that $t_{p,q}$ must commute with the product $\nabla^-_{[m-1]\cap \mathrm{Col}(1,T)} \nabla^-_{[m-1]\cap \mathrm{Col}(2,T)} \cdots \nabla^-_{[m-1]\cap \mathrm{Col}(k,T)}$ (since $t_{p,q}$ commutes with each of the elements $\nabla^-_{[m-1]\cap \mathrm{Col}(1,T)}, \nabla^-_{[m-1]\cap \mathrm{Col}(2,T)}, \ldots, \nabla^-_{[m-1]\cap \mathrm{Col}(k,T)}$). In other words, $t_{p,q}$ must commute with $N_{m-1}$ (since $N_{m-1} = \nabla^-_{[m-1]\cap \mathrm{Col}(1,T)} \nabla^-_{[m-1]\cap \mathrm{Col}(2,T)} \cdots \nabla^-_{[m-1]\cap \mathrm{Col}(k,T)}$). In other words, we have $t_{p,q} N_{m-1} = N_{m-1} t_{p,q}$. This proves Claim 9. □

*Claim 10:* Let $\ell \in \left\{ i+1, i+2, \ldots, \lambda^t_j \right\}$. Let $s \in \{j, j+1, \ldots, \lambda_i\}$. Then,

$$PN_{m-1} t_{T(i,s),\ T(\ell,j)} P = PN_{m-1} t_{m,\ T(\ell,j)} P.$$

*Proof of Claim 10.* We have $(\ell, j) \in Y(\lambda)$ (as we saw in the proof of Claim 5). Hence, the entry $T(\ell, j)$ of $T$ is well-defined. From $\ell \in \left\{ i+1, i+2, \ldots, \lambda^t_j \right\}$, we obtain $\ell \geq i+1 > i$. From $s \in \{j, j+1, \ldots, \lambda_i\}$, we obtain $s \leq \lambda_i$. Hence, $(i, s) \in Y(\lambda)$. Thus, the entry $T(i, s)$ of $T$ is well-defined.

Set $p := T(i, j) = m$ and $q := T(\ell, j)$ and $r := T(i, s)$. From $\ell > i$, we obtain $i \neq \ell$ and thus $(i, j) \neq (\ell, j)$. Hence, $T(i, j) \neq T(\ell, j)$ (since all entries of $T$ are distinct). In other words, $p \neq q$ (since $p = T(i, j)$ and $q = T(\ell, j)$). Furthermore, from $\ell > i$, we obtain $\ell \neq i$ and thus $(\ell, j) \neq (i, s)$. Hence, $T(\ell, j) \neq T(i, s)$ (since all entries of $T$ are distinct). In other words, $q \neq r$ (since $q = T(\ell, j)$ and $r = T(i, s)$). Clearly, the numbers $p$ and $r$ both lie in the $i$-th row of $T$ (since $p = T(i, j)$ and $r = T(i, s)$). Thus, the numbers $p$ and $r$ lie in the same row of $T$.

We must prove that $PN_{m-1} t_{T(i,s),\ T(\ell,j)} P = PN_{m-1} t_{m,\ T(\ell,j)} P$. If $T(i, s) = m$, then this is obvious. Thus, we WLOG assume that $T(i, s) \neq m$. Hence, $T(i, s) \neq m = T(i, j)$, so that $s \neq j$. But $s \geq j$ (since $s \in \{j, j+1, \ldots, \lambda_i\}$). Combining this with $s \neq j$, we find $s > j$, so that $j < s$.

Also, $r \neq p$ (since $r = T(i, s) \neq m = p$). Combining this with $p \neq q$ and $q \neq r$, we see that the three numbers $p, q, r$ are distinct.

The numbers $p$ and $r$ are distinct (since $r \neq p$) and satisfy $p \geq m$ (since $p = m$) and $r \geq m$ [164]. Hence, $t_{p,r} N_{m-1} = N_{m-1} t_{p,r}$ (by Claim 9, applied to $r$

[164] *Proof.* The cell $(i, j)$ lies in the same row as the cell $(i, s)$ but further west (since $j < s$).

In our construction of the tableau $T$, we filled the cells of $Y(\lambda)$ with the entries $1, 2, \ldots, n$ in a specific order – namely in the order in which these cells appeared in the list (459). But the list (459) was listing the cells of $Y(\lambda)$ row by row, where the cells of each row were listed from left to right. Thus, if a cell $c \in Y(\lambda)$ lies in the same row as another cell $d \in Y(\lambda)$ but further west, then $c$ appears before $d$ in the list (459), and therefore the cell $c$ is filled with a smaller entry in the tableau $T$ than the cell $d$ (since the cells are filled with $1, 2, \ldots, n$ in the order in which they appear in the list); in other words, we have $T(c) < T(d)$ in this case. Applying this to $c = (i, j)$ and $d = (i, s)$, we conclude that $T(i, j) < T(i, s)$ (since the cell $(i, j)$ lies in the same row as the cell $(i, s)$ but further west). In other words, $m < r$ (since $m = T(i, j)$ and $r = T(i, s)$). Thus, $r > m$, so that $r \geq m$.

instead of $q$). In other words, the element $N_{m-1}$ of $\mathcal{A}$ commutes with $t_{p,r}$. Thus, Lemma 5.24.12 (applied to $D = Y(\lambda)$ and $\mathbf{a} = N_{m-1}$) yields

$$\nabla_{\operatorname{Row} T} N_{m-1} t_{r,q} \nabla_{\operatorname{Row} T} = \nabla_{\operatorname{Row} T} N_{m-1} t_{p,q} \nabla_{\operatorname{Row} T}.$$

In view of $\nabla_{\operatorname{Row} T} = P$ and $r = T(i,s)$ and $q = T(\ell,j)$ and $p = m$, we can rewrite this as

$$PN_{m-1} t_{T(i,s),\ T(\ell,j)} P = PN_{m-1} t_{m,\ T(\ell,j)} P.$$

Thus, Claim 10 is proved. □

*Claim 11:* Let $\ell \in \left\{ i+1, i+2, \ldots, \lambda_j^t \right\}$. Then,

$$PN_{m-1}PN = -\left|\operatorname{Arm}_\lambda^+(c)\right| \cdot PN_{m-1} t_{m,\ T(\ell,j)} PN.$$

*Proof of Claim 11.* From $(i,j) \in Y(\lambda)$, we obtain $j \in [\lambda_i]$, so that $1 \leq j \leq \lambda_i$. We have

$$PN_{m-1} \underbrace{\left(1 + \sum_{s=1}^{\lambda_i} t_{T(i,s),\ T(\ell,j)}\right) PN}_{=0 \text{ (by Claim 5)}} = 0.$$

Thus,

$$\begin{aligned} 0 &= PN_{m-1} \left(1 + \sum_{s=1}^{\lambda_i} t_{T(i,s),\ T(\ell,j)}\right) PN \\ &= PN_{m-1}PN + PN_{m-1} \left(\sum_{s=1}^{\lambda_i} t_{T(i,s),\ T(\ell,j)}\right) PN. \end{aligned}$$

Subtracting $PN_{m-1}PN$ from this equality, we find

$$\begin{aligned}
-PN_{m-1}PN &= PN_{m-1}\left(\sum_{s=1}^{\lambda_i} t_{T(i,s),\ T(\ell,j)}\right)PN = \sum_{s=1}^{\lambda_i} PN_{m-1}t_{T(i,s),\ T(\ell,j)}PN \\
&= \sum_{s=1}^{j-1} P\underbrace{N_{m-1}t_{T(i,s),\ T(\ell,j)}P}_{\substack{=0 \\ \text{(by Claim 8)}}}N + \sum_{s=j}^{\lambda_i}\underbrace{PN_{m-1}t_{T(i,s),\ T(\ell,j)}P}_{\substack{=PN_{m-1}t_{m,\ T(\ell,j)}P \\ \text{(by Claim 10)}}}N \\
&\qquad\qquad (\text{since } 1 \leq j \leq \lambda_i) \\
&= \underbrace{\sum_{s=1}^{j-1} P0N}_{=0} + \underbrace{\sum_{s=j}^{\lambda_i} PN_{m-1}t_{m,\ T(\ell,j)}PN}_{\substack{=(\lambda_i - j+1)PN_{m-1}t_{m,\ T(\ell,j)}PN \\ \text{(since } j \leq \lambda_i\text{)}}} \\
&= \underbrace{(\lambda_i - j + 1)}_{\substack{=\left|\mathrm{Arm}^+_\lambda(c)\right| \\ \text{(by Proposition 5.24.8 \textbf{(a)})}}} PN_{m-1}t_{m,\ T(\ell,j)}PN \\
&= \left|\mathrm{Arm}^+_\lambda(c)\right| \cdot PN_{m-1}t_{m,\ T(\ell,j)}PN.
\end{aligned}$$

Multiplying this equality by $-1$, we obtain

$$PN_{m-1}PN = -\left|\mathrm{Arm}^+_\lambda(c)\right| \cdot PN_{m-1}t_{m,\ T(\ell,j)}PN.$$

This proves Claim 11. □

*Claim 12:* We have

$$\left|\mathrm{Leg}^-_\lambda(c)\right| \cdot PN_{m-1}PN = -\left|\mathrm{Arm}^+_\lambda(c)\right| \cdot P\left(N_{m-1} - N_m\right)PN.$$

*Proof of Claim 12.* Proposition 5.24.8 **(b)** yields $\left|\mathrm{Leg}^-_\lambda(c)\right| = \lambda^t_j - i$, so that $\lambda^t_j - i = \left|\mathrm{Leg}^-_\lambda(c)\right| \geq 0$ and therefore $i \leq \lambda^t_j$. Hence,

$$\sum_{\ell=i+1}^{\lambda^t_j} PN_{m-1}PN = \underbrace{\left(\lambda^t_j - i\right)}_{=\left|\mathrm{Leg}^-_\lambda(c)\right|} \cdot PN_{m-1}PN = \left|\mathrm{Leg}^-_\lambda(c)\right| \cdot PN_{m-1}PN.$$

Thus,

$$\begin{aligned}
\left|\mathrm{Leg}_\lambda^- (c)\right| \cdot PN_{m-1}PN &= \sum_{\ell=i+1}^{\lambda_j^t} \underbrace{PN_{m-1}PN}_{=-\left|\mathrm{Arm}_\lambda^+(c)\right| \cdot PN_{m-1}t_{m,\ T(\ell,j)}PN \text{ (by Claim 11)}} \\
&= \sum_{\ell=i+1}^{\lambda_j^t} \left(-\left|\mathrm{Arm}_\lambda^+ (c)\right| \cdot PN_{m-1}t_{m,\ T(\ell,j)}PN\right) \\
&= -\left|\mathrm{Arm}_\lambda^+ (c)\right| \cdot P \underbrace{N_{m-1}\left(\sum_{\ell=i+1}^{\lambda_j^t} t_{m,\ T(\ell,j)}\right)}_{=N_{m-1}-N_m \text{ (by Claim 4)}} PN \\
&= -\left|\mathrm{Arm}_\lambda^+ (c)\right| \cdot P\left(N_{m-1}-N_m\right)PN.
\end{aligned}$$

This proves Claim 12. □

*Claim 13:* We have

$$\left|H_\lambda (c)\right| \cdot a_{m-1} = \left|\mathrm{Arm}_\lambda^+ (c)\right| \cdot a_m.$$

*Proof of Claim 13.* The equality (466) (with the index $c$ renamed as $d$) says that

$$a_m = \prod_{d \in T^{-1}([m])} \frac{\left|H_\lambda (d)\right|}{\left|\mathrm{Arm}_\lambda^+ (d)\right|}. \tag{477}$$

The same argument (applied to $m-1$ instead of $m$) yields

$$a_{m-1} = \prod_{d \in T^{-1}([m-1])} \frac{\left|H_\lambda (d)\right|}{\left|\mathrm{Arm}_\lambda^+ (d)\right|}. \tag{478}$$

But we can easily see that $c \in T^{-1}([m])$ [165] and $T^{-1}([m]) \setminus \{c\} = T^{-1}([m-1])$ [166].

[165] *Proof.* We have $T(c) = m \in [m]$. Thus, $c \in T^{-1}([m])$.

[166] *Proof.* We know that $T$ is an $n$-tableau of shape $Y(\lambda)$, thus a bijection from $Y(\lambda)$ to $[n]$. Hence, $T^{-1}(X \setminus Y) = T^{-1}(X) \setminus T^{-1}(Y)$ for any two subsets $X$ and $Y$ of $[n]$ (indeed, this is true for any map, not just for a bijection). Applying this to $X = [m]$ and $Y = \{m\}$, we obtain

$$T^{-1}([m] \setminus \{m\}) = T^{-1}([m]) \setminus T^{-1}(\{m\}).$$

However, $T$ is a bijection from $Y(\lambda)$ to $[n]$. Thus, from $T(c) = m$, we obtain $T^{-1}(m) = c$

Now, (477) becomes

$$\begin{aligned} a_m &= \prod_{d \in T^{-1}([m])} \frac{|H_\lambda(d)|}{\left|\mathrm{Arm}^+_\lambda(d)\right|} = \frac{|H_\lambda(c)|}{\left|\mathrm{Arm}^+_\lambda(c)\right|} \cdot \prod_{d \in T^{-1}([m]) \setminus \{c\}} \frac{|H_\lambda(d)|}{\left|\mathrm{Arm}^+_\lambda(d)\right|} \\ &\qquad \left( \begin{array}{c} \text{here, we have split off the factor for } d = c \\ \text{from the product, since } c \in T^{-1}([m]) \end{array} \right) \\ &= \frac{|H_\lambda(c)|}{\left|\mathrm{Arm}^+_\lambda(c)\right|} \cdot \underbrace{\prod_{d \in T^{-1}([m-1])} \frac{|H_\lambda(d)|}{\left|\mathrm{Arm}^+_\lambda(d)\right|}}_{\substack{=a_{m-1} \\ \text{(by (478))}}} \\ &\qquad \left( \text{since } T^{-1}([m]) \setminus \{c\} = T^{-1}([m-1]) \right) \\ &= \frac{|H_\lambda(c)|}{\left|\mathrm{Arm}^+_\lambda(c)\right|} \cdot a_{m-1}. \end{aligned}$$

Multiplying this equality by $\left|\mathrm{Arm}^+_\lambda(c)\right|$, we find

$$\left|\mathrm{Arm}^+_\lambda(c)\right| \cdot a_m = |H_\lambda(c)| \cdot a_{m-1}.$$

This proves Claim 13. □

Now, Claim 12 yields

$$\begin{aligned} \left|\mathrm{Leg}^-_\lambda(c)\right| \cdot PN_{m-1}PN &= -\left|\mathrm{Arm}^+_\lambda(c)\right| \cdot \underbrace{P\left(N_{m-1} - N_m\right)PN}_{=PN_{m-1}PN - PN_mPN} \\ &= -\left|\mathrm{Arm}^+_\lambda(c)\right| \cdot \left(PN_{m-1}PN - PN_mPN\right) \\ &= -\left|\mathrm{Arm}^+_\lambda(c)\right| \cdot PN_{m-1}PN + \left|\mathrm{Arm}^+_\lambda(c)\right| \cdot PN_mPN. \end{aligned}$$

---

and $T^{-1}(\{m\}) = \left\{T^{-1}(m)\right\} = \{c\}$ (since $T^{-1}(m) = c$). Thus,

$$T^{-1}([m] \setminus \{m\}) = T^{-1}([m]) \setminus \underbrace{T^{-1}(\{m\})}_{=\{c\}} = T^{-1}([m]) \setminus \{c\}.$$

Therefore,

$$T^{-1}([m]) \setminus \{c\} = T^{-1}\left(\underbrace{[m] \setminus \{m\}}_{=[m-1]}\right) = T^{-1}([m-1]).$$

Solving this equality for $\left|\mathrm{Arm}_\lambda^+(c)\right| \cdot PN_mPN$, we obtain

$$\begin{aligned}
\left|\mathrm{Arm}_\lambda^+(c)\right| \cdot PN_mPN &= \left|\mathrm{Arm}_\lambda^+(c)\right| \cdot PN_{m-1}PN + \left|\mathrm{Leg}_\lambda^-(c)\right| \cdot PN_{m-1}PN \\
&= \underbrace{\left(\left|\mathrm{Arm}_\lambda^+(c)\right| + \left|\mathrm{Leg}_\lambda^-(c)\right|\right)}_{=|H_\lambda(c)| \text{ (by Proposition 5.24.8 \textbf{(c)})}} \cdot PN_{m-1}PN \\
&= |H_\lambda(c)| \cdot \underbrace{PN_{m-1}PN}_{=a_{m-1}P^2N \text{ (since we assumed (as our induction hypothesis) that (467) holds for } m-1 \text{ instead of } m)} \\
&= \underbrace{|H_\lambda(c)| \cdot a_{m-1}}_{=\left|\mathrm{Arm}_\lambda^+(c)\right| \cdot a_m \text{ (by Claim 13)}} P^2N \\
&= \left|\mathrm{Arm}_\lambda^+(c)\right| \cdot a_m P^2N. \qquad (479)
\end{aligned}$$

But the set $\mathrm{Arm}_\lambda^+(c)$ is nonempty (since it contains $c$ by its definition). Hence, its size $\left|\mathrm{Arm}_\lambda^+(c)\right|$ is a positive integer. Thus, we can divide the equality (479) by $\left|\mathrm{Arm}_\lambda^+(c)\right|$ (since our base ring $\mathbf{k}$ is $\mathbb{Q}$ and thus contains the rational number $\dfrac{1}{\left|\mathrm{Arm}_\lambda^+(c)\right|}$). We thus obtain

$$PN_mPN = a_mP^2N.$$

Thus, we have shown that (467) holds for $m$. This completes the induction step. Hence, (467) is proved. As explained above, Theorem 5.2.8 now follows. □

# 6. Bases of $\mathbf{k}[S_n]$

In this chapter, we will systematically study various bases of the $\mathbf{k}$-module $\mathcal{A} = \mathbf{k}[S_n]$. This includes revisiting the bases we have already encountered:

- the standard basis $(w)_{w \in S_n}$,
- the Young natural basis $(\mathbf{E}_{U,V})_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \mathrm{SYT}(\lambda)}$ from Corollary 5.15.28,
- the basis $(\mathbf{F}_U w_{U,V} \mathbf{E}_V)_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \mathrm{SYT}(\lambda)}$ from Theorem 5.21.2,

but also introducing further important bases, including the *Murphy bases* and the *seminormal basis*.

## 6.1. Notations

We will use the following notations throughout this chapter:

**Definition 6.1.1.** **(a)** We let $\mathcal{A}$ be the $\mathbf{k}$-algebra $\mathbf{k}[S_n]$.

**(b)** We write "$\lambda \vdash n$" for "$\lambda$ is a partition of $n$".

**(c)** For any partition $\lambda$ of $n$, we let $\operatorname{SYT}(\lambda)$ be the set of all standard $n$-tableaux of shape $Y(\lambda)$ (that is, of shape $\lambda$).

**(d)** For any partition $\lambda$ of $n$, we let $f^\lambda$ denote the # of standard tableaux of shape $Y(\lambda)$.

**(e)** A *standard $n$-bitableau* shall mean a triple $(\lambda, U, V)$, where $\lambda$ is a partition of $n$ and where $U$ and $V$ are two standard $n$-tableaux of shape $\lambda$. (In particular, $U$ and $V$ must be straight-shaped and of the same shape.)[167]

We let $\operatorname{SBT}(n)$ be the set of all standard $n$-bitableaux.

**(f)** We recall the sign-twist $T_{\operatorname{sign}} : \mathbf{k}[S_n] \to \mathbf{k}[S_n]$ introduced in Definition 3.11.2. This is a $\mathbf{k}$-linear map from $\mathcal{A}$ to $\mathcal{A}$ (since $\mathcal{A} = \mathbf{k}[S_n]$), and sends each $w \in S_n$ to $(-1)^w w$ (by its definition).

**(g)** We recall the antipode $S : \mathbf{k}[S_n] \to \mathbf{k}[S_n]$ introduced in Definition 3.11.12. This is a $\mathbf{k}$-linear map from $\mathcal{A}$ to $\mathcal{A}$ (since $\mathcal{A} = \mathbf{k}[S_n]$), and sends each $w \in S_n$ to $w^{-1}$ (by its definition).

The only of these notations that is new is the concept of a standard $n$-bitableau. Let us give an example:

**Example 6.1.2.** The standard 3-bitableaux are the six triples

$$\left((3), \begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array}, \begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array}\right),$$

$$\left((2,1), \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}\right), \qquad \left((2,1), \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}\right),$$

$$\left((2,1), \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}\right), \qquad \left((2,1), \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}\right),$$

$$\left((1,1,1), \begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array}, \begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array}\right).$$

[167] Note that the $\lambda$ in a standard $n$-bitableau $(\lambda, U, V)$ can be easily reconstructed from $U$ (indeed, each entry $\lambda_i$ of $\lambda$ is simply the number of entries in row $i$ of $U$). But we still include it in the $n$-bitableau for convenience.

Using the shorthand notation from Convention 5.2.4, we can rewrite them as

$$\begin{array}{llll} ((3),\ 123,\ 123), & & & \\ ((2,1),\ 12\backslash\backslash 3,\ 12\backslash\backslash 3), & & ((2,1),\ 12\backslash\backslash 3,\ 13\backslash\backslash 2), & \\ ((2,1),\ 13\backslash\backslash 2,\ 12\backslash\backslash 3), & & ((2,1),\ 13\backslash\backslash 2,\ 13\backslash\backslash 2), & \\ ((1,1,1),\ 1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3). & & & \end{array}$$

We also recall Definition 5.5.1, which introduces the row symmetrizer

$$\nabla_{\operatorname{Row} T} := \sum_{w \in \mathcal{R}(T)} w \in \mathbf{k}\left[S_n\right] = \mathcal{A}$$

and the column antisymmetrizer

$$\nabla^{-}_{\operatorname{Col} T} := \sum_{w \in \mathcal{C}(T)} (-1)^w w \in \mathbf{k}\left[S_n\right] = \mathcal{A}$$

for each $n$-tableau $T$. Here, $\mathcal{R}(T)$ and $\mathcal{C}(T)$ are two subgroups of $S_n$, defined in Definition 5.3.1. (In a nutshell: $\mathcal{R}(T)$ consists of those $w \in S_n$ such that each $i \in [n]$ lies in the same row of $T$ as $w(i)$ does, whereas $\mathcal{C}(T)$ consists of those $w \in S_n$ such that each $i \in [n]$ lies in the same column of $T$ as $w(i)$ does.)

## 6.2. What to do with a basis

### 6.2.1. The size of a basis

The $\mathbf{k}$-module $\mathcal{A} = \mathbf{k}\left[S_n\right]$ has a basis $(w)_{w \in S_n}$, and thus is free of rank $|S_n| = n!$. At least when $\mathbf{k}$ is a field, this entails that **every** basis of $\mathcal{A}$ has size $n!$ (in fact, this is even true for any nontrivial commutative ring $\mathbf{k}$; see e.g. `https://math.stackexchange.com/questions/2651777/` for a proof). Thus it should not come as a surprise that all bases of $\mathcal{A}$ we shall discuss have size $n!$. However, not all of them are indexed by the set $S_n$. Another possible index set is $\operatorname{SBT}(n)$, because of the following simple fact:

**Lemma 6.2.1.** We have $|\operatorname{SBT}(n)| = n!$.

*Proof.* This is just a restatement of Corollary 5.12.18, since the definition of standard $n$-bitableaux shows that $|\operatorname{SBT}(n)| = \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2$. For more details, see the appendix (Section A.4). □

All our bases of $\mathcal{A}$ will have indexing set $S_n$ or $\operatorname{SBT}(n)$.

### 6.2.2. The matrix representing a linear map

Bases of $\mathbf{k}$-modules allow us to represent linear maps by matrices. Let us recall how this is defined:[168]

**Definition 6.2.2.** Let $V$ and $W$ be two free $\mathbf{k}$-modules with bases $\overrightarrow{v} = \left(v_q\right)_{q\in Q}$ and $\overrightarrow{w} = \left(w_p\right)_{p\in P}$, respectively. Let $f : V \to W$ be a $\mathbf{k}$-linear map. Then, $[f]_{\overrightarrow{v}\to\overrightarrow{w}}$ shall denote the matrix representing this map $f$ with respect to the bases $\overrightarrow{v}$ and $\overrightarrow{w}$. Explicitly, this is the $P \times Q$-matrix $\left(a_{p,q}\right)_{(p,q)\in P\times Q}$, whose entries $a_{p,q}$ are determined by the requirement that

$$f\left(v_q\right) = \sum_{p\in P} a_{p,q} w_p \qquad \text{for all } q \in Q.$$

Thus, each column of this matrix $[f]_{\overrightarrow{v}\to\overrightarrow{w}}$ corresponds to some index $q \in Q$, and its entries are the coordinates of $f\left(v_q\right)$ with respect to the basis $\overrightarrow{w}$ of $W$.

We note that a $P \times Q$-matrix has its rows indexed by elements of $P$ and its columns indexed by elements of $Q$. If the sets $P$ and $Q$ are equipped with total orders, then we can arrange the rows from top to bottom (by following the order on $P$) and arrange the columns from left to right (by following the order on $Q$). This is how we will be writing down such matrices in this chapter. For instance, an $\{\alpha,\beta\} \times \{\varphi,\psi\}$-matrix $\left(a_{p,q}\right)_{(p,q)\in\{\alpha,\beta\}\times\{\varphi,\psi\}}$ can be written down as $\begin{pmatrix} a_{\alpha,\varphi} & a_{\alpha,\psi} \\ a_{\beta,\varphi} & a_{\beta,\psi} \end{pmatrix}$ if we order its indexing sets by $\alpha < \beta$ and $\varphi < \psi$.

Representing matrices have the following property: If $U, V, W$ are three free $\mathbf{k}$-modules with respective bases $\overrightarrow{u}$, $\overrightarrow{v}$, $\overrightarrow{w}$, and if $f : V \to W$ and $g : U \to V$ are two $\mathbf{k}$-linear maps, then

$$[f \circ g]_{\overrightarrow{u}\to\overrightarrow{w}} = [f]_{\overrightarrow{v}\to\overrightarrow{w}} \cdot [g]_{\overrightarrow{u}\to\overrightarrow{v}}. \tag{480}$$

As a particular case of Definition 6.2.2, if $\overrightarrow{\mathbf{b}}$ is any basis of $\mathcal{A} = \mathbf{k}\left[S_n\right]$, and if $f : \mathcal{A} \to \mathcal{A}$ is any $\mathbf{k}$-linear map, then there exists a matrix $[f]_{\overrightarrow{\mathbf{b}}\to\overrightarrow{\mathbf{b}}}$ representing $f$ with respect to the basis $\overrightarrow{\mathbf{b}}$. In particular, we can apply this to the $\mathbf{k}$-linear maps $f = S$ and $f = T_{\mathrm{sign}}$. The resulting matrices $[f]_{\overrightarrow{\mathbf{b}}\to\overrightarrow{\mathbf{b}}}$ will be $S_n \times S_n$-matrices or $\mathrm{SBT}(n) \times \mathrm{SBT}(n)$-matrices depending on the indexing set of our basis being $S_n$ or $\mathrm{SBT}(n)$; in either case, they will have $n!$ rows and $n!$ columns.

### 6.2.3. The endomorphisms $L(\mathbf{a})$ and $R(\mathbf{a})$ and the matrices $L_{\overrightarrow{\mathbf{b}}}(\mathbf{a})$ and $R_{\overrightarrow{\mathbf{b}}}(\mathbf{a})$

The following is a basic and general construction in algebra:

[168] Recall the notion of a $P \times Q$-matrix; this was defined in Section 5.6.

**Definition 6.2.3.** Let $A$ be a $\mathbf{k}$-algebra. Let $\mathbf{a} \in A$. Then, we define two $\mathbf{k}$-linear maps

$$\begin{aligned} L(\mathbf{a}) : A &\to A, \\ \mathbf{x} &\mapsto \mathbf{ax} \end{aligned}$$

and

$$\begin{aligned} R(\mathbf{a}) : A &\to A, \\ \mathbf{x} &\mapsto \mathbf{xa}. \end{aligned}$$

These maps $L(\mathbf{a})$ and $R(\mathbf{a})$ are called *left multiplication by* $\mathbf{a}$ and *right multiplication by* $\mathbf{a}$, respectively.

**Proposition 6.2.4.** Let $A$ be a $\mathbf{k}$-algebra. Then:

**(a)** We have $L(1) = R(1) = \mathrm{id}_A$.

**(b)** For any $\mathbf{a}, \mathbf{b} \in A$, we have $L(\mathbf{a}+\mathbf{b}) = L(\mathbf{a}) + L(\mathbf{b})$ and $R(\mathbf{a}+\mathbf{b}) = R(\mathbf{a}) + R(\mathbf{b})$.

**(c)** For any $\mathbf{a}, \mathbf{b} \in A$, we have $L(\mathbf{ab}) = L(\mathbf{a}) \circ L(\mathbf{b})$ and $R(\mathbf{ab}) = R(\mathbf{b}) \circ R(\mathbf{a})$.

**(d)** For any $\lambda \in \mathbf{k}$ and $\mathbf{a} \in A$, we have $L(\lambda\mathbf{a}) = \lambda L(\mathbf{a})$ and $R(\lambda\mathbf{a}) = \lambda R(\mathbf{a})$.

*Proof.* Straightforward. For instance, in order to verify the equality $R(\mathbf{ab}) = R(\mathbf{b}) \circ R(\mathbf{a})$ in part **(c)**, we need to show that $(R(\mathbf{ab}))(\mathbf{x}) = (R(\mathbf{b}) \circ R(\mathbf{a}))(\mathbf{x})$ for each $\mathbf{x} \in A$. But this follows from computing both sides:

$$(R(\mathbf{ab}))(\mathbf{x}) = \mathbf{x}(\mathbf{ab}) \qquad \text{(by the definition of } R(\mathbf{ab})\text{)}$$

and

$$\begin{aligned} (R(\mathbf{b}) \circ R(\mathbf{a}))(\mathbf{x}) &= (R(\mathbf{b}))((R(\mathbf{a}))(\mathbf{x})) \\ &= (R(\mathbf{b}))(\mathbf{xa}) \qquad \left(\begin{array}{c}\text{since the definition of } R(\mathbf{a}) \\ \text{yields } (R(\mathbf{a}))(\mathbf{x}) = \mathbf{xa}\end{array}\right) \\ &= (\mathbf{xa})\,\mathbf{b} \qquad \text{(by the definition of } R(\mathbf{b})\text{)} \\ &= \mathbf{x}(\mathbf{ab}) \qquad \text{(by associativity)}. \end{aligned}$$

Comparing these two equalities, we obtain $(R(\mathbf{ab}))(\mathbf{x}) = (R(\mathbf{b}) \circ R(\mathbf{a}))(\mathbf{x})$, as desired. □

**Corollary 6.2.5.** Let $A$ be a $\mathbf{k}$-algebra. Then:

**(a)** The map

$$\begin{aligned} L : A &\to \operatorname{End}_{\mathbf{k}} A, \\ \mathbf{a} &\mapsto L(\mathbf{a}) \end{aligned}$$

is an injective $\mathbf{k}$-algebra morphism.

**(b)** The map

$$\begin{aligned} R : A &\to \operatorname{End}_{\mathbf{k}} A, \\ \mathbf{a} &\mapsto R(\mathbf{a}) \end{aligned}$$

is an injective $\mathbf{k}$-algebra anti-morphism.

*Proof.* Proposition 6.2.4 shows that $L$ is a $\mathbf{k}$-algebra morphism and that $R$ is a $\mathbf{k}$-algebra anti-morphism. It remains to prove that both $L$ and $R$ are injective. Observe that each $\mathbf{a} \in A$ satisfies

$$\begin{aligned} (L(\mathbf{a}))(1_A) &= \mathbf{a} \cdot 1_A && \text{(by the definition of } L(\mathbf{a})) \\ &= \mathbf{a}; \end{aligned}$$

this equality allows us to reconstruct $\mathbf{a}$ from $L(\mathbf{a})$. Hence, the map $L$ is injective. Similarly, we can show that $R$ is injective (by observing that each $\mathbf{a} \in A$ satisfies $(R(\mathbf{a}))(1_A) = \mathbf{a}$). Thus, the proof of Corollary 6.2.5 is complete. □

We can apply Definition 6.2.3, Proposition 6.2.4 and Corollary 6.2.5 to the case $A = \mathcal{A}$; thus we obtain an injective $\mathbf{k}$-algebra morphism $L : \mathcal{A} \to \operatorname{End}_{\mathbf{k}} \mathcal{A}$ and an injective $\mathbf{k}$-algebra anti-morphism $R : \mathcal{A} \to \operatorname{End}_{\mathbf{k}} \mathcal{A}$. Given an element $\mathbf{a} \in \mathcal{A}$, we thus obtain two $\mathbf{k}$-linear endomorphisms $L(\mathbf{a})$ and $R(\mathbf{a})$ of $\mathcal{A}$. If we are furthermore given a basis $\overrightarrow{\mathbf{b}}$ of $\mathcal{A}$, then these endomorphisms $L(\mathbf{a})$ and $R(\mathbf{a})$ are represented by matrices:

**Definition 6.2.6.** Let $\overrightarrow{\mathbf{b}} = (\mathbf{b}_i)_{i \in I}$ be a basis of $\mathcal{A} = \mathbf{k}[S_n]$, and let $\mathbf{a} \in \mathcal{A}$. Then, we define the two $I \times I$-matrices

$$L_{\overrightarrow{\mathbf{b}}}(\mathbf{a}) := [L(\mathbf{a})]_{\overrightarrow{\mathbf{b}} \to \overrightarrow{\mathbf{b}}} \qquad \text{and} \qquad R_{\overrightarrow{\mathbf{b}}}(\mathbf{a}) := [R(\mathbf{a})]_{\overrightarrow{\mathbf{b}} \to \overrightarrow{\mathbf{b}}}.$$

These are the matrices representing the endomorphisms $L(\mathbf{a})$ and $R(\mathbf{a})$ with respect to the basis $\overrightarrow{\mathbf{b}}$.

The main advantage of these matrices compared to the endomorphisms $L(\mathbf{a})$ and $R(\mathbf{a})$ themselves is, of course, that they are concrete objects that we can just write down. Moreover, we need not write them down for all $\mathbf{a} \in \mathcal{A}$; it suffices to know them for $\mathbf{a}$ being the simple transpositions $s_1, s_2, \ldots, s_{n-1}$ as defined in (2). Here is why:

- The simple transpositions $s_1, s_2, \ldots, s_{n-1}$ generate the symmetric group $S_n$ and thus also its group algebra $\mathcal{A}$. Thus, each element of $\mathcal{A}$ is a linear combination of products of these simple transpositions.
- But the matrices $L_{\overrightarrow{\mathbf{b}}}(\mathbf{a})$ and $R_{\overrightarrow{\mathbf{b}}}(\mathbf{a})$ behave nicely under products and linear combinations: For all $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{A}$ and $\lambda_1, \lambda_2 \in \mathbf{k}$, we have

$$\begin{aligned} L_{\overrightarrow{\mathbf{b}}}(\mathbf{a}_1\mathbf{a}_2) &= L_{\overrightarrow{\mathbf{b}}}(\mathbf{a}_1) \cdot L_{\overrightarrow{\mathbf{b}}}(\mathbf{a}_2) && \text{and} \\ R_{\overrightarrow{\mathbf{b}}}(\mathbf{a}_1\mathbf{a}_2) &= R_{\overrightarrow{\mathbf{b}}}(\mathbf{a}_2) \cdot R_{\overrightarrow{\mathbf{b}}}(\mathbf{a}_1) && \text{and} \\ L_{\overrightarrow{\mathbf{b}}}(\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2) &= \lambda_1 L_{\overrightarrow{\mathbf{b}}}(\mathbf{a}_1) + \lambda_2 L_{\overrightarrow{\mathbf{b}}}(\mathbf{a}_2) && \text{and} \\ R_{\overrightarrow{\mathbf{b}}}(\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2) &= \lambda_1 R_{\overrightarrow{\mathbf{b}}}(\mathbf{a}_1) + \lambda_2 R_{\overrightarrow{\mathbf{b}}}(\mathbf{a}_2) \end{aligned}$$

  (thanks to (480) and to Proposition 6.2.4).

Hence, knowing these matrices $L_{\overrightarrow{\mathbf{b}}}(\mathbf{a})$ and $R_{\overrightarrow{\mathbf{b}}}(\mathbf{a})$ for all $\mathbf{a} \in \{s_1, s_2, \ldots, s_{n-1}\}$ is the key to comfortably computing with a given basis $\overrightarrow{\mathbf{b}}$ of $\mathcal{A}$.

### 6.2.4. Sub-spans

What else can we do with a basis $\overrightarrow{\mathbf{b}} = (\mathbf{b}_i)_{i \in I}$ of $\mathcal{A}$ ? We can consider its subfamilies $(\mathbf{b}_i)_{i \in J}$ for various subsets $J$ of $I$, and try to understand their spans $\operatorname{span}\{\mathbf{b}_i \mid i \in J\}$. Are some of these spans $\mathbf{k}$-subalgebras of $\mathcal{A}$ ? Ideals? Left or right ideals?

In the rest of this chapter, we shall analyze various bases of $\mathcal{A}$ with regard to these and other questions.

## 6.3. The standard basis of $\mathbf{k}[S_n]$, revisited

The simplest of all bases of $\mathcal{A}$ is the standard basis $\overrightarrow{\mathrm{std}} := (w)_{w \in S_n}$, which consists of all the $n!$ permutations $w \in S_n$ (or, if we are pedantic, of the corresponding basis vectors of $\mathbf{k}[S_n]$). The antipode $S$ and the sign-twist $T_{\mathrm{sign}}$ act on this basis by the simple formulas

$$S(w) = w^{-1} \qquad \text{and} \qquad T_{\mathrm{sign}}(w) = (-1)^w w \qquad \text{for each } w \in S_n;$$

thus, the matrix $[S]_{\overrightarrow{\mathrm{std}} \to \overrightarrow{\mathrm{std}}}$ is a permutation matrix, whereas the matrix $\left[T_{\mathrm{sign}}\right]_{\overrightarrow{\mathrm{std}} \to \overrightarrow{\mathrm{std}}}$ is diagonal (with diagonal entries 1 and $-1$). For each permutation $u \in S_n$, the matrices $L_{\overrightarrow{\mathrm{std}}}(u)$ and $R_{\overrightarrow{\mathrm{std}}}(u)$ (representing the linear maps $L(u)$ and $R(u)$) are permutation matrices as well, since both maps $L(u)$ and $R(u)$ just permute the standard basis $\overrightarrow{\mathrm{std}}$.

**Example 6.3.1.** For $n = 3$, the entries of the standard basis $\overrightarrow{\mathrm{std}} = (w)_{w \in S_n}$ are the elements of $S_3$, namely id, $s_1$, $s_2$, $t_{1,3}$, $\mathrm{cyc}_{1,2,3}$ and $\mathrm{cyc}_{1,3,2}$. The matrices representing the $\mathbf{k}$-linear maps $S$, $T_{\mathrm{sign}}$, $L(s_1)$, $L(s_2)$, $R(s_1)$ and $R(s_2)$ are as follows (where the rows and columns are indexed by the elements of $S_3$, in the order just listed):

$$[S]_{\overrightarrow{\mathrm{std}} \to \overrightarrow{\mathrm{std}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix};$$

$$\left[T_{\mathrm{sign}}\right]_{\overrightarrow{\mathrm{std}} \to \overrightarrow{\mathrm{std}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix};$$

$$L_{\overrightarrow{\mathrm{std}}}(s_1) = [L(s_1)]_{\overrightarrow{\mathrm{std}} \to \overrightarrow{\mathrm{std}}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix};$$

$$L_{\overrightarrow{\mathrm{std}}}(s_2) = [L(s_2)]_{\overrightarrow{\mathrm{std}} \to \overrightarrow{\mathrm{std}}} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix};$$

$$R_{\overrightarrow{\mathrm{std}}}(s_1) = [R(s_1)]_{\overrightarrow{\mathrm{std}} \to \overrightarrow{\mathrm{std}}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix};$$

$$R_{\overrightarrow{\mathrm{std}}}(s_2) = [R(s_2)]_{\overrightarrow{\mathrm{std}} \to \overrightarrow{\mathrm{std}}} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

What about spans of subfamilies of the standard basis $\overrightarrow{\mathrm{std}} = (w)_{w \in S_n}$ ? Several of these spans are subalgebras, due to the following simple exercise:

**Exercise 6.3.1.** $\boxed{1}$ Let $H$ be a subgroup of a group $G$. Prove that $\operatorname{span}\{w \mid w \in H\}$ is a $\mathbf{k}$-subalgebra of $\mathbf{k}[G]$ that is isomorphic to the group algebra $\mathbf{k}[H]$.

Hence, many subfamilies of the standard basis $\overrightarrow{\mathrm{std}}$ of $\mathcal{A}$ span $\mathbf{k}$-subalgebras of $\mathcal{A}$. But very few of them span ideals, or even left or right ideals. Indeed, if a left ideal of $\mathcal{A}$ contains any of the permutations $w \in S_n$, then it must also contain $w^{-1}w = 1$, and thus equals the whole algebra $\mathcal{A}$. The same applies to right ideals.

## 6.4. The Young natural basis of $\mathbf{k}[S_n]$, revisited

For the present Section 6.4, we assume that $n!$ is invertible in $\mathbf{k}$. Then, Corollary 5.15.28 shows that $(\mathbf{E}_{U,V})_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \mathrm{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}$ (where $\mathbf{E}_{U,V}$ is defined as in Definition 5.15.4). We called this basis the *Young natural basis*. In view of Definition 6.1.1 **(e)**, we can rewrite the condition "$\lambda$ is a partition of $n$, and $U, V \in \mathrm{SYT}(\lambda)$" as "$(\lambda, U, V)$ is a standard $n$-bitableau", or – even shorter – as "$(\lambda, U, V) \in \mathrm{SBT}(n)$". Thus, the Young natural basis can be written as

$$(\mathbf{E}_{U,V})_{(\lambda,U,V) \in \mathrm{SBT}(n)}.$$

Let us denote this basis by $\overrightarrow{\mathrm{ynb}}$.

Describing how the operators $S$, $T_{\mathrm{sign}}$, $L(\mathbf{a})$ and $R(\mathbf{a})$ (for $\mathbf{a} \in \mathbf{k}[S_n]$) act on this basis is not easy. At least Proposition 5.15.16 shows that the composition $T_{\mathrm{sign}} \circ S$ acts in a rather simple way (namely, sending each $\mathbf{E}_{P,Q}$ to $(-1)^{w_{Q,P}} \mathbf{E}_{Q\mathbf{r},P\mathbf{r}}$). No such nice formula exists for the operators $S$ and $T_{\mathrm{sign}}$ in isolation. The left- and right-multiplication maps $L(\mathbf{a})$ and $R(\mathbf{a})$ are somewhat better behaved, but still not explicit.[169] Let us see an example:

**Example 6.4.1.** Let $n = 3$. Then, the Young natural basis $\overrightarrow{\mathrm{ynb}} =$

[169] Exercise 5.15.1 allows us to rewrite a product $u\mathbf{E}_{P,Q}v$ with $u, v \in S_n$ as $\mathbf{E}_{uP,v^{-1}Q}$; but the result is usually not a basis vector of the Young natural basis, since the tableaux $uP$ and $v^{-1}Q$ may not be standard even if $P$ and $Q$ are.

$(\mathbf{E}_{U,V})_{(\lambda,U,V)\in \mathrm{SBT}(n)}$ of $\mathcal{A}$ consists of the six vectors

$$\begin{aligned}
\mathbf{E}_{123,\ 123} &= \nabla = 1 + s_1 + s_2 + t_{1,3} + \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2};\\
\mathbf{E}_{12\backslash\backslash 3,\ 12\backslash\backslash 3} &= 1 + s_1 - t_{1,3} - \mathrm{cyc}_{1,2,3};\\
\mathbf{E}_{12\backslash\backslash 3,\ 13\backslash\backslash 2} &= -s_1 + s_2 + \mathrm{cyc}_{1,2,3} - \mathrm{cyc}_{1,3,2};\\
\mathbf{E}_{13\backslash\backslash 2,\ 12\backslash\backslash 3} &= s_2 - t_{1,3} - \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2};\\
\mathbf{E}_{13\backslash\backslash 2,\ 13\backslash\backslash 2} &= 1 - s_1 + t_{1,3} - \mathrm{cyc}_{1,3,2};\\
\mathbf{E}_{1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3} &= \nabla^- = 1 - s_1 - s_2 - t_{1,3} + \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2}\,.
\end{aligned}$$

Let us agree to order the indexing set $\mathrm{SBT}(3)$ of this basis as follows:

$((3),\ 123,\ 123)$, $((2,1),\ 12\backslash\backslash 3,\ 12\backslash\backslash 3)$, $((2,1),\ 12\backslash\backslash 3,\ 13\backslash\backslash 2)$, $((2,1),\ 13\backslash\backslash 2,\ 12\backslash\backslash 3)$, $((2,1),\ 13\backslash\backslash 2,\ 13\backslash\backslash 2)$, $((1,1,1),\ 1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3)$.

Then, the matrices representing the $\mathbf{k}$-linear maps $S$, $T_{\mathrm{sign}}$, $L(s_1)$, $L(s_2)$, $R(s_1)$ and $R(s_2)$ are as follows:

$$[S]_{\overrightarrow{\mathrm{ynb}}\to\overrightarrow{\mathrm{ynb}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4/3 & -2/3 & 2/3 & -1/3 & 0 \\ 0 & 2/3 & -1/3 & 4/3 & -2/3 & 0 \\ 0 & -2/3 & 4/3 & -1/3 & 2/3 & 0 \\ 0 & -1/3 & 2/3 & -2/3 & 4/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix};$$

$$\left[T_{\mathrm{sign}}\right]_{\overrightarrow{\mathrm{ynb}}\to\overrightarrow{\mathrm{ynb}}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1/3 & 2/3 & -2/3 & 4/3 & 0 \\ 0 & -2/3 & 1/3 & -4/3 & 2/3 & 0 \\ 0 & 2/3 & -4/3 & 1/3 & -2/3 & 0 \\ 0 & 4/3 & -2/3 & 2/3 & -1/3 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix};$$

$$L_{\overrightarrow{\mathrm{ynb}}}(s_1) = [L(s_1)]_{\overrightarrow{\mathrm{ynb}}\to\overrightarrow{\mathrm{ynb}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

$$L_{\overrightarrow{\mathrm{ynb}}}(s_2) = [L(s_2)]_{\overrightarrow{\mathrm{ynb}}\to\overrightarrow{\mathrm{ynb}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

$$R_{\overrightarrow{\mathrm{ynb}}}(s_1) = \left[R(s_1)\right]_{\overrightarrow{\mathrm{ynb}} \to \overrightarrow{\mathrm{ynb}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

$$R_{\overrightarrow{\mathrm{ynb}}}(s_2) = \left[R(s_2)\right]_{\overrightarrow{\mathrm{ynb}} \to \overrightarrow{\mathrm{ynb}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

The simplicity of the last four matrices is somewhat misleading: For higher values of $n$, the matrix $L_{\overrightarrow{\mathrm{ynb}}}(s_2)$ (for example) can contain entries other than 0, 1 and $-1$. (For instance, it contains a 2 when $n = 8$.) The same applies to $R_{\overrightarrow{\mathrm{ynb}}}(s_2)$.

**Question 6.4.2.** Can this happen with $L_{\overrightarrow{\mathrm{ynb}}}(s_1)$ ?

However, some other patterns that can be spotted on the above matrices are generalizable. For instance, the matrices $R_{\overrightarrow{\mathrm{ynb}}}(s_1)$ and $R_{\overrightarrow{\mathrm{ynb}}}(s_2)$ (for $n = 3$) are block-diagonal, where the diagonal blocks have sizes $1, 2, 2, 1$. The same holds for the matrices $L_{\overrightarrow{\mathrm{ynb}}}(s_1)$ and $L_{\overrightarrow{\mathrm{ynb}}}(s_2)$ after a permutation of the rows and the columns. These observations are not a fluke; they are manifestations of Lemma 5.15.17. Namely:

- The equality (279) shows that for each partition $\lambda$ of $n$ and each standard tableau $Q \in \mathrm{SYT}(\lambda)$, the subfamily $\left(\mathbf{E}_{P,Q}\right)_{P \in \mathrm{SYT}(D)}$ of the Young natural basis spans the left ideal $\mathcal{A}\mathbf{E}_Q$. In particular, this means that its span is preserved by left multiplication with each $s_i$. Thus, the matrices $L_{\overrightarrow{\mathrm{ynb}}}(s_i)$ for all $i \in [n-1]$ (and, more generally, the matrices $L_{\overrightarrow{\mathrm{ynb}}}(\mathbf{a})$ for all $\mathbf{a} \in \mathcal{A}$) become block-diagonal if the rows and the columns are ordered in such a way that the standard $n$-bitableaux $(\lambda, P, Q)$ with equal $Q$'s appear as a contiguous block.
- The equality (280) shows that for each partition $\lambda$ of $n$ and each standard tableau $Q \in \mathrm{SYT}(\lambda)$, the subfamily $\left(\mathbf{E}_{Q,P}\right)_{P \in \mathrm{SYT}(\lambda)}$ of the Young natural basis spans the right ideal $\mathbf{E}_Q\mathcal{A}$. In particular, this means that its span is preserved by right multiplication with each $s_i$. Thus, the matrices $R_{\overrightarrow{\mathrm{ynb}}}(s_i)$ for all $i \in [n-1]$ (and, more generally, the matrices $R_{\overrightarrow{\mathrm{ynb}}}(\mathbf{a})$ for all $\mathbf{a} \in \mathcal{A}$)

become block-diagonal if the rows and the columns are ordered in such a way that the standard $n$-bitableaux $(\lambda, Q, P)$ with equal $Q$'s appear as a contiguous block.

- Finally, Theorem 5.15.26 shows that for each partition $\lambda$ of $n$, the subfamily $(\mathbf{E}_{U,V})_{U,V\in \mathrm{SYT}(\lambda)}$ of the Young natural basis spans the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_\lambda$, which is a two-sided ideal of $\mathcal{A}$ (an easy consequence of $\mathbf{E}_\lambda \in Z(\mathcal{A})$).

Even better, the entries of the matrices $L\underrightarrow{}_{\mathrm{ynb}}(\mathbf{a})$ can be described in terms of the straightening coefficients (the coefficients of the $\mathbb{Z}$-linear combination in Proposition 5.9.14):

**Proposition 6.4.3.** Let $(\lambda, U, V) \in \mathrm{SBT}(n)$. Let $\mathbf{a} \in \mathcal{A}$. Write the element $\mathbf{a}\mathbf{e}_U$ in the Specht module $\mathcal{S}^\lambda$ as a $\mathbb{Z}$-linear combination

$$\mathbf{a}\mathbf{e}_U = \sum_{Q\in \mathrm{SYT}(\lambda)} \alpha_{U,Q}\mathbf{e}_Q \tag{481}$$

of standard polytabloids $\mathbf{e}_Q$. (This can be done, by Proposition 5.9.14.) Then,

$$\mathbf{a}\mathbf{E}_{U,V} = \sum_{Q\in \mathrm{SYT}(\lambda)} \alpha_{U,Q}\mathbf{E}_{Q,V}. \tag{482}$$

This holds for arbitrary commutative rings $\mathbf{k}$, even if $(\mathbf{E}_{U,V})_{(\lambda,U,V)\in \mathrm{SBT}(n)}$ is not a basis of $\mathcal{A}$.

*Proof.* Rename the $n$-tableau $V$ as $T$. Set $D = Y(\lambda)$, so that $\mathrm{SYT}(\lambda) = \mathrm{SYT}(D)$; hence, in particular, $U$ and $V$ belong to $\mathrm{SYT}(D)$. Now, recall the proof of Corollary 5.15.9, in which we constructed a left $\mathcal{A}$-module isomorphism $\overline{\alpha} : \mathcal{A}\mathbf{E}_T \to \mathcal{S}^D$. The inverse $\overline{\alpha}^{-1}$ of this isomorphism is thus a left $\mathcal{A}$-module isomorphism $\overline{\alpha}^{-1} : \mathcal{S}^D \to \mathcal{A}\mathbf{E}_T$. Applying this isomorphism $\overline{\alpha}^{-1}$ to the equality (481), we obtain

$$\mathbf{a}\cdot\overline{\alpha}^{-1}(\mathbf{e}_U) = \sum_{Q\in \mathrm{SYT}(\lambda)} \alpha_{U,Q}\overline{\alpha}^{-1}(\mathbf{e}_Q) \tag{483}$$

(since $\overline{\alpha}^{-1}$ is left $\mathcal{A}$-linear). But the equality (273) (which we have shown in the above-mentioned proof) says that $\overline{\alpha}^{-1}(\mathbf{e}_P) = \mathbf{E}_{P,T}$ for each $P \in \mathrm{SYT}(D)$. In other words, $\overline{\alpha}^{-1}(\mathbf{e}_P) = \mathbf{E}_{P,T}$ for each $P \in \mathrm{SYT}(\lambda)$ (since $\mathrm{SYT}(D) = \mathrm{SYT}(\lambda)$). Thus, we can rewrite the equality (483) as

$$\mathbf{a}\mathbf{E}_{U,T} = \sum_{Q\in \mathrm{SYT}(\lambda)} \alpha_{U,Q}\mathbf{E}_{Q,T}.$$

This proves Proposition 6.4.3 (since $T = V$). □

The equality (482) can be rewritten as

$$\mathbf{a}\mathbf{E}_{U,V} = \sum_{(\mu,Q,P)\in \mathrm{SBT}(n)} \alpha_{(\lambda,U,V),(\mu,Q,P)} \mathbf{E}_{Q,P}, \tag{484}$$

$$\text{where } \alpha_{(\lambda,U,V),(\mu,Q,P)} = \begin{cases} \alpha_{U,Q}, & \text{if } \lambda = \mu \text{ and } V = P; \\ 0, & \text{else} \end{cases}$$

(since all addends on the right hand side of (484) are 0 except for those that satisfy $\lambda = \mu$ and $V = P$; but the latter are just the addends on the right hand side of (482)).

The coefficients $\alpha_{(\lambda,U,V),(\mu,Q,P)}$ on the right hand side of (484) are thus the entries of the matrix $L_{\overrightarrow{\mathrm{ynb}}}(\mathbf{a})$. This shows once again that the matrix $L_{\overrightarrow{\mathrm{ynb}}}(\mathbf{a})$ is block-diagonal after permuting its rows and columns, and moreover explains why its entries are integers (since the $\alpha_{U,Q}$ are integers). Furthermore, while the diagonal blocks of $L_{\overrightarrow{\mathrm{ynb}}}(\mathbf{a})$ are indexed by the pairs $(\lambda, V)$ of a partition $\lambda$ and a standard tableau $V$, the actual entries of these blocks depend only on $\lambda$ (not on $V$), since they are the integers $\alpha_{U,Q}$.

The matrices $R_{\overrightarrow{\mathrm{ynb}}}(\mathbf{a})$ can be explained similarly using the "right Specht modules" (the right $\mathcal{A}$-module analogues of the Specht modules). Alternatively, they can be obtained from the $L_{\overrightarrow{\mathrm{ynb}}}(\mathbf{a})$ by conjugating with the $\mathbf{k}$-algebra anti-automorphism $T_{\mathrm{sign}} \circ S$ (which is represented by a rather simple matrix in the basis $\overrightarrow{\mathrm{ynb}}$, thanks to Proposition 5.15.16[170]). Indeed, it is easy to see that each $\mathbf{a} \in \mathcal{A}$ satisfies the equality of endomorphisms

$$\begin{aligned} R(\mathbf{a}) &= S^{-1} \circ L(S(\mathbf{a})) \circ S \\ &= \left(T_{\mathrm{sign}} \circ S\right)^{-1} \circ L\left(T_{\mathrm{sign}}(S(\mathbf{a}))\right) \circ \left(T_{\mathrm{sign}} \circ S\right) \end{aligned}$$

and thus the equality of matrices

$$R_{\overrightarrow{\mathrm{ynb}}}(\mathbf{a}) = \Omega^{-1} \cdot L_{\overrightarrow{\mathrm{ynb}}}\left(T_{\mathrm{sign}}(S(\mathbf{a}))\right) \cdot \Omega, \qquad \text{where } \Omega := \left[T_{\mathrm{sign}} \circ S\right]_{\overrightarrow{\mathrm{ynb}} \to \overrightarrow{\mathrm{ynb}}}.$$

When $\mathbf{a} = s_i$ for some $i \in [n-1]$, then $S(\mathbf{a}) = S(s_i) = s_i^{-1} = s_i$ and thus $T_{\mathrm{sign}}(S(\mathbf{a})) = T_{\mathrm{sign}}(s_i) = -s_i$, so this equality becomes even simpler:

$$R_{\overrightarrow{\mathrm{ynb}}}(s_i) = -\Omega^{-1} \cdot L_{\overrightarrow{\mathrm{ynb}}}(s_i) \cdot \Omega.$$

Since $\Omega$ is a permutation matrix "up to sign", this ensures that the entries of $R_{\overrightarrow{\mathrm{ynb}}}(s_i)$ agree with those of $L_{\overrightarrow{\mathrm{ynb}}}(s_i)$ up to sign and position. It is worth noting that $\Omega^{-1} = \Omega$, since $T_{\mathrm{sign}} \circ S$ is an involution (by Exercise 3.11.4 **(b)**).

[170] specifically, a permutation matrix with some of the 1's replaced by $-1$'s

## 6.5. The $\mathbf{F}_U w_{U,V} \mathbf{E}_V$ basis of $\mathbf{k}[S_n]$, revisited

For the present Section 6.5, we again assume that $n!$ is invertible in $\mathbf{k}$. Then, Theorem 5.21.2 shows that $(\mathbf{F}_U w_{U,V} \mathbf{E}_V)_{\lambda \text{ is a partition of } n, \text{ and } U,V \in \mathrm{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}$ (where the notations we use are defined as in Convention 5.21.1). Let us call this basis the *FwE basis* of $\mathcal{A}$. In view of Definition 6.1.1 **(e)**, we can rewrite the condition "$\lambda$ is a partition of $n$, and $U, V \in \mathrm{SYT}(\lambda)$" as "$(\lambda, U, V)$ is a standard $n$-bitableau", or – even shorter – as "$(\lambda, U, V) \in \mathrm{SBT}(n)$". Thus, the FwE basis can be written as

$$(\mathbf{F}_U w_{U,V} \mathbf{E}_V)_{(\lambda,U,V) \in \mathrm{SBT}(n)}.$$

Let us denote this basis by $\overrightarrow{\mathrm{fwe}}$.

Proposition 5.21.4 shows that the antipode $S$ of $\mathcal{A}$ permutes this basis, sending its $(\lambda, U, V)$-entry to its $(\lambda, V, U)$-entry. The actions of the operators $T_{\mathrm{sign}}$, $L(\mathbf{a})$ and $R(\mathbf{a})$ (for $\mathbf{a} \in \mathbf{k}[S_n]$) on this basis are far less simple.

**Example 6.5.1.** Let $n = 3$. Let us agree to order the indexing set $\mathrm{SBT}(3)$ of the FwE basis as follows:

$((3), 123, 123)$, $((2,1), 12\backslash\backslash 3, 12\backslash\backslash 3)$, $((2,1), 12\backslash\backslash 3, 13\backslash\backslash 2)$, $((2,1), 13\backslash\backslash 2, 12\backslash\backslash 3)$, $((2,1), 13\backslash\backslash 2, 13\backslash\backslash 2)$, $((1,1,1), 1\backslash\backslash 2\backslash\backslash 3, 1\backslash\backslash 2\backslash\backslash 3)$.

We have already listed the six entries of this basis in Example 5.21.3. The matrices representing the $\mathbf{k}$-linear maps $S$, $T_{\mathrm{sign}}$, $L(s_1)$, $L(s_2)$, $R(s_1)$ and $R(s_2)$ are as follows:

$$[S]_{\overrightarrow{\mathrm{fwe}} \to \overrightarrow{\mathrm{fwe}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix};$$

$$\left[T_{\mathrm{sign}}\right]_{\overrightarrow{\mathrm{fwe}} \to \overrightarrow{\mathrm{fwe}}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1/3 & -2/3 & -2/3 & 4/3 & 0 \\ 0 & 2/3 & -1/3 & -4/3 & 2/3 & 0 \\ 0 & 2/3 & -4/3 & -1/3 & 2/3 & 0 \\ 0 & 4/3 & -2/3 & -2/3 & 1/3 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix};$$

$$L_{\overrightarrow{\mathrm{fwe}}}(s_1) = [L(s_1)]_{\overrightarrow{\mathrm{fwe}} \to \overrightarrow{\mathrm{fwe}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

$$L_{\overrightarrow{\text{fwe}}}(s_2) = [L(s_2)]_{\overrightarrow{\text{fwe}} \to \overrightarrow{\text{fwe}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

$$R_{\overrightarrow{\text{fwe}}}(s_1) = [R(s_1)]_{\overrightarrow{\text{fwe}} \to \overrightarrow{\text{fwe}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

$$R_{\overrightarrow{\text{fwe}}}(s_2) = [R(s_2)]_{\overrightarrow{\text{fwe}} \to \overrightarrow{\text{fwe}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

Note that the last three of these six matrices agree with those for the Young natural basis (which we saw in Example 6.4.1)! The equality $L_{\overrightarrow{\text{fwe}}}(s_2) = L_{\overrightarrow{\text{ynb}}}(s_2)$ is a fluke (it fails for $n = 5$), but the equalities $R_{\overrightarrow{\text{fwe}}}(s_i) = R_{\overrightarrow{\text{ynb}}}(s_i)$ do indeed hold for all $n \in \mathbb{N}$ and $i \in [n-1]$. This follows from the following exercise:

**Exercise 6.5.1.** [3] Show that $R_{\overrightarrow{\text{fwe}}}(\mathbf{a}) = R_{\overrightarrow{\text{ynb}}}(\mathbf{a})$ for all $n \in \mathbb{N}$ and $\mathbf{a} \in \mathcal{A}$.

[**Hint:** For each partition $\lambda$ of $n$ and each standard tableau $U \in \operatorname{SYT}(\lambda)$, construct a right $\mathcal{A}$-module morphism

$$\begin{aligned} \Xi_U : \mathbf{E}_U\mathcal{A} &\to \mathbf{F}_U\mathcal{A}, \\ \mathbf{x} &\mapsto \mathbf{F}_U\mathbf{x}. \end{aligned}$$

Show that this morphism sends the basis vectors $\mathbf{E}_{U,V}$ of $\overrightarrow{\text{ynb}}$ to the respective basis vectors $\mathbf{F}_U w_{U,V} \mathbf{E}_V$ of $\overrightarrow{\text{fwe}}$. Now apply it to an expansion of the form $\mathbf{E}_{U,V}\mathbf{a} = \sum_{Q \in \operatorname{SYT}(\lambda)} \beta_{U,Q}\mathbf{E}_{U,Q}$, which exists thanks to (280).]

Again, the matrices $L_{\overrightarrow{\text{fwe}}}(\mathbf{a})$ and $R_{\overrightarrow{\text{fwe}}}(\mathbf{a})$ for all $\mathbf{a} \in \mathcal{A}$ are block-diagonal, after appropriate permutations of rows and columns. We leave it to the reader to explain this. The two types of matrices are connected by the equality

$$R_{\overrightarrow{\text{fwe}}}(\mathbf{a}) = \Omega_{\overrightarrow{\text{fwe}}}^{-1} \cdot L_{\overrightarrow{\text{fwe}}}(S(\mathbf{a})) \cdot \Omega_{\overrightarrow{\text{fwe}}}, \qquad \text{where } \Omega_{\overrightarrow{\text{fwe}}} := [S]_{\overrightarrow{\text{fwe}} \to \overrightarrow{\text{fwe}}}$$

(why?), which is particularly nice because $\Omega_{\overrightarrow{\text{fwe}}}$ is just a permutation matrix (since the antipode $S$ permutes the basis $\overrightarrow{\text{fwe}}$) and satisfies $\Omega_{\overrightarrow{\text{fwe}}}^{-1} = \Omega_{\overrightarrow{\text{fwe}}}$ (since $S$ is an involution).

## 6.6. The Artin–Wedderburn–Specht bases

One might believe that the block-diagonality properties make the bases $\overrightarrow{\text{ynb}}$ and $\overrightarrow{\text{fwe}}$ special; but they don't. There is a general machine for constructing bases of $\mathcal{A}$ with such properties, assuming that $n!$ is invertible in $\mathbf{k}$:

For the present Section 6.6, we again assume that $n!$ is invertible in $\mathbf{k}$. The Artin–Wedderburn theorem (Theorem 5.14.1) yields that the map

$$\rho : \mathcal{A} \to \prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right),$$
$$\mathbf{a} \mapsto \left(\rho_{\lambda}\left(\mathbf{a}\right)\right)_{\lambda \text{ is a partition of } n}$$

is a $\mathbf{k}$-algebra isomorphism. But each Specht module $\mathcal{S}^{\lambda}$ is a free $\mathbf{k}$-module of rank $f^{\lambda}$ (by Lemma 5.9.17); hence, its endomorphism algebra $\operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right)$ is a free $\mathbf{k}$-module of rank $\left(f^{\lambda}\right)^2$ (by Lemma 5.14.5). Thus, we can construct a basis of $\mathcal{A}$ as follows:

Step 1: Pick any basis of $\operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right)$ for each partition $\lambda$. The simplest way to do so is to pick a basis $\left(\mathbf{b}_T\right)_{T \in \operatorname{SYT}(\lambda)}$ of $\mathcal{S}^{\lambda}$ (indexed by standard tableaux because $\mathcal{S}^{\lambda}$ is free of rank $f^{\lambda} = |\operatorname{SYT}(\lambda)|$; theoretically, other indexing sets are possible, but I have never seen a such), and then construct the "elementary matrix basis" $\left(\mathbf{M}_{U,V}\right)_{(U,V) \in \operatorname{SYT}(\lambda) \times \operatorname{SYT}(\lambda)}$ of $\operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right)$ induced by it (explicitly: $\mathbf{M}_{U,V}$ is the $\mathbf{k}$-module endomorphism of $\mathcal{S}^{\lambda}$ that sends $\mathbf{b}_V$ to $\mathbf{b}_U$ and sends all $\mathbf{b}_Q$ with $Q \neq V$ to 0).

Step 2: Combine all these bases $\left(\mathbf{M}_{U,V}\right)_{(U,V) \in \operatorname{SYT}(\lambda) \times \operatorname{SYT}(\lambda)}$ (for all partitions $\lambda$ of $n$) into a basis $\left(\mathbf{M}_{\lambda,U,V}\right)_{(\lambda,U,V) \in \operatorname{SBT}(n)}$ of $\prod_{\lambda \text{ is a partition of } n} \operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right)$.

Step 3: Then, apply $\rho^{-1}$ to this basis – that is, set $\mathbf{M}'_{\lambda,U,V} := \rho^{-1}\left(\mathbf{M}_{\lambda,U,V}\right)$ for each $(\lambda, U, V) \in \operatorname{SBT}(n)$. This yields a basis $\left(\mathbf{M}'_{\lambda,U,V}\right)_{(\lambda,U,V) \in \operatorname{SBT}(n)}$ of $\mathcal{A}$.

Bases of $\mathcal{A}$ obtained in such a way have some properties in common. In particular, they all rely on the invertibility of $n!$ (since $\rho$ is usually not an isomorphism otherwise), and they are all indexed by the standard $n$-bitableaux $(\lambda, U, V)$. Their subfamilies often span ideals: For each given partition $\lambda$ of $n$, the span of $\left\{\mathbf{M}'_{\lambda,U,V} \mid (U,V) \in \operatorname{SYT}(\lambda) \times \operatorname{SYT}(\lambda)\right\}$ is the two-sided ideal $\mathcal{A}\mathbf{E}_{\lambda}$ of $\mathcal{A}$. Moreover, if the underlying bases $\left(\mathbf{M}_{U,V}\right)_{(U,V) \in \operatorname{SYT}(\lambda) \times \operatorname{SYT}(\lambda)}$ of

$\operatorname{End}_{\mathbf{k}}\left(\mathcal{S}^{\lambda}\right)$ are "elementary matrix bases" (see Step 1 above), then the span of $\left\{\mathbf{M}'_{\lambda,U,V} \mid V \in \operatorname{SYT}(\lambda)\right\}$ for fixed $\lambda$ and $U$ is a right ideal of $\mathcal{A}$, while the span of $\left\{\mathbf{M}'_{\lambda,U,V} \mid U \in \operatorname{SYT}(\lambda)\right\}$ for fixed $\lambda$ and $V$ is a left ideal of $\mathcal{A}$. All of this is rather easy to prove. I refer to such bases as *Artin–Wedderburn bases*.

Let us apply the above construction, picking the easiest option in Step 1: As our basis $(\mathbf{b}_T)_{T\in\operatorname{SYT}(\lambda)}$ of $\mathcal{S}^{\lambda}$, we choose the basis $(\mathbf{e}_T)_{T\in\operatorname{SYT}(\lambda)}$ consisting of the standard polytabloids (see Theorem 5.9.1). We then construct the "elementary matrix basis" $(\mathbf{M}_{U,V})_{(U,V)\in\operatorname{SYT}(\lambda)\times\operatorname{SYT}(\lambda)}$, and proceed as explained above. Let us denote the resulting basis $\left(\mathbf{M}'_{\lambda,U,V}\right)_{(\lambda,U,V)\in\operatorname{SBT}(n)}$ of $\mathcal{A}$ by $(\mathbf{W}_{U,V})_{(\lambda,U,V)\in\operatorname{SBT}(n)}$ (we omit the mention of $\lambda$ since $U$ uniquely determines $\lambda$) and by $\overrightarrow{\text{aws}}$ (short for "Artin–Wedderburn–Specht" or "Artin–Wedderburn–standard"). Explicitly, $\mathbf{W}_{U,V}$ can be characterized as the unique element of $\mathcal{A}$ that satisfies $\mathbf{W}_{U,V}\mathbf{e}_V = \mathbf{e}_U$ in the Specht module $\mathcal{S}^{\lambda}$ (where $V \in \operatorname{SYT}(\lambda)$) and satisfies $\mathbf{W}_{U,V}\mathbf{e}_W = 0$ for all standard $n$-tableaux $W \neq V$ of partition shapes.

**Example 6.6.1.** For $n = 3$, this basis $\overrightarrow{\text{aws}} = (\mathbf{W}_{U,V})_{(\lambda,U,V)\in\operatorname{SBT}(n)}$ of $\mathcal{A}$ consists of the vectors

$$\begin{aligned}
\mathbf{W}_{123,\ 123} &= \frac{1}{6}\nabla = \frac{1}{6} + \frac{1}{6}s_1 + \frac{1}{6}s_2 + \frac{1}{6}t_{1,3} + \frac{1}{6}\operatorname{cyc}_{1,2,3} + \frac{1}{6}\operatorname{cyc}_{1,3,2};\\
\mathbf{W}_{12\backslash\backslash 3,\ 12\backslash\backslash 3} &= \frac{1}{3} + \frac{1}{3}s_1 - \frac{1}{3}t_{1,3} - \frac{1}{3}\operatorname{cyc}_{1,2,3};\\
\mathbf{W}_{12\backslash\backslash 3,\ 13\backslash\backslash 2} &= -\frac{1}{3}s_1 + \frac{1}{3}s_2 + \frac{1}{3}\operatorname{cyc}_{1,2,3} - \frac{1}{3}\operatorname{cyc}_{1,3,2};\\
\mathbf{W}_{13\backslash\backslash 2,\ 12\backslash\backslash 3} &= \frac{1}{3}s_2 - \frac{1}{3}t_{1,3} - \frac{1}{3}\operatorname{cyc}_{1,2,3} + \frac{1}{3}\operatorname{cyc}_{1,3,2};\\
\mathbf{W}_{13\backslash\backslash 2,\ 13\backslash\backslash 2} &= \frac{1}{3} - \frac{1}{3}s_1 + \frac{1}{3}t_{1,3} - \frac{1}{3}\operatorname{cyc}_{1,3,2};\\
\mathbf{W}_{1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3} &= \frac{1}{6}\nabla^{-} = \frac{1}{6} - \frac{1}{6}s_1 - \frac{1}{6}s_2 - \frac{1}{6}t_{1,3} + \frac{1}{6}\operatorname{cyc}_{1,2,3} + \frac{1}{6}\operatorname{cyc}_{1,3,2}.
\end{aligned}$$

The attentive reader will have certainly noticed that these vectors are just scalar multiples of the respective entries of the Young natural basis $\overrightarrow{\text{ynb}} = (\mathbf{E}_{U,V})_{(\lambda,U,V)\in\operatorname{SBT}(n)}$ from Example 6.4.1! The scalars are $\frac{1}{6}$ for the $n$-bitableaux $(\lambda, U, V)$ with $\lambda = (3)$ or $\lambda = (1,1,1)$ and are $\frac{1}{3}$ for those with $\lambda = (2,1)$. Thus, the matrices representing the $\mathbf{k}$-linear maps $S$, $T_{\text{sign}}$, $L(s_1)$, $L(s_2)$, $R(s_1)$ and $R(s_2)$ with respect to the basis $\overrightarrow{\text{aws}}$ are exactly the same as for the basis $\overrightarrow{\text{ynb}}$. (Normally, the fact that two different scalars are involved would cause some entries to be rescaled; but this is immaterial to us here, since these entries are all 0.)

Alas, this pattern does not generalize. The vectors $\mathbf{W}_{U,V}$ are scalar multiples

of the $\mathbf{E}_{U,V}$ for all $n \leq 4$, but not for $n > 4$. Here is an easy way to see the latter: For any $n \in \mathbb{N}$, the basis $\overrightarrow{\text{aws}} = (\mathbf{W}_{U,V})_{(\lambda,U,V)\in \mathrm{SBT}(n)}$ has the property that its entries multiply like elementary matrices:

$$\begin{aligned} &\mathbf{W}_{P,Q}\mathbf{W}_{Q,R} = \mathbf{W}_{P,R}; \\ &\mathbf{W}_{P,Q}\mathbf{W}_{T,R} = 0 \qquad \text{whenever } Q \neq T. \end{aligned}$$

Up to a scalar factor, the Young natural basis $\overrightarrow{\text{ynb}} = (\mathbf{E}_{U,V})_{(\lambda,U,V)\in \mathrm{SBT}(n)}$ also satisfies the former of these two equalities (by Exercise 5.15.2), but not the latter, unless $n \leq 4$ (see Remark 5.11.17 and Exercise 5.11.2, and recall that $\mathbf{E}_P = \mathbf{E}_{P,P}$). Thus, the two bases cannot be equal – even up to scalar – for $n \geq 5$.

Nevertheless, the matrices representing the $\mathbf{k}$-linear maps $L(\mathbf{a})$ for all $\mathbf{a} \in \mathcal{A}$ with respect to the basis $\overrightarrow{\text{aws}}$ agree with those for the basis $\overrightarrow{\text{ynb}}$ for general $n$:

**Exercise 6.6.1.** [3] Show that $L_{\overrightarrow{\text{aws}}}(\mathbf{a}) = L_{\overrightarrow{\text{ynb}}}(\mathbf{a})$ for all $n \in \mathbb{N}$ and $\mathbf{a} \in \mathcal{A}$.

[**Hint:** By linearity, it suffices to prove this for $\mathbf{a} = \mathbf{W}_{P,Q}$. Show that $\mathbf{W}_{P,Q}\mathbf{W}_{U,V} = \delta_{Q,U}\mathbf{W}_{P,V}$ and $\mathbf{W}_{P,Q}\mathbf{E}_{U,V} = \delta_{Q,U}\mathbf{E}_{P,V}$.]

However, $R_{\overrightarrow{\text{aws}}}(\mathbf{a})$ generally does not equal $R_{\overrightarrow{\text{ynb}}}(\mathbf{a})$ for $n \in \mathbb{N}$ and $\mathbf{a} \in \mathcal{A}$. For example, for $n = 5$, we have $R_{\overrightarrow{\text{aws}}}(s_1) \neq R_{\overrightarrow{\text{ynb}}}(s_1)$ as well as $[S]_{\overrightarrow{\text{aws}}\to\overrightarrow{\text{aws}}} \neq [S]_{\overrightarrow{\text{ynb}}\to\overrightarrow{\text{ynb}}}$ and $\left[T_{\text{sign}}\right]_{\overrightarrow{\text{aws}}\to\overrightarrow{\text{aws}}} \neq \left[T_{\text{sign}}\right]_{\overrightarrow{\text{ynb}}\to\overrightarrow{\text{ynb}}}$.

## 6.7. Combinatorial interlude: Lexicographic orders

We now take a break from algebra and introduce some total orders on the sets of partitions, of standard tableaux and of standard bitableaux.

### 6.7.1. Partitions

Recall that a partition is a weakly decreasing finite list of positive integers, such as $(3,2,2)$. The length of a partition $\lambda$ is denoted by $\ell(\lambda)$. If $\lambda$ is a partition and $i$ is a positive integer, then $\lambda_i$ denotes the $i$-th entry of $\lambda$; we understand this to mean 0 if $i > \ell(\lambda)$ (see Definition 5.1.1 **(h)**). Thus, in a sense, we are pretending that a partition $\lambda$ has infinitely many entries (rather than just $\ell(\lambda)$ many), starting with its $\ell(\lambda)$ actual entries and then continuing with infinitely many zeroes. This legal fiction will come convenient in the following definition:

**Definition 6.7.1.** Define a binary relation $<$ on the set of all partitions as follows:

Let $\lambda$ and $\mu$ be two partitions. Then, we declare that $\lambda < \mu$ if and only if

- there exists at least one $i \geq 1$ such that $\lambda_i \neq \mu_i$;

- the **smallest** such $i$ satisfies $\lambda_i < \mu_i$.

In other words, we say that $\lambda < \mu$ if the **leftmost** position in which $\lambda$ and $\mu$ differ is occupied by a smaller entry in $\lambda$ than in $\mu$.

For example:

- We have $(5,3,3,1) < (5,4,1)$, because the leftmost position in which the partitions $(5,3,3,1)$ and $(5,4,1)$ differ is the second position, and the former partition has a smaller entry in this position than the latter partition does (since $3 < 4$).
- We have $(5,2) < (5,2,1)$, since the leftmost position in which the partitions $(5,2)$ and $(5,2,1)$ differ is the third position (recall that we pretend that the third entry of the partition $(5,2)$ is 0), and the former partition has a smaller entry in this position than the latter partition does (since $0 < 1$).
- The empty partition $\varnothing = ()$ satisfies $\varnothing < \lambda$ for any partition $\lambda \neq \varnothing$, because all entries of $\varnothing$ are 0.

We shall prove the following fact shortly:

**Proposition 6.7.2.** The relation $<$ defined in Definition 6.7.1 is the smaller relation of a total order on the set of all partitions.

**Definition 6.7.3.** This total order is called the *lexicographic order on partitions*. We shall use it whenever we need a total order on the set of all partitions (unless we explicitly introduce a different order). Thus, the symbol $<$ when applied to partitions will always be understood as in Definition 6.7.1. Moreover, the symbols $\leq$, $>$ and $\geq$ will be understood accordingly (e.g., we let "$\lambda \leq \mu$" mean "$\lambda = \mu$ or $\lambda < \mu$", and we let "$\lambda > \mu$" mean "$\mu < \lambda$").

**Example 6.7.4.** **(a)** Here are all seven partitions of 5, ordered in lexicographic order:

$$(1,1,1,1,1) < (2,1,1,1) < (2,2,1) < (3,1,1) < (3,2) < (4,1) < (5).$$

**(b)** Here is a chain of some partitions ordered in lexicographic order:

$$() < (1) < (1,1) < (2) < (2,1) < (2,2) < (2,2,1) < (3) < (3,1).$$

It is easy to extend this chain to the right. We can also interject any number of further partitions between $(1)$ and $(2)$: namely,

$$(1) < (1,1) < (1,1,1) < (1,1,1,1) < \cdots < (2).$$

*Proof of Proposition 6.7.2 (sketched).* For a detailed proof, see the Appendix (Section A.5). In a nutshell: The relation $<$ is one of many instances of a lexicographic order (i.e., a way of comparing tuples by first comparing the first entries, then resolving ties using the second entries, then resolving any remaining ties using the third entries, and so on). The only unusual feature of our current situation is that different partitions can have different lengths; but we consider them as having infinitely many entries (thanks to Definition 5.1.1 **(h)**). □

**Exercise 6.7.1.** $\boxed{1}$ Let $\lambda$ be a partition of $n$. Let $k \in [n]$. Prove that $\lambda_1 \geq k$ holds if and only if $\lambda \geq \left(k, \underbrace{1,1,\ldots,1}_{n-k \text{ times}}\right)$ in the lexicographic order on partitions.

**Exercise 6.7.2.** We shall use *exponential notation*, meaning that a sequence $\underbrace{p,p,\ldots,p}_{k \text{ times}}$ of equal numbers will be denoted by $p^k$ (not to be mistaken for the $k$-th power of $p$). Thus, for example, the partition $(4,4,3,1,1,1)$ can be written as $\left(4^2,3,1^3\right)$ or even as $\left(4^2,3^1,1^3\right)$.

Consider the lexicographic order on partitions, restricted to the set of all partitions of $n$.

**(a)** $\boxed{1}$ Prove that the smallest $\lfloor n/2 \rfloor + 1$ partitions in this order are

$$(1^n) < \left(2,1^{n-2}\right) < \left(2,2,1^{n-4}\right) < \cdots < \left(2^{\lfloor n/2 \rfloor},1\right) \text{ or } \left(2^{\lfloor n/2 \rfloor}\right)$$

(the last partition here is $\left(2^{\lfloor n/2 \rfloor},1\right)$ if $n$ is odd and $\left(2^{\lfloor n/2 \rfloor}\right)$ if $n$ is even).

**(b)** $\boxed{1}$ What are the largest four partitions in this order (assuming $n \geq 4$) ?

The lexicographic order is just one of many orders that can be imposed on partitions. Several other orders are known (both total and partial orders), such as the colexicographic order, the dominance order (see, e.g., [GriRei14, Definition 2.2.7], [JamKer81, §1.4] or [Wildon18, Definition 4.1]) and the refinement order (see, e.g., [Macdon95, §I.6, around (6.10)]). But we shall have no use for these other orders in the present chapter.

### 6.7.2. Standard tableaux of partition shapes

Next, we move on from partitions to tableaux.

If $\kappa$ is a partition and if $U$ is an $n$-tableau of shape $Y(\kappa)$, then we can uniquely recover $\kappa$ from $U$ (in fact, each entry $\kappa_i$ of $\kappa$ is the number of entries in the $i$-th row of $U$). Hence, the sets $\operatorname{SYT}(\kappa)$ for different partitions $\kappa$ are disjoint. Therefore, the union $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ is a disjoint union. (Recall that "$\kappa \vdash n$" means

"$\kappa$ is a partition of $n$".) Explicitly, the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ consists of all standard $n$-tableaux of partition shape (i.e., of shape $Y(\kappa)$ for some partition $\kappa$ of $n$). This set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ is finite (since all sets $\operatorname{SYT}(\kappa)$ are finite, and there are finitely many partitions $\kappa$ of $n$).

For example, for $n = 3$, the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ consists of the four standard tableaux 123, 12\\3, 13\\2, 1\\2\\3 (where we are using the shorthand notation from Convention 5.2.4).

Now we shall define a total order on the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ as follows:

**Definition 6.7.5.** We define a binary relation $<$ on the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ as follows:

Let $P, Q \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ be two $n$-tableaux. Thus, $P \in \operatorname{SYT}(\lambda)$ and $Q \in \operatorname{SYT}(\mu)$ for some partitions $\lambda$ and $\mu$ of $n$. Consider these two partitions $\lambda$ and $\mu$. (They are uniquely determined by $P$ and $Q$, since the sets $\operatorname{SYT}(\kappa)$ for different partitions $\kappa$ are disjoint.) Consider also the $n$-tabloids $\overline{P}$ and $\overline{Q}$ of shapes $Y(\lambda)$ and $Y(\mu)$. Now, we declare that $P < Q$ if and only if

$$\left(\lambda < \mu \text{ or } \left(\lambda = \mu \text{ and } \overline{P} < \overline{Q}\right)\right),$$

where:

- the "$<$" sign in "$\lambda < \mu$" refers to the lexicographic order on partitions (defined in Definition 6.7.1);
- the "$<$" sign in "$\overline{P} < \overline{Q}$" refers to the Young last letter order on the set $\{n\text{-tabloids of shape } Y(\lambda)\}$ (defined in Definition 5.7.5 **(b)**), which makes sense here because the preceding statement $\lambda = \mu$ ensures that both $\overline{P}$ and $\overline{Q}$ are $n$-tabloids of shape $Y(\lambda)$ (in fact, $\overline{Q}$ is an $n$-tabloid of shape $Y(\mu)$, but because of $\lambda = \mu$, this is the same as being an $n$-tabloid of shape $Y(\lambda)$).

For example, for $n = 5$:

- We have $\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1} 5 \\ \cline{1-1}\end{array} < \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2}\end{array}$, because the shapes $(3,1,1)$ and $(3,2)$ of these two tableaux satisfy $(3,1,1) < (3,2)$.
- We have $\begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 \\ \cline{1-2}\end{array} < \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2}\end{array}$, because the shapes $(3,2)$ and $(3,2)$ of

these two tableaux are equal whereas the corresponding $n$-tabloids satisfy $\begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4\\\cline{1-2}\end{array} < \begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5\\\cline{1-2}\end{array}$.

We shall prove the following fact shortly:

**Proposition 6.7.6.** The relation $<$ defined in Definition 6.7.5 is the smaller relation of a total order on the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$.

**Definition 6.7.7.** We shall use this total order whenever we need a total order on the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ (unless we explicitly introduce a different order). Thus, the symbol $<$ when applied to standard $n$-tableaux will always be understood as in Definition 6.7.5. Moreover, the symbols $\leq$, $>$ and $\geq$ will be understood accordingly (e.g., we let "$P \leq Q$" mean "$P = Q$ or $P < Q$", and we let "$P > Q$" mean "$Q < P$").

**Example 6.7.8.** For $n = 3$, here are the elements of the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ listed in the order we just introduced:

$$\begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array} < \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array} < \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array} < \begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array}.$$

*Proof of Proposition 6.7.6.* See the appendix (Section A.5) for a detailed proof. Here is a quick outline: Again, the relation $<$ is a kind of lexicographic order, although not in the most direct way, since it is defined not on tuples but rather on single standard tableaux. To reveal its lexicographic nature, we need to re-encode any standard $n$-tableau $T \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ as the pair $\left(\lambda, \overline{T}\right)$, where $\lambda$ is the partition of $n$ satisfying $T \in \operatorname{SYT}(\lambda)$, and where $\overline{T}$ is the $n$-tabloid of $T$. This encoding $T \mapsto \left(\lambda, \overline{T}\right)$ is injective, since each standard (or even just row-standard) $n$-tableau $T$ can be uniquely reconstructed from its $n$-tabloid $\overline{T}$ (see Proposition 5.3.8). Once we replace the tableaux $T \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ by the corresponding pairs $\left(\lambda, \overline{T}\right)$ (where $\lambda$ is the partition such that $T \in \operatorname{SYT}(\lambda)$), our definition of the relation $<$ takes the form

$$\left(\left(\lambda, \overline{P}\right) < \left(\mu, \overline{Q}\right)\right) \iff \left(\lambda < \mu \text{ or } \left(\lambda = \mu \text{ and } \overline{P} < \overline{Q}\right)\right);$$

but this is just a lexicographic order on pairs, and thus is well-known to be a total order. So the original relation $<$ is a total order as well. $\square$

The total order we constructed on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ has a rather convenient property:

**Lemma 6.7.9** (Combined zero-sandwich lemma)**.** Let $P$ and $Q$ be two standard $n$-tableaux in $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$. Assume that $P < Q$ (where the relation $<$ is the one defined in Definition 6.7.5). Then,

$$\nabla^{-}_{\operatorname{Col} P} \nabla_{\operatorname{Row} Q} = 0 \qquad \text{and} \qquad \nabla_{\operatorname{Row} Q} \nabla^{-}_{\operatorname{Col} P} = 0.$$

*Proof.* We have $P \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$. That is, we have $P \in \operatorname{SYT}(\lambda)$ for some partition $\lambda$ of $n$. Similarly, we have $Q \in \operatorname{SYT}(\mu)$ for some partition $\mu$ of $n$. Consider these two partitions $\lambda$ and $\mu$. Thus, $P$ is an $n$-tableau of shape $Y(\lambda)$, whereas $Q$ is an $n$-tableau of shape $Y(\mu)$.

Recall that $P < Q$. By Definition 6.7.5, this means that

$$\left(\lambda < \mu \text{ or } \left(\lambda = \mu \text{ and } \overline{P} < \overline{Q}\right)\right)$$

(since $P \in \operatorname{SYT}(\lambda)$ and $Q \in \operatorname{SYT}(\mu)$). Hence, we are in one of the following two cases:

*Case 1:* We have $\lambda < \mu$.

*Case 2:* We have $\lambda = \mu$ and $\overline{P} < \overline{Q}$.

Let us first consider Case 1. In this case, we have $\lambda < \mu$. By Definition 6.7.1, this means that

- there exists at least one $i \geq 1$ such that $\lambda_i \neq \mu_i$;
- the **smallest** such $i$ satisfies $\lambda_i < \mu_i$.

Consider this smallest $i$. Thus, $i \geq 1$ and $\lambda_i < \mu_i$, but each positive integer $j < i$ satisfies

$$\lambda_j = \mu_j \tag{485}$$

(since $i$ is the **smallest** $i \geq 1$ such that $\lambda_i \neq \mu_i$). Adding the equalities (485) for all $j \in \{1, 2, \ldots, i-1\}$ together, we find

$$\lambda_1 + \lambda_2 + \cdots + \lambda_{i-1} = \mu_1 + \mu_2 + \cdots + \mu_{i-1}.$$

Adding the inequality $\lambda_i < \mu_i$ to this equality, we find

$$(\lambda_1 + \lambda_2 + \cdots + \lambda_{i-1}) + \lambda_i < (\mu_1 + \mu_2 + \cdots + \mu_{i-1}) + \mu_i.$$

In other words,

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i < \mu_1 + \mu_2 + \cdots + \mu_i.$$

In other words, $\mu_1 + \mu_2 + \cdots + \mu_i > \lambda_1 + \lambda_2 + \cdots + \lambda_i$. Therefore:

- Theorem 5.10.8 **(b)** (applied to $\mu$, $\lambda$, $Q$ and $P$ instead of $\lambda$, $\mu$, $S$ and $T$) yields $\nabla^{-}_{\operatorname{Col} P}\mathbf{a}\nabla_{\operatorname{Row} Q} = 0$ for each $\mathbf{a} \in \mathbf{k}[S_n]$. Applying this to $\mathbf{a} = 1$, we obtain $\nabla^{-}_{\operatorname{Col} P} 1 \nabla_{\operatorname{Row} Q} = 0$. In other words, $\nabla^{-}_{\operatorname{Col} P}\nabla_{\operatorname{Row} Q} = 0$.
- Theorem 5.10.8 **(d)** (applied to $\mu$, $\lambda$, $Q$ and $P$ instead of $\lambda$, $\mu$, $S$ and $T$) yields $\nabla_{\operatorname{Row} Q}\mathbf{a}\nabla^{-}_{\operatorname{Col} P} = 0$ for each $\mathbf{a} \in \mathbf{k}[S_n]$. Applying this to $\mathbf{a} = 1$, we obtain $\nabla_{\operatorname{Row} Q} 1 \nabla^{-}_{\operatorname{Col} P} = 0$. In other words, $\nabla_{\operatorname{Row} Q}\nabla^{-}_{\operatorname{Col} P} = 0$.

Hence, the proof of Lemma 6.7.9 is complete in Case 1.

Now, let us consider Case 2. In this case, we have $\lambda = \mu$ and $\overline{P} < \overline{Q}$. From $\lambda = \mu$, we obtain $\operatorname{SYT}(\lambda) = \operatorname{SYT}(\mu)$, so that $Q \in \operatorname{SYT}(\mu) = \operatorname{SYT}(\lambda)$. Hence, both tableaux $P$ and $Q$ belong to $\operatorname{SYT}(\lambda)$, that is, are $n$-tableaux of shape $Y(\lambda)$. Moreover, they are standard, hence column-standard. Now, recall that $\overline{P} < \overline{Q}$ in the Young last letter order. Hence, Lemma 5.10.9 **(c)** (applied to $S = Q$ and $T = P$) yields

$$\nabla^{-}_{\operatorname{Col} P}\nabla_{\operatorname{Row} Q} = 0 \qquad \text{and} \qquad \nabla_{\operatorname{Row} Q}\nabla^{-}_{\operatorname{Col} P} = 0.$$

Thus, Lemma 6.7.9 is proved in Case 2.

We have now proved Lemma 6.7.9 in both Cases 1 and 2. Hence, the proof is complete. □

### 6.7.3. Standard $n$-bitableaux

Finally, we define a total order on the set $\operatorname{SBT}(n)$ of all standard $n$-bitableaux. We recall that any standard $n$-bitableau $(\lambda, U, V) \in \operatorname{SBT}(n)$ satisfies $U, V \in \operatorname{SYT}(\lambda) \subseteq \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$.

**Definition 6.7.10.** We define a binary relation $<$ on the set $\operatorname{SBT}(n)$ as follows:

For two standard $n$-bitableaux $(\lambda, A, B)$ and $(\mu, C, D)$ in $\operatorname{SBT}(n)$, we declare that $(\lambda, A, B) < (\mu, C, D)$ if and only if

$$(A < C \text{ or } (A = C \text{ and } B < D)).$$

Here, the symbol "$<$" in "$A < C$" and "$B < D$" refers to the binary relation $<$ introduced in Definition 6.7.5.

For example, for $n = 5$:

- We have

$$\left((3,2), \begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 \\ \cline{1-2}\end{array}, \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 & 5 \\ \cline{1-2}\end{array}\right) < \left((3,2), \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2}\end{array}, \begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 \\ \cline{1-2}\end{array}\right),$$

because the first tableaux in these two $n$-bitableaux satisfy $\begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4\\\cline{1-2}\end{array} < \begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5\\\cline{1-2}\end{array}$.

- We have

$$\left((3,2), \begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2&5\\\cline{1-2}\end{array}, \begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4\\\cline{1-2}\end{array}\right) < \left((3,2), \begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2&5\\\cline{1-2}\end{array}, \begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5\\\cline{1-2}\end{array}\right),$$

because the first tableaux in these two $n$-bitableaux are equal whereas the second tableaux satisfy $\begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4\\\cline{1-2}\end{array} < \begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5\\\cline{1-2}\end{array}$.

We shall prove the following fact shortly:

**Proposition 6.7.11.** The relation $<$ defined in Definition 6.7.10 is the smaller relation of a total order on the set $\mathrm{SBT}(n)$.

**Definition 6.7.12.** We shall use this total order whenever we need a total order on the set $\mathrm{SBT}(n)$ (unless we explicitly introduce a different order). Thus, the symbol $<$ when applied to standard $n$-bitableaux will always be understood as in Definition 6.7.10. Moreover, the symbols $\leq$, $>$ and $\geq$ will be understood accordingly (e.g., we let "$\mathfrak{P} \leq \mathfrak{Q}$" mean "$\mathfrak{P} = \mathfrak{Q}$ or $\mathfrak{P} < \mathfrak{Q}$").

**Example 6.7.13.** For $n = 3$, here are the six standard $n$-bitableaux in $\mathrm{SBT}(n)$ listed in the order we just introduced:

$$\left((1,1,1), \begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\cline{1-1}\end{array}, \begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\cline{1-1}\end{array}\right),$$

$$\left((2,1), \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}\right), \qquad \left((2,1), \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}\right),$$

$$\left((2,1), \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}\right), \qquad \left((2,1), \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}\right),$$

$$\left((3), \begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array}, \begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array}\right).$$

*Proof of Proposition 6.7.11.* See the appendix (Section A.5) for a detailed proof. But again, this relation is just a lexicographic order in disguise. Indeed, a standard $n$-bitableau $(\lambda, U, V) \in \operatorname{SBT}(n)$ can be re-encoded as the pair $(U, V) \in \left(\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)\right)^2$, because the partition $\lambda$ is uniquely determined by the tableau $U$. But the condition that defines the binary relation $<$ on $\operatorname{SBT}(n)$, namely

$$(A < C \text{ or } (A = C \text{ and } B < D)),$$

is precisely the lexicographic inequality $(A, B) < (C, D)$ between the latter pairs. □

Let us state another simple property of the relation $<$ from Definition 6.7.10:

**Lemma 6.7.14.** Let $(\lambda, A, B)$ and $(\mu, C, D)$ be two standard $n$-bitableaux in $\operatorname{SBT}(n)$ satisfying $(\lambda, A, B) < (\mu, C, D)$ (with respect to the relation $<$ from Definition 6.7.10). Then:

**(a)** We have $A \leq C$. (Here, the symbol "$\leq$" refers to the total order on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ introduced in Definition 6.7.5.)

**(b)** We have $A < C$ or $B < D$. (Here, the symbol "$<$" refers to the total order on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ introduced in Definition 6.7.5.)

*Proof.* See the appendix (Section A.5) for the very simple proof. □

## 6.8. The Murphy bases

We now come to some new bases of $\mathcal{A}$. The *Murphy bases* (also known as the *Murphy cellular bases*) are two closely related bases of $\mathcal{A}$ that share a few of the nice properties of the Young natural basis (Section 6.4), the FwE basis (Section 6.5) and the Artin–Wedderburn bases (Section 6.6), although in a weakened form (the matrices representing the $L(\mathbf{a})$ and $R(\mathbf{a})$ maps shall only be block-triangular, not block-diagonal). In exchange, the bases are much simpler to define, and exist **for every commutative ring k**. In a sense, these bases are the best one can hope for without assuming that $n!$ be invertible in $\mathbf{k}$.

### 6.8.1. Introductory example: $n = 3$

We have already seen one of the Murphy bases in Subsection 4.4.4 for $n = 3$. This was the basis

$$\overrightarrow{\mathrm{rmu}}_3 := (\nabla,\ \nabla_{3,3},\ \nabla_{2,3},\ \nabla_{3,2},\ \nabla_{2,2},\ 1),$$

where we set

$$\nabla_{b,a} := \sum_{\substack{w \in S_n; \\ w(a) = b}} w \qquad \text{for all } a, b \in [n].$$

We called this basis the "row Murphy basis" of $\mathbf{k}[S_3]$. It consists of the six vectors

$$\begin{aligned}
\nabla &= 1 + s_1 + s_2 + t_{1,3} + \mathrm{cyc}_{1,2,3} + \mathrm{cyc}_{1,3,2}, \\
\nabla_{3,3} &= 1 + s_1, \\
\nabla_{2,3} &= s_2 + \mathrm{cyc}_{1,3,2}, \\
\nabla_{3,2} &= s_2 + \mathrm{cyc}_{1,2,3}, \\
\nabla_{2,2} &= 1 + t_{1,3}, \\
1 &= 1.
\end{aligned}$$

The matrices representing the $\mathbf{k}$-linear maps $S$, $T_{\mathrm{sign}}$, $L(s_1)$, $L(s_2)$, $R(s_1)$ and $R(s_2)$ with respect to this basis are as follows:

$$[S]_{\overrightarrow{\mathrm{rmú}} \to \overrightarrow{\mathrm{rmú}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix};$$

$$\left[T_{\mathrm{sign}}\right]_{\overrightarrow{\mathrm{rmú}} \to \overrightarrow{\mathrm{rmú}}} = \begin{pmatrix} 3 & 0 & 2 & 2 & 0 & 0 \\ -4 & -1 & -2 & -2 & 0 & 0 \\ -2 & 0 & -1 & -2 & 0 & 0 \\ -2 & 0 & -2 & -1 & 0 & 0 \\ -4 & 0 & -2 & -2 & -1 & 0 \\ 6 & 2 & 2 & 2 & 2 & 1 \end{pmatrix};$$

$$L_{\overrightarrow{\mathrm{rmú}}}(s_1) = [L(s_1)]_{\overrightarrow{\mathrm{rmú}} \to \overrightarrow{\mathrm{rmú}}} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

$$L_{\overrightarrow{\mathrm{rmú}}}(s_2) = [L(s_2)]_{\overrightarrow{\mathrm{rmú}} \to \overrightarrow{\mathrm{rmú}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

$$R_{\overrightarrow{\mathrm{rmú}}}(s_1) = [R(s_1)]_{\overrightarrow{\mathrm{rmú}} \to \overrightarrow{\mathrm{rmú}}} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix};$$

$$R_{\overrightarrow{\text{rmu}}}(s_2) = \left[R(s_2)\right]_{\overrightarrow{\text{rmu}} \to \overrightarrow{\text{rmu}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

We observe a few properties:

- All the matrices have integer entries. This is because the basis is defined over any commutative ring $\mathbf{k}$, including $\mathbb{Z}$.
- The matrix $[S]_{\overrightarrow{\text{rmu}} \to \overrightarrow{\text{rmu}}}$ is a permutation matrix. In other words, the antipode $S$ permutes the row Murphy basis.
- The matrix $\left[T_{\text{sign}}\right]_{\overrightarrow{\text{rmu}} \to \overrightarrow{\text{rmu}}}$ is not so nice. Indeed, the sign-twist $T_{\text{sign}}$ does not permute the row Murphy basis, but rather sends it to (a permuted) column Murphy basis, which we will soon define as well.
- The matrices $L_{\overrightarrow{\text{rmu}}}(\mathbf{a})$ and $R_{\overrightarrow{\text{rmu}}}(\mathbf{a})$ are not block-diagonal any more, but they are block-upper-triangular (if the rows/columns are appropriately ordered).

We will now define a row Murphy basis for all $n \in \mathbb{N}$, and then show that it satisfies all these properties.

### 6.8.2. Definitions

We shall define the Murphy bases soon. First, we recall the notions of an $n$-tabloid (Definition 5.3.6) and of an $n$-column-tabloid (Definition 5.9.3). For any diagram $D$, we introduced a left $S_n$-action on the set $\{n\text{-tabloids of shape } D\}$ in Definition 5.3.10. We now make a similar definition for $n$-column-tabloids:

**Definition 6.8.1.** Let $D$ be any diagram. Then, we define a left $S_n$-action on the set $\{n\text{-column-tabloids of shape } D\}$ by setting

$$w \rightharpoonup \widetilde{T} := \widetilde{w \rightharpoonup T} = \widetilde{w \circ T} \qquad \text{for all } w \in S_n \text{ and all } n\text{-tableaux } T.$$

This action is well-defined (we will justify this in a moment), and is again called the *action on entries* (for column-tabloids). This action makes the set $\{n\text{-column-tabloids of shape } D\}$ into a left $S_n$-set.

The reason why this action is well-defined is the same as for the action in Definition 5.3.10 (except that now, "row-equivalence" must be replaced by "column-equivalence").

Next, we define a lot of elements of $\mathcal{A}$, some (not all) of which will constitute the Murphy bases:

**Definition 6.8.2.** Let $P$ and $Q$ be two $n$-tableaux of the same shape. Then:

**(a)** We define the *row-to-row sum* $\nabla_{P,Q}^{\mathrm{Row}}$ to be the element

$$\nabla_{P,Q}^{\mathrm{Row}} := \sum_{\substack{u \in S_n; \\ u\overline{Q} = \overline{P}}} u \in \mathcal{A} \tag{486}$$

(recall that $\overline{T}$ denotes the $n$-tabloid corresponding to the $n$-tableau $T$, as defined in Definition 5.3.6).

**(b)** We define the *column-to-column signed sum* $\nabla_{P,Q}^{-\mathrm{Col}}$ by

$$\nabla_{P,Q}^{-\mathrm{Col}} = \sum_{\substack{u \in S_n; \\ u\widetilde{Q} = \widetilde{P}}} (-1)^u u \in \mathcal{A} \tag{487}$$

(recall that $\widetilde{T}$ denotes the $n$-column-tabloid corresponding to the $n$-tableau $T$, as defined in Definition 5.9.3).

Using the notations $\mathrm{Row}\,(i,T)$ and $\mathrm{Col}\,(j,T)$ from Proposition 5.5.3, we can rewrite the definitions of $\nabla_{P,Q}^{\mathrm{Row}}$ and $\nabla_{P,Q}^{-\mathrm{Col}}$ in the following more concrete way:

**Definition 6.8.3.** For any $n$-tableau $T$ and any $i \in \mathbb{Z}$, we let $\mathrm{Row}\,(i,T)$ denote the set of all entries in the $i$-th row of $T$.

For any $n$-tableau $T$ and any $j \in \mathbb{Z}$, we let $\mathrm{Col}\,(j,T)$ denote the set of all entries in the $j$-th column of $T$.

**Proposition 6.8.4.** Let $P$ and $Q$ be two $n$-tableaux of the same shape. Then,

$$\nabla_{P,Q}^{\mathrm{Row}} = \sum_{\substack{u \in S_n; \\ u(\mathrm{Row}(i,Q)) = \mathrm{Row}(i,P) \text{ for all } i \in \mathbb{Z}}} u \tag{488}$$

and

$$\nabla_{P,Q}^{-\mathrm{Col}} = \sum_{\substack{u \in S_n; \\ u(\mathrm{Col}(j,Q)) = \mathrm{Col}(j,P) \text{ for all } j \in \mathbb{Z}}} (-1)^u u. \tag{489}$$

**Example 6.8.5.** Let $P = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array}$ and $Q = \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 3 \\ \hline \end{array}$. Then, we can use Proposition 6.8.4 to compute $\nabla_{P,Q}^{\mathrm{Row}}$ and $\nabla_{P,Q}^{-\mathrm{Col}}$ as follows: Let us write $\mathrm{oln}\,(i_1 i_2 \ldots i_n)$ for the permutation of $[n]$ that sends $1, 2, \ldots, n$ to $i_1, i_2, \ldots, i_n$, respectively. Then, $\mathrm{Row}\,(1,Q) = \{1,4\}$ and $\mathrm{Col}\,(1,Q) = \{1,2\}$, and so on. Thus, (488) yields

$$\nabla_{P,Q}^{\mathrm{Row}} = \sum_{\substack{u \in S_n; \\ u(\{1,4\}) = \{1,2\}; \\ u(\{2,3\}) = \{3,4\}}} u = \mathrm{oln}\,(1342) + \mathrm{oln}\,(1432) + \mathrm{oln}\,(2341) + \mathrm{oln}\,(2431)\,,$$

whereas (489) yields

$$\begin{aligned}\nabla_{P,Q}^{-\mathrm{Col}} &= \sum_{\substack{u\in S_n;\\ u(\{1,2\})=\{1,3\};\\ u(\{4,3\})=\{2,4\}}} (-1)^u\, u \\ &= -\operatorname{oln}(1324)+\operatorname{oln}(1342)+\operatorname{oln}(3124)-\operatorname{oln}(3142).\end{aligned}$$

To prove Proposition 6.8.4, we need a simple combinatorial lemma:

**Lemma 6.8.6.** Let $U$ and $V$ be two $n$-tableaux of the same shape. Then:

**(a)** We have the logical equivalence

$$\left(\overline{U}=\overline{V}\right) \iff \left(\operatorname{Row}(i,U)=\operatorname{Row}(i,V) \text{ for all } i\in\mathbb{Z}\right).$$

**(b)** We have the logical equivalence

$$\left(\widetilde{U}=\widetilde{V}\right) \iff \left(\operatorname{Col}(j,U)=\operatorname{Col}(j,V) \text{ for all } j\in\mathbb{Z}\right).$$

The proof of Lemma 6.8.6 should be completely routine by now; it can be found in the appendix (Section A.6).

*Proof of Proposition 6.8.4.* The core of the proof are the following two claims:

*Claim 1:* Let $u\in S_n$. Then, we have the logical equivalence

$$\left(u\overline{Q}=\overline{P}\right) \iff \left(u\left(\operatorname{Row}(i,Q)\right)=\operatorname{Row}(i,P) \text{ for all } i\in\mathbb{Z}\right).$$

*Claim 2:* Let $u\in S_n$. Then, we have the logical equivalence

$$\left(u\widetilde{Q}=\widetilde{P}\right) \iff \left(u\left(\operatorname{Col}(j,Q)\right)=\operatorname{Col}(j,P) \text{ for all } j\in\mathbb{Z}\right).$$

*Proof of Claim 1.* Definition 5.3.10 yields $u\rightharpoonup\overline{Q}=\overline{u\rightharpoonup Q}$.

Let $i\in\mathbb{Z}$. By Definition 5.3.9, each entry of the tableau $u\rightharpoonup Q$ is obtained from the corresponding entry of $Q$ by applying the permutation $u$. Hence, in particular, the set of all entries in the $i$-th row of $u\rightharpoonup Q$ is obtained from the set of all entries in the $i$-th row of $Q$ by applying the permutation $u$ to each element. In other words,

$$\operatorname{Row}(i,u\rightharpoonup Q)=u\left(\operatorname{Row}(i,Q)\right) \tag{490}$$

(since the former set is $\operatorname{Row}(i,u\rightharpoonup Q)$ whereas the latter set is $\operatorname{Row}(i,Q)$).

Forget that we fixed $i$. We thus have proved the equality (490) for each $i\in\mathbb{Z}$.

Now, we have the following chain of logical equivalences:

$$\begin{aligned}
\left(u\overline{Q} = \overline{P}\right) &\iff \left(\overline{u \rightharpoonup Q} = \overline{P}\right) \qquad \left(\text{since } u\overline{Q} = u \rightharpoonup \overline{Q} = \overline{u \rightharpoonup Q}\right) \\
&\iff \left(\operatorname{Row}(i, u \rightharpoonup Q) = \operatorname{Row}(i, P) \text{ for all } i \in \mathbb{Z}\right) \\
&\qquad \left(\text{by Lemma 6.8.6 } \textbf{(a)}\text{, applied to } U = u \rightharpoonup Q \text{ and } V = P\right) \\
&\iff \left(u\left(\operatorname{Row}(i, Q)\right) = \operatorname{Row}(i, P) \text{ for all } i \in \mathbb{Z}\right)
\end{aligned}$$

(by (490)). This proves Claim 1. □

*Proof of Claim 2.* This is analogous to Claim 1 (using Lemma 6.8.6 **(b)** instead of Lemma 6.8.6 **(a)**). □

Now, (486) becomes

$$\nabla_{P,Q}^{\operatorname{Row}} = \sum_{\substack{u \in S_n; \\ u\overline{Q} = \overline{P}}} u = \sum_{\substack{u \in S_n; \\ u(\operatorname{Row}(i,Q)) = \operatorname{Row}(i,P) \text{ for all } i \in \mathbb{Z}}} u$$

(here, we have replaced the condition "$u\overline{Q} = \overline{P}$" under the summation sign by the condition "$u\left(\operatorname{Row}(i, Q)\right) = \operatorname{Row}(i, P)$ for all $i \in \mathbb{Z}$", which is equivalent according to Claim 1). Thus, (488) is proved. Likewise, we can obtain (489) from (487) using Claim 2. This proves Proposition 6.8.4. □

**Remark 6.8.7.** We have encountered the row-to-row sums $\nabla_{P,Q}^{\operatorname{Row}}$ before: Let $P$ and $Q$ be two $n$-tableaux of the same shape. Then, in Definition 5.5.15, we have defined $\nabla_{\overline{P},\overline{Q}} = \sum_{\substack{u \in S_n; \\ u\overline{Q} = \overline{P}}} u$. But in Definition 6.8.2, we have defined $\nabla_{P,Q}^{\operatorname{Row}} = \sum_{\substack{u \in S_n; \\ u\overline{Q} = \overline{P}}} u$. Comparing these two equalities, we obtain

$$\nabla_{\overline{P},\overline{Q}} = \nabla_{P,Q}^{\operatorname{Row}}. \tag{491}$$

### 6.8.3. Basic identities

We shall continue with a series of simple but useful properties of the $\nabla_{P,Q}^{\operatorname{Row}}$ and $\nabla_{P,Q}^{-\operatorname{Col}}$ elements. We begin by identifying the row-symmetrizers $\nabla_{\operatorname{Row} T}$ and the column-antisymmetrizers $\nabla_{\operatorname{Col} T}^{-}$ as particular cases of row-to-row sums and signed column-to-column sums, respectively:

**Proposition 6.8.8.** Let $T$ be an $n$-tableau of any shape. Then,

$$\nabla_{T,T}^{\operatorname{Row}} = \nabla_{\operatorname{Row} T} \qquad \text{and} \tag{492}$$

$$\nabla_{T,T}^{-\operatorname{Col}} = \nabla_{\operatorname{Col} T}^{-}. \tag{493}$$

*Proof.* The equality (492) is just Proposition 5.5.17 (since (491) tells us that $\nabla_{\overline{T},\overline{T}}$ is precisely our row-to-row sum $\nabla^{\mathrm{Row}}_{T,T}$). Thus, let us prove (493).

Definition 5.5.1 yields $\nabla^{-}_{\mathrm{Col}\,T} = \sum_{w \in \mathcal{C}(T)} (-1)^w w$. Recall from Proposition 5.3.13 **(a)** that $\mathcal{R}(T) = \left\{ w \in S_n \mid \overline{w \rightharpoonup T} = \overline{T} \right\}$. The same argument (but applied to columns instead of rows, and to $n$-column-tabloids instead of $n$-tabloids) shows that

$$\mathcal{C}(T) = \left\{ w \in S_n \mid \widetilde{w \rightharpoonup T} = \widetilde{T} \right\}.$$

In other words,

$$\mathcal{C}(T) = \left\{ w \in S_n \mid w\widetilde{T} = \widetilde{T} \right\}$$

(since $\widetilde{w \rightharpoonup T} = w \rightharpoonup \widetilde{T} = w\widetilde{T}$ for each $w \in S_n$). Thus, the summation sign $\sum_{w \in \mathcal{C}(T)}$ can be rewritten as $\sum_{\substack{w \in S_n; \\ w\widetilde{T} = \widetilde{T}}}$. Hence,

$$\sum_{w \in \mathcal{C}(T)} (-1)^w w = \sum_{\substack{w \in S_n; \\ w\widetilde{T} = \widetilde{T}}} (-1)^w w = \sum_{\substack{u \in S_n; \\ u\widetilde{T} = \widetilde{T}}} (-1)^u u$$

(here, we have renamed the summation index $w$ as $u$). Therefore,

$$\nabla^{-}_{\mathrm{Col}\,T} = \sum_{w \in \mathcal{C}(T)} (-1)^w w = \sum_{\substack{u \in S_n; \\ u\widetilde{T} = \widetilde{T}}} (-1)^u u = \nabla^{-\mathrm{Col}}_{T,T}$$

(since $\nabla^{-\mathrm{Col}}_{T,T}$ was defined as $\sum_{\substack{u \in S_n; \\ u\widetilde{T} = \widetilde{T}}} (-1)^u u$). This proves (493). The proof of (492) is analogous (but, as we said, it was already done when we proved Proposition 5.5.17). Thus, Proposition 6.8.8 is proved. □

Next, we prove an equivariance property:

**Proposition 6.8.9.** Let $P$ and $Q$ be two $n$-tableaux of the same shape. Let $g \in S_n$ and $h \in S_n$ be two permutations. Then,

$$\nabla^{\mathrm{Row}}_{gP,hQ} = g \nabla^{\mathrm{Row}}_{P,Q} h^{-1} \qquad \text{and} \tag{494}$$

$$\nabla^{-\mathrm{Col}}_{gP,hQ} = (-1)^g (-1)^h g \nabla^{-\mathrm{Col}}_{P,Q} h^{-1}. \tag{495}$$

*Proof.* Definition 6.8.1 yields $h \rightharpoonup \widetilde{Q} = \widetilde{h \rightharpoonup Q}$. In other words, $h\widetilde{Q} = \widetilde{hQ}$. Thus, $\widetilde{hQ} = h\widetilde{Q}$. Similarly, $\widetilde{gP} = g\widetilde{P}$.

For any permutation $u \in S_n$, we have the following chain of logical equivalences:

$$\begin{aligned}
&\left(\left(guh^{-1}\right)\widetilde{hQ} = \widetilde{gP}\right) \\
&\iff \left(\left(guh^{-1}\right)h\widetilde{Q} = g\widetilde{P}\right) \qquad \left(\text{since } \widetilde{hQ} = h\widetilde{Q} \text{ and } \widetilde{gP} = g\widetilde{P}\right) \\
&\iff \left(gu\widetilde{Q} = g\widetilde{P}\right) \qquad \left(\text{since } \left(guh^{-1}\right)h\widetilde{Q} = gu\underbrace{h^{-1}h}_{=1}\widetilde{Q} = gu\widetilde{Q}\right) \\
&\iff \left(u\widetilde{Q} = \widetilde{P}\right) \qquad (496)
\end{aligned}$$

(here, we have cancelled the "factor" $g$, since the action of $g$ on the set of all $n$-column-tabloids is a permutation and thus injective). Now, the definition of $\nabla^{-\mathrm{Col}}_{gP,hQ}$ yields

$$\begin{aligned}
\nabla^{-\mathrm{Col}}_{gP,hQ} &= \sum_{\substack{u\in S_n; \\ u\widetilde{hQ}=\widetilde{gP}}} (-1)^u\, u \\
&= \sum_{\substack{u\in S_n; \\ \left(guh^{-1}\right)\widetilde{hQ}=\widetilde{gP}}} (-1)^{guh^{-1}}\, guh^{-1} \qquad (497)
\end{aligned}$$

(here, we have substituted $guh^{-1}$ for $u$ in the sum, since the map $S_n \to S_n,\ u \mapsto guh^{-1}$ is a bijection (its inverse is the map $S_n \to S_n,\ v \mapsto g^{-1}vh$)).

On the other hand, $(-1)^h \in \{1,-1\}$, so that $\left((-1)^h\right)^{-1} = (-1)^h$ (since each $x \in \{1,-1\}$ satisfies $x^{-1} = x$). Next, recall that the sign homomorphism

$$\begin{aligned}
S_n &\to \{1,-1\}, \\
\sigma &\mapsto (-1)^\sigma
\end{aligned}$$

is a group morphism. Thus, each $u \in S_n$ satisfies

$$(-1)^{guh^{-1}} = (-1)^g\,(-1)^u\,\underbrace{\left((-1)^h\right)^{-1}}_{=(-1)^h} = (-1)^g\,(-1)^u\,(-1)^h.$$

Hence, we can rewrite (497) as

$$\begin{aligned}
\nabla_{gP,hQ}^{-\mathrm{Col}} &= \sum_{\substack{u\in S_n; \\ \left(guh^{-1}\right)\widetilde{hQ}=\widetilde{gP}}} (-1)^g (-1)^u (-1)^h guh^{-1} \\
&= \sum_{\substack{u\in S_n; \\ u\widetilde{Q}=\widetilde{P}}} (-1)^g (-1)^u (-1)^h guh^{-1} \\
&\qquad \left( \begin{array}{c} \text{here, we have replaced the condition ``} \left(guh^{-1}\right)\widetilde{hQ} = \widetilde{gP}\text{''} \\ \text{under the summation sign by ``} u\widetilde{Q} = \widetilde{P}\text{'',} \\ \text{which is equivalent because of (496)} \end{array} \right) \\
&= (-1)^g (-1)^h g \underbrace{\left( \sum_{\substack{u\in S_n; \\ u\widetilde{Q}=\widetilde{P}}} (-1)^u u \right)}_{\substack{=\nabla_{P,Q}^{-\mathrm{Col}} \\ \text{(by (487))}}} h^{-1} = (-1)^g (-1)^h g \nabla_{P,Q}^{-\mathrm{Col}} h^{-1}.
\end{aligned}$$

This proves (495). A similar argument (but easier due to absence of signs) proves (494). Hence, Proposition 6.8.9 is established. □

The row-to-row sums and the column-to-column signed sums are actually just shifted versions of row symmetrizers and the column antisymmetrizers:

**Lemma 6.8.10** (permutation lemma)**.** Let $P$ and $Q$ be two $n$-tableaux of the same shape. Let $u \in S_n$ be a permutation such that $P = uQ$. Then,

$$\nabla_{P,Q}^{\mathrm{Row}} = \nabla_{\mathrm{Row}\, P} u = u \nabla_{\mathrm{Row}\, Q} \tag{498}$$

and

$$\nabla_{P,Q}^{-\mathrm{Col}} = (-1)^u \nabla_{\mathrm{Col}\, P}^{-} u = (-1)^u u \nabla_{\mathrm{Col}\, Q}^{-}. \tag{499}$$

*Proof.* We observe that $(-1)^u \in \{1, -1\}$ and thus $\left((-1)^u\right)^{-1} = (-1)^u$ (since $x^{-1} = x$ for each $x \in \{1, -1\}$).

Applying (495) to $g = \mathrm{id}$ and $h = u$, we obtain

$$\nabla_{\mathrm{id}\, P,uQ}^{-\mathrm{Col}} = \underbrace{(-1)^{\mathrm{id}}}_{=1} (-1)^u \underbrace{\mathrm{id}}_{=1} \nabla_{P,Q}^{-\mathrm{Col}} u^{-1} = (-1)^u \nabla_{P,Q}^{-\mathrm{Col}} u^{-1}.$$

In view of $\mathrm{id}\, P = P$ and $uQ = P$, we can rewrite this as $\nabla_{P,P}^{-\mathrm{Col}} = (-1)^u \nabla_{P,Q}^{-\mathrm{Col}} u^{-1}$. Comparing this with $\nabla_{P,P}^{-\mathrm{Col}} = \nabla_{\mathrm{Col}\, P}^{-}$ (which follows from (493), applied to

$T = P$), we obtain $(-1)^u \nabla^{-\mathrm{Col}}_{P,Q} u^{-1} = \nabla^{-}_{\mathrm{Col}\, P}$. Solving this for $\nabla^{-\mathrm{Col}}_{P,Q}$, we find

$$\nabla^{-\mathrm{Col}}_{P,Q} = \underbrace{\frac{1}{(-1)^u}}_{=\left((-1)^u\right)^{-1}=(-1)^u} \nabla^{-}_{\mathrm{Col}\, P} u = (-1)^u \nabla^{-}_{\mathrm{Col}\, P} u. \tag{500}$$

This proves the first equality sign in (499).

On the other hand, from $P = uQ$, we obtain $Q = u^{-1}P$. But (495) (applied to $g = u^{-1}$ and $h = \mathrm{id}$) yields

$$\nabla^{-\mathrm{Col}}_{u^{-1}P,\mathrm{id}\, Q} = \underbrace{(-1)^{u^{-1}}}_{=\left((-1)^u\right)^{-1}} \underbrace{(-1)^{\mathrm{id}}}_{=1} u^{-1} \nabla^{-\mathrm{Col}}_{P,Q} \underbrace{\mathrm{id}^{-1}}_{=\mathrm{id}=1} = \left((-1)^u\right)^{-1} u^{-1} \nabla^{-\mathrm{Col}}_{P,Q}.$$

In view of $u^{-1}P = Q$ and $\mathrm{id}\, Q = Q$, we can rewrite this equality as $\nabla^{-\mathrm{Col}}_{Q,Q} = \left((-1)^u\right)^{-1} u^{-1} \nabla^{-\mathrm{Col}}_{P,Q}$. Comparing this with $\nabla^{-\mathrm{Col}}_{Q,Q} = \nabla^{-}_{\mathrm{Col}\, Q}$ (which follows from (493), applied to $T = Q$), we obtain $\left((-1)^u\right)^{-1} u^{-1} \nabla^{-\mathrm{Col}}_{P,Q} = \nabla^{-}_{\mathrm{Col}\, Q}$. Solving this for $\nabla^{-\mathrm{Col}}_{P,Q}$, we find

$$\nabla^{-\mathrm{Col}}_{P,Q} = (-1)^u\, u \nabla^{-}_{\mathrm{Col}\, Q}. \tag{501}$$

Combining this with (500), we obtain the equality (499).

An analogous argument (but without any signs, and using (492) and (494)) yields (498). Thus, Lemma 6.8.10 is proved. □

We can easily describe the actions of $S$ and $T_{\mathrm{sign}}$ on the $\nabla^{\mathrm{Row}}_{P,Q}$ and $\nabla^{-\mathrm{Col}}_{P,Q}$ elements:

**Proposition 6.8.11.** Let $P$ and $Q$ be two $n$-tableaux of the same shape. Then,

$$S\left(\nabla^{\mathrm{Row}}_{P,Q}\right) = \nabla^{\mathrm{Row}}_{Q,P} \qquad \text{and} \tag{502}$$

$$S\left(\nabla^{-\mathrm{Col}}_{P,Q}\right) = \nabla^{-\mathrm{Col}}_{Q,P}. \tag{503}$$

*Proof.* Let $D$ be the shape of the $n$-tableaux $P$ and $Q$. Let $g$ be the permutation $w_{P,Q} \in S_n$ as in Definition 5.15.2. Thus, (269) says that $w_{P,Q} Q = P$. In other words, $gQ = P$ (since we defined $g$ to be $w_{P,Q}$), so that $P = gQ$. Note also that $(-1)^{g^{-1}} = (-1)^g$ (since any permutation $\sigma \in S_n$ satisfies $(-1)^{\sigma^{-1}} = (-1)^\sigma$).

Proposition 5.5.10 shows that every $n$-tableau $T$ of shape $D$ satisfies $S\left(\nabla^{-}_{\mathrm{Col}\, T}\right) = \nabla^{-}_{\mathrm{Col}\, T}$. Applying this to $T = P$, we obtain $S\left(\nabla^{-}_{\mathrm{Col}\, P}\right) = \nabla^{-}_{\mathrm{Col}\, P}$.

Now, $P = gQ$; hence, (499) (applied to $u = g$) yields

$$\nabla^{-\mathrm{Col}}_{P,Q} = (-1)^g \nabla^{-}_{\mathrm{Col}\, P} g = (-1)^g\, g \nabla^{-}_{\mathrm{Col}\, Q}.$$

Hence, in particular, $\nabla_{P,Q}^{-\mathrm{Col}} = (-1)^g \nabla_{\mathrm{Col}\,P}^{-} g$. Applying the map $S$ to this equality, we find

$$S\left(\nabla_{P,Q}^{-\mathrm{Col}}\right) = S\left((-1)^g \nabla_{\mathrm{Col}\,P}^{-} g\right) = (-1)^g \cdot S(g) \cdot S\left(\nabla_{\mathrm{Col}\,P}^{-}\right)$$

(since $S$ is a $\mathbf{k}$-algebra anti-automorphism (by Theorem 3.11.14 **(a)**)). Thus,

$$\begin{aligned} S\left(\nabla_{P,Q}^{-\mathrm{Col}}\right) &= (-1)^g \cdot \underbrace{S(g)}_{\substack{=g^{-1} \\ \text{(by the definition of } S\text{)}}} \cdot \underbrace{S\left(\nabla_{\mathrm{Col}\,P}^{-}\right)}_{=\nabla_{\mathrm{Col}\,P}^{-}} \\ &= (-1)^g g^{-1} \nabla_{\mathrm{Col}\,P}^{-}. \end{aligned} \tag{504}$$

On the other hand, $Q = g^{-1}P$ (since $P = gQ$); thus, (498) (applied to $Q$, $P$ and $g^{-1}$ instead of $P$, $Q$ and $u$) yields

$$\nabla_{Q,P}^{-\mathrm{Col}} = (-1)^{g^{-1}} \nabla_{\mathrm{Col}\,Q}^{-} g^{-1} = (-1)^{g^{-1}} g^{-1} \nabla_{\mathrm{Col}\,P}^{-}.$$

Hence,

$$\nabla_{Q,P}^{-\mathrm{Col}} = \underbrace{(-1)^{g^{-1}}}_{=(-1)^g} g^{-1} \nabla_{\mathrm{Col}\,P}^{-} = (-1)^g g^{-1} \nabla_{\mathrm{Col}\,P}^{-}.$$

Comparing this with (504), we obtain $S\left(\nabla_{P,Q}^{-\mathrm{Col}}\right) = \nabla_{Q,P}^{-\mathrm{Col}}$. This proves (503). A similar argument (but without any signs) yields (502). Thus, Proposition 6.8.11 is proved. □

**Proposition 6.8.12.** Let $P$ and $Q$ be two $n$-tableaux of the same shape. Then,

$$T_{\mathrm{sign}}\left(\nabla_{P,Q}^{\mathrm{Row}}\right) = \nabla_{P\mathbf{r},Q\mathbf{r}}^{-\mathrm{Col}} \qquad \text{and} \tag{505}$$

$$T_{\mathrm{sign}}\left(\nabla_{P,Q}^{-\mathrm{Col}}\right) = \nabla_{P\mathbf{r},Q\mathbf{r}}^{\mathrm{Row}}. \tag{506}$$

(Recall from Definition 5.15.10 that $T\mathbf{r}$ denotes the reflection of a tableau $T$ across the main diagonal.)

*Proof.* Let $D$ be the shape of the $n$-tableaux $P$ and $Q$. Let $g$ be the permutation $w_{P,Q} \in S_n$ as in Definition 5.15.2. Thus, (269) says that $w_{P,Q}Q = P$. In other words, $gQ = P$ (since we defined $g$ to be $w_{P,Q}$), so that $P = gQ$. Note also that $\left((-1)^g\right)^2 = 1$ (since any permutation $\sigma \in S_n$ satisfies $\left((-1)^\sigma\right)^2 = 1$).

Now, $P = gQ$; hence, (499) (applied to $u = g$) yields

$$\nabla_{P,Q}^{-\mathrm{Col}} = (-1)^g \nabla_{\mathrm{Col}\,P}^{-} g = (-1)^g g \nabla_{\mathrm{Col}\,Q}^{-}.$$

Hence, in particular, $\nabla_{P,Q}^{-\mathrm{Col}} = (-1)^g \nabla_{\mathrm{Col}\, P}^{-} g$. Applying the map $T_{\mathrm{sign}}$ to this equality, we find

$$\begin{aligned} T_{\mathrm{sign}}\left(\nabla_{P,Q}^{-\mathrm{Col}}\right) &= T_{\mathrm{sign}}\left((-1)^g \nabla_{\mathrm{Col}\, P}^{-} g\right) \\ &= (-1)^g \cdot \underbrace{T_{\mathrm{sign}}\left(\nabla_{\mathrm{Col}\, P}^{-}\right)}_{\substack{=\nabla_{\mathrm{Row}(P\mathbf{r})} \\ \text{(by (276))}}} \cdot \underbrace{T_{\mathrm{sign}}(g)}_{\substack{=(-1)^g g \\ \text{(by the definition of } T_{\mathrm{sign}}\text{)}}} \\ &\qquad \left(\begin{array}{c} \text{since } T_{\mathrm{sign}} \text{ is a } \mathbf{k}\text{-algebra morphism} \\ \text{(by Theorem 3.11.5 } \mathbf{(a)}\text{)} \end{array}\right) \\ &= (-1)^g \cdot \nabla_{\mathrm{Row}(P\mathbf{r})} \cdot (-1)^g g = \underbrace{\left((-1)^g\right)^2}_{=1} \cdot \nabla_{\mathrm{Row}(P\mathbf{r})} g \\ &= \nabla_{\mathrm{Row}(P\mathbf{r})} g. \end{aligned} \tag{507}$$

On the other hand, Proposition 5.15.12 **(d)** (applied to $Q$ and $g$ instead of $T$ and $w$) yields $(gQ)\,\mathbf{r} = g\,(Q\mathbf{r})$. Since $P = gQ$, we can rewrite this as $P\mathbf{r} = g\,(Q\mathbf{r})$. Of course, this shows that the tableaux $P\mathbf{r}$ and $Q\mathbf{r}$ have the same shape. Thus, (498) (applied to $P\mathbf{r}$, $Q\mathbf{r}$ and $g$ instead of $P$, $Q$ and $u$) yields

$$\nabla_{P\mathbf{r},Q\mathbf{r}}^{\mathrm{Row}} = \nabla_{\mathrm{Row}(P\mathbf{r})} g = g \nabla_{\mathrm{Row}(Q\mathbf{r})}.$$

In particular, $\nabla_{P\mathbf{r},Q\mathbf{r}}^{\mathrm{Row}} = \nabla_{\mathrm{Row}(P\mathbf{r})} g$. Comparing this equality with (507), we obtain $T_{\mathrm{sign}}\left(\nabla_{P,Q}^{-\mathrm{Col}}\right) = \nabla_{P\mathbf{r},Q\mathbf{r}}^{\mathrm{Row}}$. This proves (506).

The equality (505) can be proved similarly. Alternatively, we can derive it from (506) easily: Apply (506) to $P\mathbf{r}$ and $Q\mathbf{r}$ instead of $P$ and $Q$. Thus we find

$$T_{\mathrm{sign}}\left(\nabla_{P\mathbf{r},Q\mathbf{r}}^{-\mathrm{Col}}\right) = \nabla_{(P\mathbf{r})\mathbf{r},(Q\mathbf{r})\mathbf{r}}^{\mathrm{Row}} = \nabla_{P,Q}^{\mathrm{Row}}$$

(since Proposition 5.15.12 **(b)** yields $(P\mathbf{r})\,\mathbf{r} = P$ and $(Q\mathbf{r})\,\mathbf{r} = Q$). Applying the map $T_{\mathrm{sign}}$ to this equality, we obtain

$$T_{\mathrm{sign}}\left(T_{\mathrm{sign}}\left(\nabla_{P\mathbf{r},Q\mathbf{r}}^{-\mathrm{Col}}\right)\right) = T_{\mathrm{sign}}\left(\nabla_{P,Q}^{\mathrm{Row}}\right).$$

Hence,

$$\begin{aligned} T_{\mathrm{sign}}\left(\nabla_{P,Q}^{\mathrm{Row}}\right) &= T_{\mathrm{sign}}\left(T_{\mathrm{sign}}\left(\nabla_{P\mathbf{r},Q\mathbf{r}}^{-\mathrm{Col}}\right)\right) = \underbrace{\left(T_{\mathrm{sign}} \circ T_{\mathrm{sign}}\right)}_{\substack{=\mathrm{id} \\ \text{(by Theorem 3.11.5 } \mathbf{(b)}\text{)}}} \left(\nabla_{P\mathbf{r},Q\mathbf{r}}^{-\mathrm{Col}}\right) \\ &= \mathrm{id}\left(\nabla_{P\mathbf{r},Q\mathbf{r}}^{-\mathrm{Col}}\right) = \nabla_{P\mathbf{r},Q\mathbf{r}}^{-\mathrm{Col}}. \end{aligned}$$

This proves (505). Thus, the proof of Proposition 6.8.12 is complete. □

For future use, let us prove another identity:

**Lemma 6.8.13.** Let $P$ and $Q$ be two $n$-tableaux of the same shape. Then,

$$\nabla_{P,Q}^{-\mathrm{Col}} \nabla_{Q,P}^{\mathrm{Row}} = \pm \mathbf{E}_P,$$

where $\mathbf{E}_P$ is the Young symmetrizer (as defined in Definition 5.11.1 **(b)**).

*Proof.* Let $D$ be the shape of the $n$-tableaux $P$ and $Q$. Then, Proposition 5.3.13 **(b)** (applied to $T = Q$) shows that any $n$-tableau of shape $D$ can be written as $w \rightharpoonup Q$ for some $w \in S_n$. Thus, in particular, $P$ can be written in this way. In other words, there exists some $w \in S_n$ such that $P = w \rightharpoonup Q$. Consider this $w$.

Thus, $P = w \rightharpoonup Q = wQ$. Hence, (499) (applied to $u = w$) yields

$$\nabla_{P,Q}^{-\mathrm{Col}} = (-1)^w \nabla_{\mathrm{Col}\, P}^{-} w = (-1)^w w \nabla_{\mathrm{Col}\, Q}^{-}.$$

Moreover, $Q = w^{-1}P$ (since $P = wQ$). Hence, (498) (applied to $Q$, $P$ and $w^{-1}$ instead of $P$, $Q$ and $u$) yields

$$\nabla_{Q,P}^{\mathrm{Row}} = \nabla_{\mathrm{Row}\, Q} w^{-1} = w^{-1} \nabla_{\mathrm{Row}\, P}.$$

Hence,

$$\begin{aligned} \underbrace{\nabla_{P,Q}^{-\mathrm{Col}}}_{=(-1)^w \nabla_{\mathrm{Col}\, P}^{-} w} \underbrace{\nabla_{Q,P}^{\mathrm{Row}}}_{=w^{-1}\nabla_{\mathrm{Row}\, P}} &= (-1)^w \nabla_{\mathrm{Col}\, P}^{-} \underbrace{ww^{-1}}_{=1} \nabla_{\mathrm{Row}\, P} \\ &= (-1)^w \underbrace{\nabla_{\mathrm{Col}\, P}^{-} \nabla_{\mathrm{Row}\, P}}_{=\mathbf{E}_P \text{ (by the definition of } \mathbf{E}_P)} = \underbrace{(-1)^w}_{=\pm 1} \mathbf{E}_P = \pm \mathbf{E}_P. \end{aligned}$$

This proves Lemma 6.8.13. □

### 6.8.4. The basis theorem

Recall from Definition 6.1.1 **(e)** that $\mathrm{SBT}(n)$ is the set of all standard $n$-bitableaux, i.e., of all triples $(\lambda, U, V)$, where $\lambda$ is a partition of $n$ and where $U$ and $V$ are two standard $n$-tableaux of shape $\lambda$. We now state one of the main results in this section (the proof will be given later):

**Theorem 6.8.14.** Both families

$$\left(\nabla_{V,U}^{\mathrm{Row}}\right)_{(\lambda,U,V)\in \mathrm{SBT}(n)} \qquad \text{and} \qquad \left(\nabla_{V,U}^{-\mathrm{Col}}\right)_{(\lambda,U,V)\in \mathrm{SBT}(n)}$$

are bases of $\mathcal{A}$.

**Definition 6.8.15.** These two bases are called the *Murphy bases* (or the *Murphy cellular bases*), specifically the *row Murphy basis* and the *column Murphy basis*.

For $n = 3$, the row Murphy basis is the basis $\overrightarrow{\mathrm{rmu}}_3$ from Subsection 6.8.1, since its basis vectors are

$$\begin{aligned}
\nabla^{\mathrm{Row}}_{123,\ 123} &= \nabla; \\
\nabla^{\mathrm{Row}}_{12\backslash\backslash 3,\ 12\backslash\backslash 3} &= \nabla_{3,3}; \\
\nabla^{\mathrm{Row}}_{12\backslash\backslash 3,\ 13\backslash\backslash 2} &= \nabla_{3,2}; \\
\nabla^{\mathrm{Row}}_{13\backslash\backslash 2,\ 12\backslash\backslash 3} &= \nabla_{2,3}; \\
\nabla^{\mathrm{Row}}_{13\backslash\backslash 2,\ 13\backslash\backslash 2} &= \nabla_{2,2}; \\
\nabla^{\mathrm{Row}}_{1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3} &= \mathrm{id} = 1.
\end{aligned}$$

One of our main goals in this section is to prove Theorem 6.8.14, and then to study the Murphy bases in more detail, in particular their nice block-triangularity properties.

The Murphy bases were discovered by Canfield and Williamson in 1989 [CanWil89, Theorems 3.12, 3.14 and 6.13] (where the $\nabla^{\mathrm{Row}}_{V,U}$ are called *matched sums*, and the $\nabla^{-\mathrm{Col}}_{V,U}$ are called *signed sums*, although both are indexed by pairs of ordered set partitions rather than of tableaux) and – in the more general context of a Hecke algebra – by Murphy in 1992 [Murphy92, §3], [Murphy95, Theorem 4.17] (whose $x_{st}$ generalize our $\nabla^{\mathrm{Row}}_{V,U}$, although their definition is somewhat subtler).

One more little remark: The unity 1 of $\mathcal{A}$ is always an element of each of the two Murphy bases. Namely:

**Exercise 6.8.1.** $\boxed{1}$ **(a)** Prove that $\nabla^{\mathrm{Row}}_{P,P} = 1$, where $P$ is any $n$-tableau of shape $(1, 1, \ldots, 1)$.

**(b)** Prove that $\nabla^{-\mathrm{Col}}_{Q,Q} = 1$, where $Q$ is any $n$-tableau of shape $(n)$.

### 6.8.5. The dot product on $\mathcal{A}$

Recall the notion of a bilinear form (Definition 5.19.9).

**Definition 6.8.16.** Recall that $(w)_{w \in S_n}$ is the standard basis of $\mathbf{k}[S_n] = \mathcal{A}$. Thus, by Proposition 5.19.10, we can define a bilinear form $f : \mathcal{A} \times \mathcal{A} \to \mathbf{k}$ by setting

$$f(u, v) = \delta_{u,v} \qquad \text{for all } u \in S_n \text{ and } v \in S_n$$

(where $\delta_{u,v}$ denotes the Kronecker delta). We shall denote this form $f$ by $\langle \cdot, \cdot \rangle$, which means that we shall denote each value $f(\mathbf{a}, \mathbf{b})$ of $f$ as $\langle \mathbf{a}, \mathbf{b} \rangle$. Thus, for all $u \in S_n$ and $v \in S_n$, we have

$$\langle u, v \rangle = f(u, v) = \delta_{u,v} \tag{508}$$

(by the definition of $f$).

This form $\langle \cdot, \cdot \rangle$ will be called the *dot product* on $\mathcal{A}$ (even though our notation for it is a pair of angular brackets rather than a dot).

As the name "dot product" suggests, this form $\langle \cdot, \cdot \rangle$ is the obvious analogue of the dot product on $\mathbf{k}^n$ on the free $\mathbf{k}$-module $\mathbf{k}[S_n]$. Here is another way to say this:

**Exercise 6.8.2.** $\boxed{1}$ Let $(a_w)_{w \in S_n} \in \mathbf{k}^{S_n}$ and $(b_w)_{w \in S_n} \in \mathbf{k}^{S_n}$ be two families of scalars. Prove that

$$\left\langle \sum_{w \in S_n} a_w w, \sum_{w \in S_n} b_w w \right\rangle = \sum_{w \in S_n} a_w b_w.$$

We shall give several more explicit formulas for $\langle \mathbf{a}, \mathbf{b} \rangle$ in a moment. First, we recall from Definition 5.11.6 that if $\mathbf{a}$ is any element of $\mathcal{A} = \mathbf{k}[S_n]$ and if $w \in S_n$ is any permutation, then $[w]\,\mathbf{a}$ denotes the coefficient of $w$ in $\mathbf{a}$ (when $\mathbf{a}$ is expanded as a linear combination of the standard basis vectors of $\mathbf{k}[S_n]$). We now use this notation to rewrite the bilinear form $\langle \cdot, \cdot \rangle$ as follows:

**Proposition 6.8.17.** Let id mean the identity permutation $\mathrm{id}_{[n]} \in S_n$. Let $\mathbf{a}, \mathbf{b} \in \mathcal{A}$. Then,

$$\begin{aligned}
\langle \mathbf{a}, \mathbf{b} \rangle &= [\mathrm{id}]\,(S(\mathbf{a})\,\mathbf{b}) && (509)\\
&= [\mathrm{id}]\,(\mathbf{b} S(\mathbf{a})) && (510)\\
&= [\mathrm{id}]\,(S(\mathbf{b})\,\mathbf{a}) && (511)\\
&= [\mathrm{id}]\,(\mathbf{a} S(\mathbf{b})) && (512)\\
&= \langle \mathbf{b}, \mathbf{a} \rangle\,. && (513)
\end{aligned}$$

*Proof.* We must prove that the scalars $\langle \mathbf{a}, \mathbf{b} \rangle$, $[\mathrm{id}]\,(S(\mathbf{a})\,\mathbf{b})$, $[\mathrm{id}]\,(\mathbf{b} S(\mathbf{a}))$, $[\mathrm{id}]\,(S(\mathbf{b})\,\mathbf{a})$, $[\mathrm{id}]\,(\mathbf{a} S(\mathbf{b}))$ and $\langle \mathbf{b}, \mathbf{a} \rangle$ are equal. All these scalars depend $\mathbf{k}$-linearly on each of $\mathbf{a}$ and $\mathbf{b}$ (since $\langle \cdot, \cdot \rangle$ is a bilinear form and since $S$ is a $\mathbf{k}$-linear map). Hence, by linearity, we can WLOG assume that both $\mathbf{a}$ and $\mathbf{b}$ belong to the standard basis $(w)_{w \in S_n}$ of $\mathcal{A}$. Assume this. Thus, $\mathbf{a} = u$ and $\mathbf{b} = v$ for some $u, v \in S_n$. Consider these $u, v$.

From $\mathbf{a} = u$ and $\mathbf{b} = v$, we obtain

$$\begin{aligned}
\langle \mathbf{a}, \mathbf{b} \rangle &= \langle u, v \rangle = \delta_{u,v} \qquad \text{(by (508))}\\
&= \begin{cases} 1, & \text{if } u = v; \\ 0, & \text{if } u \neq v \end{cases} && (514)
\end{aligned}$$

and $S(\mathbf{a}) = S(u) = u^{-1}$ (by the definition of $S$) and $S(\mathbf{b}) = S(v) = v^{-1}$ (again by the definition of $S$).

Next, we recall that each permutation $w \in S_n$ satisfies

$$[\mathrm{id}]\, w = \begin{cases} 1, & \text{if } w = \mathrm{id}; \\ 0, & \text{if } w \neq \mathrm{id} \end{cases} \tag{515}$$

(by the very definition of $[\mathrm{id}]\,\mathbf{a}$ as the coefficient of id in $\mathbf{a}$). However, from $S(\mathbf{a}) = u^{-1}$ and $\mathbf{b} = v$, we obtain $S(\mathbf{a})\,\mathbf{b} = u^{-1}v$ and thus

$$\begin{aligned} [\mathrm{id}]\,(S(\mathbf{a})\,\mathbf{b}) &= [\mathrm{id}]\left(u^{-1}v\right) \\ &= \begin{cases} 1, & \text{if } u^{-1}v = \mathrm{id}; \\ 0, & \text{if } u^{-1}v \neq \mathrm{id} \end{cases} \qquad \left(\text{by (515), applied to } w = u^{-1}v\right) \\ &= \begin{cases} 1, & \text{if } u = v; \\ 0, & \text{if } u \neq v \end{cases} \end{aligned}$$

(since the equation $u^{-1}v = \mathrm{id}$ in $S_n$ is equivalent to $u = v$). Comparing this with (514), we obtain $\langle \mathbf{a}, \mathbf{b} \rangle = [\mathrm{id}]\,(S(\mathbf{a})\,\mathbf{b})$. Hence, (509) is proved.

A similar argument can be used to prove (510) (here, we have to argue that $\mathbf{b}S(\mathbf{a}) = vu^{-1}$, and that the equation $vu^{-1} = \mathrm{id}$ in $S_n$ is equivalent to $u = v$). Likewise, we can prove (511) (here, we must argue that $S(\mathbf{b})\,\mathbf{a} = v^{-1}u$, and that the equation $v^{-1}u = \mathrm{id}$ in $S_n$ is equivalent to $u = v$). Similarly, we can prove (512) (here, we must argue that $\mathbf{a}S(\mathbf{b}) = uv^{-1}$, and that the equation $uv^{-1} = \mathrm{id}$ in $S_n$ is equivalent to $u = v$).

It remains to prove (513). For this purpose, let us apply (509) to $\mathbf{b}$ and $\mathbf{a}$ instead of $\mathbf{a}$ and $\mathbf{b}$ (we can do this, since we have already proved (509)). Thus we obtain $\langle \mathbf{b}, \mathbf{a} \rangle = [\mathrm{id}]\,(S(\mathbf{b})\,\mathbf{a})$. Comparing this with (511), we obtain $\langle \mathbf{a}, \mathbf{b} \rangle = \langle \mathbf{b}, \mathbf{a} \rangle$. Thus, (513) is proved, and so the proof of Proposition 6.8.17 is complete. □

Note that the dot product on $\mathcal{A}$ is precisely the $G$-invariant form $f$ constructed in the proof of Theorem 5.19.30, where we take $G$ to be the symmetric group $S_n$ and we take the $G$-set $X$ to be $G = S_n$ itself (with the left regular $G$-action).

In the standard language of bilinear algebra, the equality (513) says that the dot product $\langle \cdot, \cdot \rangle$ is symmetric, whereas the equality (508) says that the standard basis $(w)_{w \in S_n}$ of $\mathcal{A}$ is orthonormal with respect to the dot product. Here are some further properties of the dot product:

**Definition 6.8.18.** Recall the concepts of left-curried and right-curried forms (Definition 5.19.7). A $\mathbf{k}$-bilinear form $f : U \times V \to \mathbf{k}$ on two $\mathbf{k}$-modules $U$ and $V$ is said to be *perfect* if its left-curried form $f_L : U \to V^*$ and its right-curried form $f_R : V \to U^*$ are $\mathbf{k}$-module isomorphisms. The dot product

(i.e., the bilinear form $\langle \cdot, \cdot \rangle$ from Definition 6.8.16) is perfect; this is easy to see from Exercise 6.8.2.

**Exercise 6.8.3.** $\boxed{1}$ **(a)** Prove that the bilinear form $\langle \cdot, \cdot \rangle$ is $S_n$-invariant, meaning that all $w \in S_n$ and $\mathbf{a}, \mathbf{b} \in \mathcal{A}$ satisfy

$$\langle \mathbf{a}, \mathbf{b} \rangle = \langle w\mathbf{a}, w\mathbf{b} \rangle .$$

**(b)** Prove that the bilinear form $\langle \cdot, \cdot \rangle$ is invariant under $S$ and $T_{\mathrm{sign}}$, meaning that all $\mathbf{a}, \mathbf{b} \in \mathcal{A}$ satisfy

$$\langle \mathbf{a}, \mathbf{b} \rangle = \langle S(\mathbf{a}), S(\mathbf{b}) \rangle = \left\langle T_{\mathrm{sign}}(\mathbf{a}), T_{\mathrm{sign}}(\mathbf{b}) \right\rangle .$$

### 6.8.6. Some lemmas from linear algebra

Next we shall state a few facts from linear algebra that will be used later on. We recall the notion of a $P \times Q$-matrix (as defined in Section 5.6).

**Lemma 6.8.19.** Let $I$ be a finite totally ordered set. Let $A = \left(a_{i,j}\right)_{i,j \in I} \in \mathbf{k}^{I \times I}$ be an $I \times I$-matrix. Assume that

- the scalar $a_{i,i} \in \mathbf{k}$ is invertible for each $i \in I$;
- we have $a_{i,j} = 0$ for all $i, j \in I$ that satisfy $i < j$.

[171] Then, the matrix $A$ is invertible.

*Proof.* For each $m \in \mathbb{N}$, the set $[m] = \{1, 2, \ldots, m\}$ is totally ordered (with the usual order: $1 < 2 < \cdots < m$).

Any finite totally ordered set is isomorphic to the totally ordered set $[m]$ for some $m \in \mathbb{N}$. In particular, this must hold for our finite totally ordered set $I$. Thus, we can WLOG assume that this totally ordered set $I$ **is** the totally ordered set $[m]$ for some $m \in \mathbb{N}$ [172]. Assume this.

[171] These two assumptions can be reworded as "the matrix $A$ is upper-triangular (if its rows and columns are ordered according to the total order on $I$) and its diagonal entries are invertible".

[172] In more detail: Let us label the elements of $I$ as $t_1, t_2, \ldots, t_m$ in increasing order. (This can be done, since $I$ is a finite totally ordered set.) Then, the map $[m] \to I$, $i \mapsto t_i$ is an isomorphism of totally ordered sets. Hence, the $I \times I$-matrix $A = \left(a_{i,j}\right)_{i,j \in I}$ can be transformed into the $m \times m$-matrix $A' := \left(a_{t_i, t_j}\right)_{i,j \in [m]}$ by merely changing the indexing of its rows and columns (and this change does not affect their ordering). Replacing the matrix $A$ by the matrix $A'$ (and the totally ordered set $I$ by the totally ordered set $[m]$) does not change the claim of Lemma 6.8.19 (since the matrix $A$ is invertible if and only if its reindexed version $A'$ is

Now, recall our assumption that we have $a_{i,j} = 0$ for all $i, j \in I$ that satisfy $i < j$. In view of $I = [m]$, this is just saying that $a_{i,j} = 0$ for all $i, j \in [m]$ that satisfy $i < j$. In other words, the matrix $A$ is lower-triangular. Hence, its determinant $\det A$ is the product of its diagonal entries (since the determinant of a lower-triangular matrix is the product of its diagonal entries). In other words,

$$\det A = a_{1,1}a_{2,2}\cdots a_{m,m} = \prod_{i\in[m]} a_{i,i} = \prod_{i\in I} a_{i,i} \qquad (\text{since } [m] = I).$$

But we assumed that the scalar $a_{i,i} \in \mathbf{k}$ is invertible for each $i \in I$. Hence, the product $\prod_{i\in I} a_{i,i}$ of these scalars is also invertible (since any product of invertible scalars is invertible). In other words, $\det A$ is invertible (since $\det A = \prod_{i\in I} a_{i,i}$). Thus, the matrix $A$ is itself invertible (because a square matrix whose determinant is invertible must itself be invertible[173]). This proves Lemma 6.8.19. □

**Lemma 6.8.20** (Gram trick)**.** Let $n \in \mathbb{N}$. Let $V$ be a free $\mathbf{k}$-module of rank $n$, and let $(u_1, u_2, \ldots, u_n)$ and $(v_1, v_2, \ldots, v_n)$ be two $n$-tuples of elements of $V$. Let $f : V \times V \to \mathbf{k}$ be a bilinear form. Assume that the $n \times n$-matrix

$$\left(f\left(u_i, v_j\right)\right)_{i,j\in[n]} \in \mathbf{k}^{n\times n}$$

is invertible. (This matrix is called the *Gram matrix* of $f$ with respect to the $n$-tuples $(u_1, u_2, \ldots, u_n)$ and $(v_1, v_2, \ldots, v_n)$.)

Then, both $n$-tuples $(u_1, u_2, \ldots, u_n)$ and $(v_1, v_2, \ldots, v_n)$ are bases of $V$.

*Proof.* We are given that the $\mathbf{k}$-module $V$ is free of rank $n$. In other words, $V$ has a basis $\overrightarrow{b} = (b_1, b_2, \ldots, b_n)$ consisting of $n$ vectors. Consider this basis.

For each $i \in [n]$, we let $x_{i,1}, x_{i,2}, \ldots, x_{i,n}$ be the coordinates of the vector $u_i \in V$ with respect to the basis $\overrightarrow{b}$; that is, we write $u_i$ as

$$u_i = \sum_{k\in[n]} x_{i,k} b_k, \qquad \text{where } x_{i,k} \in \mathbf{k}. \tag{516}$$

For each $j \in [n]$, we let $y_{j,1}, y_{j,2}, \ldots, y_{j,n}$ be the coordinates of the vector $v_j \in V$ with respect to the basis $\overrightarrow{b}$; that is, we write $v_j$ as

$$v_j = \sum_{\ell\in[n]} y_{j,\ell} b_\ell, \qquad \text{where } y_{j,\ell} \in \mathbf{k}. \tag{517}$$

invertible), nor does it change the assumptions (for example, the assumption "$a_{i,j} = 0$ for all $i, j \in I$ that satisfy $i < j$" is equivalent to "$a_{t_i,t_j} = 0$ for all $i, j \in [m]$ that satisfy $i < j$", because the map $[m] \to I$, $i \mapsto t_i$ is an isomorphism of totally ordered sets). Hence, we can WLOG assume that $I = [m]$ (otherwise, we just replace $I$ and $A$ by $[m]$ and $A'$).

[173] This is a classical fact (see, e.g., [Grinbe15, Theorem 6.110 **(a)**]).

We let $A_{i,j}$ denote the $(i,j)$-th entry of a matrix $A$.

Let $X$ be the $n \times n$-matrix $\left(x_{i,k}\right)_{i,k\in[n]} \in \mathbf{k}^{n\times n}$, and let $Y$ be the $n \times n$-matrix $\left(y_{j,\ell}\right)_{j,\ell\in[n]} \in \mathbf{k}^{n\times n}$. Thus, we have

$$X_{i,k} = x_{i,k} \qquad \text{for each } i,k \in [n] \tag{518}$$

and

$$Y_{j,\ell} = y_{j,\ell} \qquad \text{for each } j,\ell \in [n]\,. \tag{519}$$

Let $G$ be the matrix $\left(f\left(b_k, b_\ell\right)\right)_{k,\ell\in[n]} \in \mathbf{k}^{n\times n}$. Thus,

$$G_{k,\ell} = f\left(b_k, b_\ell\right) \qquad \text{for each } k,\ell \in [n]\,. \tag{520}$$

Now, we claim that

$$\left(f\left(u_i, v_j\right)\right)_{i,j\in[n]} = XGY^T. \tag{521}$$

*Proof of (521).* If $A$ and $B$ are two $n \times n$-matrices, then the entries of their product $AB$ are given by the formula

$$(AB)_{i,j} = \sum_{k\in[n]} A_{i,k}B_{k,j} \qquad \text{for all } i,j \in [n]$$

(by the definition of the product of two matrices). Applying this formula twice, we obtain the following variant of it: If $A$, $B$ and $C$ are three $n \times n$-matrices, then the entries of their product $ABC$ are given by the formula

$$(ABC)_{i,j} = \sum_{k\in[n]} \sum_{\ell\in[n]} A_{i,k}B_{k,\ell}C_{\ell,j} \qquad \text{for all } i,j \in [n]\,.$$

Applying this formula to $A = X$ and $B = G$ and $C = Y^T$, we obtain

$$\begin{aligned}
\left(XGY^T\right)_{i,j} &= \sum_{k\in[n]} \sum_{\ell\in[n]} X_{i,k}G_{k,\ell} \underbrace{\left(Y^T\right)_{\ell,j}}_{\substack{=Y_{j,\ell} \\ \text{(by the definition of} \\ \text{the transpose of a matrix)}}} \\
&= \sum_{k\in[n]} \sum_{\ell\in[n]} \underbrace{X_{i,k}}_{\substack{=x_{i,k} \\ \text{(by (518))}}} \underbrace{G_{k,\ell}}_{\substack{=f(b_k,b_\ell) \\ \text{(by (520))}}} \underbrace{Y_{j,\ell}}_{\substack{=y_{j,\ell} \\ \text{(by (519))}}} \\
&= \sum_{k\in[n]} \sum_{\ell\in[n]} x_{i,k} f\left(b_k, b_\ell\right) y_{j,\ell} \qquad \text{for all } i,j \in [n]\,.
\end{aligned}$$

On the other hand, for all $i, j \in [n]$, we have

$$\begin{aligned} f\left(u_i, v_j\right) &= f\left(\sum_{k\in[n]} x_{i,k} b_k, \sum_{\ell\in[n]} y_{j,\ell} b_\ell\right) && \text{(by (516) and (517))} \\ &= \sum_{k\in[n]} \sum_{\ell\in[n]} x_{i,k} \underbrace{y_{j,\ell} f\left(b_k, b_\ell\right)}_{=f(b_k,b_\ell) y_{j,\ell}} && \text{(since } f \text{ is } \mathbf{k}\text{-bilinear)} \\ &= \sum_{k\in[n]} \sum_{\ell\in[n]} x_{i,k} f\left(b_k, b_\ell\right) y_{j,\ell}. \end{aligned}$$

Comparing these two equalities, we obtain

$$\left(XGY^T\right)_{i,j} = f\left(u_i, v_j\right) \qquad \text{for all } i, j \in [n].$$

Hence, the $n \times n$-matrix $XGY^T$ equals $\left(f\left(u_i, v_j\right)\right)_{i,j\in[n]}$. This proves (521). □

But we assumed that the $n \times n$-matrix $\left(f\left(u_i, v_j\right)\right)_{i,j\in[n]}$ is invertible. Because of (521), we can rewrite this as follows: The $n \times n$-matrix $XGY^T$ is invertible. Hence, its determinant $\det\left(XGY^T\right)$ is invertible in $\mathbf{k}$ (because a square matrix is invertible if and only if its determinant is invertible[174]).

But it is well-known that the determinant function is multiplicative (i.e., the determinant of a product of $n \times n$-matrices is the product of their determinants). Since $X$, $G$ and $Y^T$ are square matrices, we thus obtain

$$\det\left(XGY^T\right) = \det X \cdot \det G \cdot \underbrace{\det\left(Y^T\right)}_{\substack{=\det Y \\ (\text{since } \det\left(A^T\right)=\det A \\ \text{for any square matrix } A)}} = \det X \cdot \det G \cdot \det Y.$$

Thus, the product $\det X \cdot \det G \cdot \det Y$ is invertible in $\mathbf{k}$ (since $\det\left(XGY^T\right)$ is invertible in $\mathbf{k}$). Therefore, its three factors $\det X$ and $\det G$ and $\det Y$ are all invertible in $\mathbf{k}$ (since a product of elements of a commutative ring is invertible if and only if all its factors are invertible).

In particular, $\det X$ is invertible. Hence, the matrix $X$ is invertible (since a square matrix is invertible if and only if its determinant is invertible). Thus, its transpose $X^T$ is invertible as well (since the transpose of an invertible matrix is invertible). Therefore, the $n$-tuple $(u_1, u_2, \ldots, u_n)$ is a basis of $V$ [175]. A similar argument (using the invertibility of $\det Y$) shows that $(v_1, v_2, \ldots, v_n)$ is a basis of $V$ as well. Thus, Lemma 6.8.20 is proved. □

[174] This is a classical fact (see, e.g., [Grinbe15, Theorem 6.110 **(a)**]).

[175] *Proof.* Let $\alpha : V \to V$ be the $\mathbf{k}$-linear map that sends the basis vectors $b_1, b_2, \ldots, b_n$ of the basis $\overrightarrow{b}$ to the vectors $u_1, u_2, \ldots, u_n$, respectively (this map is well-defined, since a $\mathbf{k}$-linear map from a free $\mathbf{k}$-module $M$ can be defined by specifying its values on a given basis of $M$).

**Exercise 6.8.4.** [2] Under the assumptions of Lemma 6.8.20, prove that the bilinear form $f$ is perfect (i.e., its left-curried form $f_L : V \to V^*$ and its right-curried form $f_R : V \to V^*$ are $\mathbf{k}$-module isomorphisms).

**Exercise 6.8.5.** [2] Consider the setting of Lemma 6.8.20 again, but drop the assumption that $V$ be free of rank $n$. Prove that both $n$-tuples $(u_1, u_2, \ldots, u_n)$ and $(v_1, v_2, \ldots, v_n)$ are $\mathbf{k}$-linearly independent (although no longer necessarily bases).

For convenience, let us state a version of Lemma 6.8.20 in which the two tuples $(u_1, u_2, \ldots, u_n)$ and $(v_1, v_2, \ldots, v_n)$ have been reindexed as families $(u_i)_{i \in I}$ and $(v_i)_{i \in I}$:

**Lemma 6.8.21** (Gram trick, redux)**.** Let $n \in \mathbb{N}$. Let $V$ be a free $\mathbf{k}$-module of rank $n$, and let $(u_i)_{i \in I}$ and $(v_i)_{i \in I}$ be two families of elements of $V$, where $I$ is a finite set of size $|I| = n$. Let $f : V \times V \to \mathbf{k}$ be a bilinear form. Assume that the $I \times I$-matrix

$$\left(f\left(u_i, v_j\right)\right)_{i,j \in I} \in \mathbf{k}^{I \times I}$$

is invertible. (This matrix is called the *Gram matrix* of $f$ with respect to the families $(u_i)_{i \in I}$ and $(v_i)_{i \in I}$.)

Then, both families $(u_i)_{i \in I}$ and $(v_i)_{i \in I}$ are bases of $V$.

*Proof.* Rename the $n$ elements of $I$ as $1, 2, \ldots, n$; then, this is just Lemma 6.8.20. □

### 6.8.7. Proof of the basis property

We shall now show two lemmas about the bilinear form $\langle \cdot, \cdot \rangle$ on $\mathcal{A}$ and its interaction with the Murphy bases. These lemmas will be crucial for proving

---

Then, for each $i \in [n]$, we have

$$\begin{aligned} \alpha\left(b_i\right) &= u_i \qquad \text{(since the map } \alpha \text{ sends } b_i \text{ to } u_i \text{, by its definition)} \\ &= \sum_{k \in [n]} x_{i,k} b_k \qquad \text{(by (516))}. \end{aligned}$$

Thus, the matrix representing the $\mathbf{k}$-linear map $\alpha : V \to V$ with respect to the basis $\overrightarrow{b} = (b_1, b_2, \ldots, b_n)$ is the matrix $(x_{i,k})_{k,i \in [n]}$ (by Definition 6.2.2). But this matrix $(x_{i,k})_{k,i \in [n]}$ is just $X^T$ (since $X = (x_{i,k})_{i,k \in [n]}$ and thus $X^T = (x_{i,k})_{k,i \in [n]}$), and thus is invertible (since we have shown that $X^T$ is invertible). Hence, we have shown that the matrix representing the $\mathbf{k}$-linear map $\alpha : V \to V$ with respect to the basis $\overrightarrow{b}$ is invertible. Therefore, the $\mathbf{k}$-linear map $\alpha : V \to V$ is invertible as well (since a $\mathbf{k}$-linear map that is represented by an invertible matrix is itself invertible). In other words, $\alpha : V \to V$ is a $\mathbf{k}$-module isomorphism.

This isomorphism $\alpha : V \to V$ sends the $n$ vectors $b_1, b_2, \ldots, b_n$ to the $n$ vectors $u_1, u_2, \ldots, u_n$ (by its definition). But a $\mathbf{k}$-module isomorphism always sends a basis of its domain to a basis of its target. Since the isomorphism $\alpha$ sends $b_1, b_2, \ldots, b_n$ to $u_1, u_2, \ldots, u_n$, we thus conclude that $(u_1, u_2, \ldots, u_n)$ is a basis of $V$ (since $(b_1, b_2, \ldots, b_n)$ is a basis of $V$).

Theorem 6.8.14:

**Lemma 6.8.22** (Murphy diagonal lemma)**.** Let $(\lambda, A, B) \in \operatorname{SBT}(n)$. Then,

$$\left\langle \nabla_{B,A}^{-\operatorname{Col}},\ \nabla_{B,A}^{\operatorname{Row}} \right\rangle = \pm 1.$$

*Proof of Lemma 6.8.22.* Let $\mathbf{E}_P$ denote the Young symmetrizer of an $n$-tableau $P$ (as defined in Definition 5.11.1 **(b)**).

We have

$$\underbrace{S\left(\nabla_{B,A}^{-\operatorname{Col}}\right)}_{\substack{=\nabla_{A,B}^{-\operatorname{Col}} \\ \text{(by (503))}}} \nabla_{B,A}^{\operatorname{Row}} = \nabla_{A,B}^{-\operatorname{Col}} \nabla_{B,A}^{\operatorname{Row}} = \pm \mathbf{E}_A$$

(by Lemma 6.8.13, applied to $P = A$ and $Q = B$). Now, from (509), we obtain

$$\begin{aligned}\left\langle \nabla_{B,A}^{-\operatorname{Col}},\ \nabla_{B,A}^{\operatorname{Row}} \right\rangle &= \left[\underbrace{\mathrm{id}}_{=1}\right]\left(\underbrace{S\left(\nabla_{B,A}^{-\operatorname{Col}}\right)\nabla_{B,A}^{\operatorname{Row}}}_{=\pm\mathbf{E}_A}\right) = [1]\left(\pm\mathbf{E}_A\right) \\ &= \pm \underbrace{[1]\left(\mathbf{E}_A\right)}_{\substack{=1 \\ \text{(by Lemma 5.11.13)}}} = \pm 1.\end{aligned}$$

This proves Lemma 6.8.22. □

**Lemma 6.8.23** (Murphy orthogonality lemma)**.** Let $(\lambda, A, B) \in \operatorname{SBT}(n)$ and $(\mu, C, D) \in \operatorname{SBT}(n)$. Assume that $A < C$ or $B < D$, where the symbol "$<$" refers to the relation $<$ from Definition 6.7.5. Then,

$$\left\langle \nabla_{B,A}^{-\operatorname{Col}},\ \nabla_{D,C}^{\operatorname{Row}} \right\rangle = 0.$$

*Proof of Lemma 6.8.23.* Both $A$ and $B$ are standard $n$-tableaux of shape $Y(\lambda)$ (since $(\lambda, A, B) \in \operatorname{SBT}(n)$). Hence, Proposition 5.3.13 **(b)** (applied to $Y(\lambda)$ and $B$ instead of $D$ and $T$) shows that any $n$-tableau of shape $Y(\lambda)$ can be written as $w \rightharpoonup B$ for some $w \in S_n$. Hence, in particular, the $n$-tableau $A$ can be written in this way. In other words, $A = w \rightharpoonup B$ for some $w \in S_n$. Let us denote this $w$ by $u$. Thus, $A = u \rightharpoonup B = uB$. Applying (499) to $P = A$ and $Q = B$, we thus obtain

$$\nabla_{A,B}^{-\operatorname{Col}} = (-1)^u \nabla_{\operatorname{Col} A}^{-} u = (-1)^u u \nabla_{\operatorname{Col} B}^{-}.$$

Both $C$ and $D$ are standard $n$-tableaux of shape $Y(\mu)$ (since $(\mu, C, D) \in \operatorname{SBT}(n)$). Hence, Proposition 5.3.13 **(b)** (applied to $Y(\mu)$ and $C$ instead of $D$ and $T$) shows that any $n$-tableau of shape $Y(\mu)$ can be written as $w \rightharpoonup C$ for

some $w \in S_n$. Hence, in particular, the $n$-tableau $D$ can be written in this way. In other words, $D = w \rightharpoonup C$ for some $w \in S_n$. Let us denote this $w$ by $v$. Thus, $D = v \rightharpoonup C = vC$. Applying (498) to $D$, $C$ and $v$ instead of $P$, $Q$ and $u$, we thus obtain

$$\nabla^{\mathrm{Row}}_{D,C} = \nabla_{\mathrm{Row}\, D} v = v \nabla_{\mathrm{Row}\, C}.$$

Furthermore, $S\left(\nabla^{-\mathrm{Col}}_{B,A}\right) = \nabla^{-\mathrm{Col}}_{A,B}$ (by (503)).

We have assumed that $A < C$ or $B < D$. Thus, we are in one of the following two cases (which may overlap):

*Case 1:* We have $A < C$.

*Case 2:* We have $B < D$.

Let us consider Case 1 first. In this case, we have $A < C$. Hence, Lemma 6.7.9 (applied to $P = A$ and $Q = C$) yields

$$\nabla^{-}_{\mathrm{Col}\, A} \nabla_{\mathrm{Row}\, C} = 0 \qquad \text{and} \qquad \nabla_{\mathrm{Row}\, C} \nabla^{-}_{\mathrm{Col}\, A} = 0.$$

Now,

$$\underbrace{\nabla^{\mathrm{Row}}_{D,C}}_{= v \nabla_{\mathrm{Row}\, C}} \underbrace{S\left(\nabla^{-\mathrm{Col}}_{B,A}\right)}_{\substack{= \nabla^{-\mathrm{Col}}_{A,B} \\ = (-1)^u \nabla^{-}_{\mathrm{Col}\, A} u}} = (-1)^u \cdot v \underbrace{\nabla_{\mathrm{Row}\, C} \nabla^{-}_{\mathrm{Col}\, A}}_{=0} u = 0.$$

Hence, (510) yields $\left\langle \nabla^{-\mathrm{Col}}_{B,A},\ \nabla^{\mathrm{Row}}_{D,C} \right\rangle = [\mathrm{id}] \left( \underbrace{\nabla^{\mathrm{Row}}_{D,C} S\left(\nabla^{-\mathrm{Col}}_{B,A}\right)}_{=0} \right) = [\mathrm{id}]\, 0 = 0.$

Thus, Lemma 6.8.23 is proved in Case 1.

Let us consider Case 2 next. In this case, we have $B < D$. Hence, Lemma 6.7.9 (applied to $P = B$ and $Q = D$) yields

$$\nabla^{-}_{\mathrm{Col}\, B} \nabla_{\mathrm{Row}\, D} = 0 \qquad \text{and} \qquad \nabla_{\mathrm{Row}\, D} \nabla^{-}_{\mathrm{Col}\, B} = 0.$$

Now,

$$\underbrace{S\left(\nabla^{-\mathrm{Col}}_{B,A}\right)}_{\substack{= \nabla^{-\mathrm{Col}}_{A,B} \\ = (-1)^u u \nabla^{-}_{\mathrm{Col}\, B}}} \underbrace{\nabla^{\mathrm{Row}}_{D,C}}_{= \nabla_{\mathrm{Row}\, D} v} = (-1)^u\, u \underbrace{\nabla^{-}_{\mathrm{Col}\, B} \nabla_{\mathrm{Row}\, D}}_{=0} v = 0.$$

Hence, (509) yields $\left\langle \nabla^{-\mathrm{Col}}_{B,A},\ \nabla^{\mathrm{Row}}_{D,C} \right\rangle = [\mathrm{id}] \left( \underbrace{S\left(\nabla^{-\mathrm{Col}}_{B,A}\right) \nabla^{\mathrm{Row}}_{D,C}}_{=0} \right) = [\mathrm{id}]\, 0 = 0.$

Thus, Lemma 6.8.23 is proved in Case 2.

So the proof of Lemma 6.8.23 is complete in both Cases 1 and 2. □

Now, we can prove Theorem 6.8.14:

*Proof of Theorem 6.8.14.* Recall the total order on $\mathrm{SBT}(n)$ defined in Definition 6.7.12. Thus, $\mathrm{SBT}(n)$ is a finite totally ordered set with size $|\mathrm{SBT}(n)| = n!$ (by Lemma 6.2.1).

Now, let $M$ be the $\mathrm{SBT}(n) \times \mathrm{SBT}(n)$-matrix

$$\left(\left\langle \nabla^{-\mathrm{Col}}_{B,A},\ \nabla^{\mathrm{Row}}_{D,C} \right\rangle\right)_{(\lambda,A,B),(\mu,C,D)\in\mathrm{SBT}(n)} \in \mathbf{k}^{\mathrm{SBT}(n)\times\mathrm{SBT}(n)}.$$

We claim that this matrix $M$ is "upper-triangular with all its diagonal entries being invertible". Here is what this means in concrete terms:

- The scalar $\left\langle \nabla^{-\mathrm{Col}}_{B,A},\ \nabla^{\mathrm{Row}}_{B,A} \right\rangle \in \mathbf{k}$ is invertible for each $(\lambda, A, B) \in \mathrm{SBT}(n)$ (because Lemma 6.8.22 shows that this scalar is $\pm 1$, but $\pm 1$ is always invertible).
- We have $\left\langle \nabla^{-\mathrm{Col}}_{B,A},\ \nabla^{\mathrm{Row}}_{D,C} \right\rangle = 0$ for all $(\lambda, A, B), (\mu, C, D) \in \mathrm{SBT}(n)$ that satisfy $(\lambda, A, B) < (\mu, C, D)$. (Indeed, if $(\lambda, A, B), (\mu, C, D) \in \mathrm{SBT}(n)$ satisfy $(\lambda, A, B) < (\mu, C, D)$, then we have $A < C$ or $B < D$ (by Lemma 6.7.14 **(b)**), and therefore Lemma 6.8.23 yields $\left\langle \nabla^{-\mathrm{Col}}_{B,A},\ \nabla^{\mathrm{Row}}_{D,C} \right\rangle = 0$.)

Thus, Lemma 6.8.19 (applied to the set $\mathrm{SBT}(n)$ and the matrix $M = \left(\left\langle \nabla^{-\mathrm{Col}}_{B,A},\ \nabla^{\mathrm{Row}}_{D,C} \right\rangle\right)_{(\lambda,A,B),(\mu,C,D)\in\mathrm{SBT}(n)}$ instead of the set $I$ and the matrix $A = \left(a_{i,j}\right)_{i,j\in I}$) shows that the matrix $M$ is invertible.

But the definition of $M$ shows that $M$ is actually the Gram matrix of the form $\langle \cdot, \cdot \rangle : \mathcal{A} \times \mathcal{A} \to \mathbf{k}$ with respect to the two families

$$\left(\nabla^{-\mathrm{Col}}_{V,U}\right)_{(\lambda,U,V)\in\mathrm{SBT}(n)} \qquad \text{and} \qquad \left(\nabla^{\mathrm{Row}}_{V,U}\right)_{(\lambda,U,V)\in\mathrm{SBT}(n)}$$

(see Lemma 6.8.21 for the definition of "Gram matrix"). Hence, Lemma 6.8.21 (applied to the number $n!$, the free $\mathbf{k}$-module $\mathcal{A}$, the two families $\left(\nabla^{-\mathrm{Col}}_{V,U}\right)_{(\lambda,U,V)\in\mathrm{SBT}(n)}$ and $\left(\nabla^{\mathrm{Row}}_{V,U}\right)_{(\lambda,U,V)\in\mathrm{SBT}(n)}$ and the bilinear form $\langle \cdot, \cdot \rangle$ instead of the number $n$, the free $\mathbf{k}$-module $V$, the two families $(u_i)_{i\in I}$ and $(v_i)_{i\in I}$ and the bilinear form $f$) shows that both families

$$\left(\nabla^{-\mathrm{Col}}_{V,U}\right)_{(\lambda,U,V)\in\mathrm{SBT}(n)} \qquad \text{and} \qquad \left(\nabla^{\mathrm{Row}}_{V,U}\right)_{(\lambda,U,V)\in\mathrm{SBT}(n)}$$

are bases of $\mathcal{A}$ (since $\mathcal{A}$ is a free $\mathbf{k}$-module of rank $n!$, and since the Gram matrix $\left(\left\langle \nabla^{-\mathrm{Col}}_{B,A},\ \nabla^{\mathrm{Row}}_{D,C} \right\rangle\right)_{(\lambda,A,B),(\mu,C,D)\in\mathrm{SBT}(n)} = M$ is invertible). Hence, Theorem 6.8.14 is proved. □

In other words, we now know that the Murphy bases are bases indeed.

**Remark 6.8.24.** Other proofs of Theorem 6.8.14 appear in [CanWil89, Theorems 3.12, 3.14 and 6.13] and [Murphy92, Lemma 3.5 and Lemma 3.7]. These proofs have the advantage of giving a more efficient algorithm to expand a standard basis vector $w$ in the Murphy bases. (Our proof gives a rather horrible algorithm, since it relies on the proof of Lemma 6.8.21 and thus involves inverting an $n! \times n!$-matrix.)

**Exercise 6.8.6.** [1] Reprove Corollary 5.21.11 using Theorem 6.8.14.

### 6.8.8. The Murphy spans

Now we come to the study of spans of subfamilies of the Murphy bases. We will show that some of these spans are left ideals (i.e., left $\mathcal{A}$-submodules) of $\mathcal{A}$, which explains the block-upper-triangularity of the matrices $L_{\overrightarrow{\mathrm{rmu}}}(\mathbf{a})$ in Subsection 6.8.1. (The block-upper-triangularity of the matrices $R_{\overrightarrow{\mathrm{rmu}}}(\mathbf{a})$ can be explained similarly, or derived from the former via the antipode.)

We recall that $\mathrm{SBT}(n)$ is the set of all standard $n$-bitableaux, i.e., of all triples $(\lambda, U, V)$, where $\lambda$ is a partition of $n$ and where $U$ and $V$ are two standard $n$-tableaux of shape $\lambda$. We shall now define a close relative of this set: the set $\mathrm{HSBT}(n)$ of all *half-standard $n$-bitableaux*:

**Definition 6.8.25.** A *half-standard $n$-bitableau* shall mean a triple $(\lambda, U, V)$, where $\lambda$ is a partition of $n$ and where $U$ and $V$ are two $n$-tableaux of shape $\lambda$ such that $U$ is standard. Thus, the difference between a half-standard $n$-bitableau $(\lambda, U, V)$ and a standard one is that in the former, $V$ is not required to be standard (but $U$ is), whereas both $U$ and $V$ are required to be standard in the latter.

We let $\mathrm{HSBT}(n)$ denote the set of all half-standard $n$-bitableaux.

For example, $\left((2,1), \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array}\right)$ is a half-standard 3-bitableau, thus belongs to $\mathrm{HSBT}(3)$.

Obviously, $\mathrm{SBT}(n) \subseteq \mathrm{HSBT}(n)$, since any half-standard $n$-bitableau is a standard $n$-bitableau as well.

**Definition 6.8.26.** Let $T \in \bigcup_{\kappa \vdash n} \mathrm{SYT}(\kappa)$ be a standard $n$-tableau. Then, we define the following eight $\mathbf{k}$-submodules of $\mathcal{A}$:

- the *standard row-Murphy spans*

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} := \mathrm{span}\left\{\nabla^{\mathrm{Row}}_{V,U} \mid (\lambda, U, V) \in \mathrm{SBT}(n) \text{ and } U \geq T\right\};$$

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},> T} := \mathrm{span}\left\{\nabla^{\mathrm{Row}}_{V,U} \mid (\lambda, U, V) \in \mathrm{SBT}(n) \text{ and } U > T\right\};$$

- the *standard column-Murphy spans*

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T} := \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{V,U} \;\mid\; (\lambda, U, V) \in \mathrm{SBT}(n) \text{ and } U \leq T \right\};$$

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},< T} := \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{V,U} \;\mid\; (\lambda, U, V) \in \mathrm{SBT}(n) \text{ and } U < T \right\};$$

- the *non-standard row-Murphy spans*

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T} := \operatorname{span}\left\{ \nabla^{\mathrm{Row}}_{V,U} \;\mid\; (\lambda, U, V) \in \mathrm{HSBT}(n) \text{ and } U \geq T \right\};$$

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},> T} := \operatorname{span}\left\{ \nabla^{\mathrm{Row}}_{V,U} \;\mid\; (\lambda, U, V) \in \mathrm{HSBT}(n) \text{ and } U > T \right\};$$

- the *non-standard column-Murphy spans*

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},\leq T} := \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{V,U} \;\mid\; (\lambda, U, V) \in \mathrm{HSBT}(n) \text{ and } U \leq T \right\};$$

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},< T} := \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{V,U} \;\mid\; (\lambda, U, V) \in \mathrm{HSBT}(n) \text{ and } U < T \right\},$$

where all the inequality signs ($\geq$, $>$, $\leq$ and $<$) are understood with respect to the total order on $\bigcup_{\kappa \vdash n} \mathrm{SYT}(\kappa)$ (see Definition 6.7.7).

(The reader can easily see the pattern behind these naming conventions: The "std" and "all" subscripts refer to the $n$-tableau $V$ being either standard or arbitrary (respectively), whereas the "$\geq T$", "$> T$", "$\leq T$" and "$< T$" subscripts refer to the inequalities that $U$ is required to satisfy. Other combinations of requirements are possible, but we will only use the above eight.)

**Example 6.8.27.** For $n = 3$ and $T = 13\backslash\backslash 2$, we have

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} = \operatorname{span}\left\{ \nabla^{\mathrm{Row}}_{13\backslash\backslash 2,\ 13\backslash\backslash 2},\ \nabla^{\mathrm{Row}}_{13\backslash\backslash 2,\ 12\backslash\backslash 3},\ \nabla^{\mathrm{Row}}_{12\backslash\backslash 3,\ 13\backslash\backslash 2},\ \nabla^{\mathrm{Row}}_{12\backslash\backslash 3,\ 12\backslash\backslash 3},\ \nabla^{\mathrm{Row}}_{123,\ 123} \right\};$$

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},> T} = \operatorname{span}\left\{ \nabla^{\mathrm{Row}}_{13\backslash\backslash 2,\ 12\backslash\backslash 3},\ \nabla^{\mathrm{Row}}_{12\backslash\backslash 3,\ 12\backslash\backslash 3},\ \nabla^{\mathrm{Row}}_{123,\ 123} \right\};$$

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T} = \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3},\ \nabla^{-\mathrm{Col}}_{13\backslash\backslash 2,\ 13\backslash\backslash 2},\ \nabla^{-\mathrm{Col}}_{12\backslash\backslash 3,\ 13\backslash\backslash 2} \right\};$$

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},< T} = \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3} \right\}.$$

The $\mathbf{k}$-modules $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$, $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},>T}$, $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},\leq T}$, $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},<T}$ are similar but have additional generators $\nabla^{\mathrm{Row}}_{V,U}$ or $\nabla^{-\mathrm{Col}}_{V,U}$ in which $V$ is a non-standard tableau; for instance, one of the additional generators of $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T}$ is $\nabla^{\mathrm{Row}}_{23\backslash\backslash 1,\ 12\backslash\backslash 3}$.

Thus we have defined eight submodules of $\mathcal{A}$. Soon we will see (Theorem 6.8.29) that there are in fact only four of them, since each standard span equals the corresponding non-standard one (e.g., we have $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} = \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$). This is why many of our results about these submodules will come in four variants.

To state the main result of this subsection, we need a classical notion from the linear algebra of bilinear forms:

**Definition 6.8.28.** For each subset $V$ of $\mathcal{A}$, we define a subset $V^{\perp}$ of $\mathcal{A}$ by

$$V^{\perp} := \{\mathbf{a} \in \mathcal{A} \mid \langle \mathbf{a}, \mathbf{v} \rangle = 0 \text{ for each } \mathbf{v} \in V\}. \tag{522}$$

This subset $V^{\perp}$ is easily seen to be a $\mathbf{k}$-submodule of $\mathcal{A}$ (since the dot product $\langle \cdot, \cdot \rangle$ is bilinear, and thus $\langle \mathbf{a}, \mathbf{v} \rangle$ depends $\mathbf{k}$-linearly on $\mathbf{a}$). We call $V^{\perp}$ the *orthogonal complement* of $V$.

Because of the equality (513), we can replace the condition "$\langle \mathbf{a}, \mathbf{v} \rangle = 0$" in (522) by the equivalent condition "$\langle \mathbf{v}, \mathbf{a} \rangle = 0$".

We can now state the main claim of this subsection:

**Theorem 6.8.29.** Let $T \in \bigcup_{\kappa \vdash n} \mathrm{SYT}(\kappa)$ be a standard $n$-tableau. Then:

**(a)** The set $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$ and satisfies

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} = \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T} = \left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right)^{\perp}.$$

**(b)** The set $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$ and satisfies

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T} = \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},>T} = \left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}\right)^{\perp}.$$

**(c)** The set $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$ and satisfies

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} = \mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},<T} = \left(\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}\right)^{\perp}.$$

**(d)** The set $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$ and satisfies

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T} = \mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},\leq T} = \left(\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T}\right)^{\perp}.$$

It will take us a while to prove this, not least because of the similar-but-not-quite-the-same nature of the four parts. First, however, we observe some near-obvious equalities:

**Proposition 6.8.30.** Let $T_1, T_2, \ldots, T_m$ be all the standard tableaux in the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, listed in increasing order (with respect to the total order defined in Definition 6.7.7). Then:

**(a)** We have

$$\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T_1} = \mathcal{A} \qquad \text{and} \qquad \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T_m} = \mathcal{A}.$$

**(b)** We have

$$\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},> T_m} = 0 \qquad \text{and} \qquad \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},< T_1} = 0.$$

**(c)** Let $i \in [m-1]$. Then,

$$\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T_{i+1}} = \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},> T_i} \qquad \text{and} \qquad \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T_i} = \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},< T_{i+1}} \qquad \text{and}$$

$$\mathcal{F}^{\operatorname{Row}}_{\operatorname{all},\geq T_{i+1}} = \mathcal{F}^{\operatorname{Row}}_{\operatorname{all},> T_i} \qquad \text{and} \qquad \mathcal{F}^{-\operatorname{Col}}_{\operatorname{all},\leq T_i} = \mathcal{F}^{-\operatorname{Col}}_{\operatorname{all},< T_{i+1}}.$$

*Proof.* We first observe that the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ is nonempty[176]. However, $T_1, T_2, \ldots, T_m$ were defined to be all the standard tableaux in this set, listed without repetition. Hence, $m = \left| \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa) \right| > 0$ (since the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ is nonempty). This shows that $T_1$ and $T_m$ are well-defined.

Recall that $T_1, T_2, \ldots, T_m$ are all the standard tableaux in the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, listed in increasing order. Hence, $T_1$ is the smallest element of the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, and $T_m$ is the largest element of this set.

**(a)** Each tableau $U \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ satisfies $U \geq T_1$ (since $T_1$ is the smallest element of the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$). Thus,

$$\text{each } (\lambda, U, V) \in \operatorname{SBT}(n) \text{ satisfies } U \geq T_1 \tag{523}$$

(because each $(\lambda, U, V) \in \operatorname{SBT}(n)$ satisfies $U \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$).

[176] Indeed, the $n$-tableau of shape $(n)$ with entries $1, 2, \ldots, n$ arranged from left to right in its only row is clearly an element of this set. Thus, this set has at least one element.

On the other hand, each tableau $U \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ satisfies $U \leq T_m$ (since $T_m$ is the largest element of the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$). Thus,

$$\text{each } (\lambda, U, V) \in \operatorname{SBT}(n) \text{ satisfies } U \leq T_m \tag{524}$$

(because each $(\lambda, U, V) \in \operatorname{SBT}(n)$ satisfies $U \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$).

Theorem 6.8.14 shows that both families

$$\left(\nabla^{\operatorname{Row}}_{V,U}\right)_{(\lambda,U,V)\in\operatorname{SBT}(n)} \qquad \text{and} \qquad \left(\nabla^{-\operatorname{Col}}_{V,U}\right)_{(\lambda,U,V)\in\operatorname{SBT}(n)}$$

are bases of $\mathcal{A}$. Thus, they span $\mathcal{A}$. In other words,

$$\operatorname{span}\left\{\nabla^{\operatorname{Row}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n)\right\} = \mathcal{A} \qquad \text{and}$$
$$\operatorname{span}\left\{\nabla^{-\operatorname{Col}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n)\right\} = \mathcal{A}.$$

Now, the definition of $\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T_1}$ shows that

$$\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T_1} = \operatorname{span} \underbrace{\left\{\nabla^{\operatorname{Row}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n) \text{ and } U \geq T_1\right\}}_{\substack{=\left\{\nabla^{\operatorname{Row}}_{V,U} \mid (\lambda,U,V)\in\operatorname{SBT}(n)\right\} \\ \text{(here, we have removed the condition ``}U\geq T_1\text{'',} \\ \text{since (523) shows that this condition} \\ \text{is automatically satisfied for each } (\lambda,U,V)\in\operatorname{SBT}(n))}}$$
$$= \operatorname{span}\left\{\nabla^{\operatorname{Row}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n)\right\} = \mathcal{A}.$$

Likewise, the definition of $\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T_m}$ shows that

$$\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T_m} = \operatorname{span} \underbrace{\left\{\nabla^{-\operatorname{Col}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n) \text{ and } U \leq T_m\right\}}_{\substack{=\left\{\nabla^{-\operatorname{Col}}_{V,U} \mid (\lambda,U,V)\in\operatorname{SBT}(n)\right\} \\ \text{(here, we have removed the condition ``}U\leq T_m\text{'',} \\ \text{since (524) shows that this condition} \\ \text{is automatically satisfied for each } (\lambda,U,V)\in\operatorname{SBT}(n))}}$$
$$= \operatorname{span}\left\{\nabla^{-\operatorname{Col}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n)\right\} = \mathcal{A}.$$

Hence, the proof of Proposition 6.8.30 **(a)** is complete.

**(b)** There exists no $(\lambda, U, V) \in \operatorname{SBT}(n)$ that satisfies $U < T_1$ (since every $(\lambda, U, V) \in \operatorname{SBT}(n)$ satisfies $U \geq T_1$ by (523), and thus cannot satisfy $U < T_1$). Now, the definition of $\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T_1}$ shows that

$$\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T_1} = \operatorname{span} \underbrace{\left\{\nabla^{-\operatorname{Col}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n) \text{ and } U < T_1\right\}}_{\substack{=\varnothing \\ \text{(since there exists no } (\lambda,U,V)\in\operatorname{SBT}(n) \text{ that satisfies } U<T_1)}} = \operatorname{span}\varnothing = 0.$$

There exists no $(\lambda, U, V) \in \operatorname{SBT}(n)$ that satisfies $U > T_m$ (since every $(\lambda, U, V) \in \operatorname{SBT}(n)$ satisfies $U \leq T_m$ by (524), and thus cannot satisfy $U > T_m$). Now, the definition of $\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},>T_m}$ shows that

$$\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},>T_m} = \operatorname{span} \underbrace{\left\{ \nabla^{\operatorname{Row}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n) \text{ and } U > T_m \right\}}_{=\varnothing} = \operatorname{span} \varnothing = 0.$$

(since there exists no $(\lambda,U,V) \in \operatorname{SBT}(n)$ that satisfies $U > T_m$)

Thus, the proof of Proposition 6.8.30 **(b)** is complete.

**(c)** The tableaux $T_i$ and $T_{i+1}$ are consecutive entries of the list $(T_1, T_2, \ldots, T_m)$, which contains each standard tableau in $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$. Hence, for an arbitrary standard $n$-tableau $U \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, we have the logical equivalence

$$(U \geq T_{i+1}) \iff (U > T_i) \tag{525}$$

[177]. But the definition of $\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T_{i+1}}$ shows that

$$\begin{aligned}
\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T_{i+1}} &= \operatorname{span} \underbrace{\left\{ \nabla^{\operatorname{Row}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n) \text{ and } U \geq T_{i+1} \right\}}_{= \left\{ \nabla^{\operatorname{Row}}_{V,U} \mid (\lambda,U,V) \in \operatorname{SBT}(n) \text{ and } U > T_i \right\} \text{ (because of the equivalence (525))}} \\
&= \operatorname{span} \left\{ \nabla^{\operatorname{Row}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n) \text{ and } U > T_i \right\} \\
&= \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},>T_i} \qquad \left( \text{by the definition of } \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},>T_i} \right).
\end{aligned}$$

Hence, we have proved that $\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T_{i+1}} = \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},>T_i}$. Similarly, we can show that $\mathcal{F}^{\operatorname{Row}}_{\operatorname{all},\geq T_{i+1}} = \mathcal{F}^{\operatorname{Row}}_{\operatorname{all},>T_i}$. The remaining two equalities $\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T_i} = \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T_{i+1}}$ and $\mathcal{F}^{-\operatorname{Col}}_{\operatorname{all},\leq T_i} = \mathcal{F}^{-\operatorname{Col}}_{\operatorname{all},<T_{i+1}}$ can be proved in an analogous way, where instead of the equivalence (525) we must now use the equivalence

$$(U \leq T_i) \iff (U < T_{i+1}). \tag{526}$$

[177] *Proof.* Let $U \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ be arbitrary. Then, $U = T_j$ for some $j \in [m]$ (since $T_1, T_2, \ldots, T_m$ are all the standard tableaux in the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$). Consider this $j$.

But the list $(T_1, T_2, \ldots, T_m)$ is ordered in increasing order (since this is how we defined $T_1, T_2, \ldots, T_m$). Hence, we have $T_j > T_i$ if and only if $j > i$. For the same reason, we have $T_j \geq T_{i+1}$ if and only if $j \geq i+1$. Thus, we have the following chain of logical equivalences:

$$\begin{aligned}
(T_j \geq T_{i+1}) &\iff (j \geq i+1) \iff (j > i) \qquad (\text{since } j \text{ and } i \text{ are integers}) \\
&\iff (T_j > T_i) \qquad (\text{since we have } T_j > T_i \text{ if and only if } j > i).
\end{aligned}$$

In other words,

$$(U \geq T_{i+1}) \iff (U > T_i)$$

(since $U = T_j$). Qed.

(The proof of the equivalence (526) is similar to the proof of (525); alternatively, it can be easily derived from (525)[178].) Altogether, all claims of Proposition 6.8.30 **(c)** are now proved. □

The following two facts are also easy to see:

**Proposition 6.8.31.** Let $T \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ be a standard $n$-tableau. Then, the four sets $\mathcal{F}^{\operatorname{Row}}_{\operatorname{all},\geq T}$, $\mathcal{F}^{\operatorname{Row}}_{\operatorname{all},>T}$, $\mathcal{F}^{-\operatorname{Col}}_{\operatorname{all},\leq T}$ and $\mathcal{F}^{-\operatorname{Col}}_{\operatorname{all},<T}$ are left $\mathcal{A}$-submodules of $\mathcal{A}$ (that is, subrepresentations of the left regular representation $\mathcal{A}$ of $S_n$).

*Proof.* Here is a sketch of the (easy) proof; details can be proved in the appendix (Section A.6).

We shall only prove the claim about $\mathcal{F}^{\operatorname{Row}}_{\operatorname{all},\geq T}$, since the other three claims are analogous. We thus need to show that $\mathcal{F}^{\operatorname{Row}}_{\operatorname{all},\geq T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$, that is, a subrepresentation of the $S_n$-representation $\mathcal{A}$ (since $\mathcal{A} = \mathbf{k}[S_n]$).

We defined the $\mathbf{k}$-module $\mathcal{F}^{\operatorname{Row}}_{\operatorname{all},\geq T}$ as the span of the vectors $\nabla^{\operatorname{Row}}_{V,U}$ for $(\lambda, U, V) \in \operatorname{HSBT}(n)$ satisfying $U \geq T$; let us refer to these vectors as the "good nablas". But every $w \in S_n$ and every $(\lambda, U, V) \in \operatorname{HSBT}(n)$ satisfy

$$\nabla^{\operatorname{Row}}_{wV,U} = w \nabla^{\operatorname{Row}}_{V,U}$$

(by (494), applied to $P = V$ and $Q = U$ and $g = w$ and $h = \mathrm{id}$). Hence, if $\nabla^{\operatorname{Row}}_{V,U}$ is a good nabla and $w \in S_n$ is a permutation, then $w\nabla^{\operatorname{Row}}_{V,U}$ is again a good nabla (namely, $\nabla^{\operatorname{Row}}_{wV,U}$). Thus, the span of all good nablas is an $S_n$-subset of $\mathcal{A}$, and hence a subrepresentation of the $S_n$-representation $\mathcal{A}$ (since it is a $\mathbf{k}$-submodule of $\mathcal{A}$ as well). But this span is $\mathcal{F}^{\operatorname{Row}}_{\operatorname{all},\geq T}$. Hence, $\mathcal{F}^{\operatorname{Row}}_{\operatorname{all},\geq T}$ is a subrepresentation of the $S_n$-representation $\mathcal{A}$, just as we wanted to show. □

**Proposition 6.8.32.** Let $T \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ be a standard $n$-tableau. Then,

$$\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},>T} \subseteq \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T} \qquad \text{and} \qquad \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T} \subseteq \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T}.$$

[178] Here is how: Let $U \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ be arbitrary. Then, we have the logical equivalence $(U \leq T_i) \iff (\text{not } U > T_i)$ (since the order on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ is a total order) and the logical equivalence $(U < T_{i+1}) \iff (\text{not } U \geq T_{i+1})$ (similarly). Hence, we have the following chain of equivalences:

$$\begin{aligned}(U < T_{i+1}) &\iff (\text{not } U \geq T_{i+1}) \iff (\text{not } U > T_i) \qquad (\text{by (525)}) \\ &\iff (U \leq T_i) \qquad (\text{since } (U \leq T_i) \iff (\text{not } U > T_i)).\end{aligned}$$

In other words, $(U \leq T_i) \iff (U < T_{i+1})$. This proves (526).

*Proof.* We have

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T} = \operatorname{span}\underbrace{\left\{\nabla^{\mathrm{Row}}_{V,U} \;\mid\; (\lambda,U,V)\in \mathrm{SBT}(n) \text{ and } U>T\right\}}_{\subseteq \left\{\nabla^{\mathrm{Row}}_{V,U} \mid (\lambda,U,V)\in \mathrm{SBT}(n) \text{ and } U\geq T\right\} \\ \text{(since the condition ``}U>T\text{'' implies ``}U\geq T\text{'')}}$$

$$\left(\text{by the definition of } \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T}\right)$$

$$\subseteq \operatorname{span}\left\{\nabla^{\mathrm{Row}}_{V,U} \;\mid\; (\lambda,U,V)\in \mathrm{SBT}(n) \text{ and } U\geq T\right\} = \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$$

(by the definition of $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$). Thus, $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T} \subseteq \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$ is proved. A similar argument establishes $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} \subseteq \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$. This completes the proof of Proposition 6.8.32. □

To proceed, we need a variant of Lemma 6.8.23:

**Lemma 6.8.33** (Murphy orthogonality lemma, take 2)**.** Let $(\lambda, A, B) \in \mathrm{HSBT}(n)$ and $(\mu, C, D) \in \mathrm{HSBT}(n)$. Assume that $A < C$, where the symbol "$<$" refers to the relation $<$ from Definition 6.7.5. Then,

$$\left\langle \nabla^{-\mathrm{Col}}_{B,A},\ \nabla^{\mathrm{Row}}_{D,C} \right\rangle = 0.$$

*Proof.* This lemma is very similar to Lemma 6.8.23 except for two differences: We are no longer requiring that the tableaux $B$ and $D$ are standard (since $(\lambda, A, B)$ and $(\mu, C, D)$ now belong to $\mathrm{HSBT}(n)$ rather than to $\mathrm{SBT}(n)$), but in exchange, we are now requiring that $A < C$ (rather than merely $A < C$ or $B < D$). Fortunately, the proof of Lemma 6.8.23 still works, except that Case 2 no longer exists. □

An easy consequence of Lemma 6.8.33 is the following:

**Lemma 6.8.34.** Let $T \in \bigcup_{\kappa\vdash n} \mathrm{SYT}(\kappa)$ be a standard $n$-tableau. Then:

**(a)** We have $\langle \mathbf{a}, \mathbf{b}\rangle = 0$ for any $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},<T}$ and $\mathbf{b} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$.

**(b)** We have $\langle \mathbf{a}, \mathbf{b}\rangle = 0$ for any $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},\leq T}$ and $\mathbf{b} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},>T}$.

*Proof.* **(a)** Let $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},<T}$ and $\mathbf{b} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$. We must prove that $\langle \mathbf{a}, \mathbf{b}\rangle = 0$.

This claim depends $\mathbf{k}$-linearly on each of $\mathbf{a}$ and $\mathbf{b}$ (because $\langle \cdot, \cdot\rangle$ is a bilinear form). Hence, we can WLOG assume

- that $\mathbf{a}$ is one of the vectors $\nabla^{-\mathrm{Col}}_{V,U}$ with $(\lambda, U, V) \in \mathrm{HSBT}(n)$ satisfying $U < T$ (since $\mathbf{a}$ belongs to $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},<T}$, which is defined to be the span of these vectors), and

- that $\mathbf{b}$ is one of the vectors $\nabla^{\mathrm{Row}}_{V,U}$ with $(\lambda, U, V) \in \mathrm{HSBT}(n)$ satisfying $U \geq T$ (since $\mathbf{b}$ belongs to $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$, which is defined to be the span of these vectors).

Assume this. Thus,

$$\mathbf{a} = \nabla^{-\mathrm{Col}}_{B,A} \qquad \text{for some } (\lambda, A, B) \in \mathrm{HSBT}(n) \text{ satisfying } A < T,$$

and

$$\mathbf{b} = \nabla^{\mathrm{Row}}_{D,C} \qquad \text{for some } (\mu, C, D) \in \mathrm{HSBT}(n) \text{ satisfying } C \geq T.$$

Consider these $(\lambda, A, B)$ and $(\mu, C, D)$. Then, $A < T \leq C$ (since $C \geq T$), so that $A < C$ (since the relation $<$ is a total order). Thus, $\left\langle \nabla^{-\mathrm{Col}}_{B,A}, \nabla^{\mathrm{Row}}_{D,C} \right\rangle = 0$ (by Lemma 6.8.33). In other words, $\langle \mathbf{a}, \mathbf{b} \rangle = 0$ (since $\mathbf{a} = \nabla^{-\mathrm{Col}}_{B,A}$ and $\mathbf{b} = \nabla^{\mathrm{Row}}_{D,C}$). This proves Lemma 6.8.34 **(a)**.

**(b)** This is analogous to part **(a)**. (The main difference is that the chain of inequalities $A < T \leq C$ must now be replaced by $A \leq T < C$.) □

**Lemma 6.8.35.** Let $T \in \bigcup_{\kappa \vdash n} \mathrm{SYT}(\kappa)$ be a standard $n$-tableau. Then:

**(a)** We have $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T} \subseteq \left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right)^{\perp}$.

**(b)** We have $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},>T} \subseteq \left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}\right)^{\perp}$.

**(c)** We have $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},<T} \subseteq \left(\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}\right)^{\perp}$.

**(d)** We have $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},\leq T} \subseteq \left(\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T}\right)^{\perp}$.

*Proof.* **(a)** We must prove that $\mathbf{b} \in \left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right)^{\perp}$ for each $\mathbf{b} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$.

So let $\mathbf{b} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ be arbitrary. We must then prove that $\mathbf{b} \in \left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right)^{\perp}$. In other words, we must prove that $\langle \mathbf{b}, \mathbf{v} \rangle = 0$ for each $\mathbf{v} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ (because $\left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right)^{\perp}$ is defined as $\left\{\mathbf{a} \in \mathcal{A} \mid \langle \mathbf{a}, \mathbf{v} \rangle = 0 \text{ for each } \mathbf{v} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right\}$).

So let $\mathbf{v} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ be arbitrary. We must prove that $\langle \mathbf{b}, \mathbf{v} \rangle = 0$.

We have

$$\mathbf{v} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} = \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{V,U} \ \mid \ (\lambda, U, V) \in \underbrace{\mathrm{SBT}(n)}_{\subseteq \mathrm{HSBT}(n)} \text{ and } U < T \right\}$$
$$\left(\text{by the definition of } \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right)$$
$$\subseteq \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{V,U} \ \mid \ (\lambda, U, V) \in \mathrm{HSBT}(n) \text{ and } U < T \right\} = \mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},<T}$$

(by the definition of $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},\leq T}$). Hence, Lemma 6.8.34 **(a)** (applied to $\mathbf{a} = \mathbf{v}$) yields $\langle \mathbf{v}, \mathbf{b} \rangle = 0$. Finally, (513) shows that $\langle \mathbf{v}, \mathbf{b} \rangle = \langle \mathbf{b}, \mathbf{v} \rangle$, so that $\langle \mathbf{b}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{b} \rangle = 0$. As we explained above, this completes the proof of Lemma 6.8.35 **(a)**.

**(b)** This is analogous to part **(a)**, but using Lemma 6.8.34 **(b)** instead of Lemma 6.8.34 **(a)**.

**(c)** This proof, too, is similar to that of part **(a)**, but there are some minor changes, so we give it in full.

We must prove that $\mathbf{b} \in \left(\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}\right)^{\perp}$ for each $\mathbf{b} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},<T}$.

So let $\mathbf{b} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},<T}$ be arbitrary. We must then prove that $\mathbf{b} \in \left(\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}\right)^{\perp}$. In other words, we must prove that $\langle \mathbf{b}, \mathbf{v} \rangle = 0$ for each $\mathbf{v} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$ (because $\left(\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}\right)^{\perp}$ is defined as $\left\{\mathbf{a} \in \mathcal{A} \ \mid \ \langle \mathbf{a}, \mathbf{v} \rangle = 0 \text{ for each } \mathbf{v} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}\right\}$).

So let $\mathbf{v} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$ be arbitrary. We must prove that $\langle \mathbf{b}, \mathbf{v} \rangle = 0$.

We have

$$\mathbf{v} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} = \operatorname{span}\left\{ \nabla^{\mathrm{Row}}_{V,U} \ \mid \ (\lambda, U, V) \in \underbrace{\mathrm{SBT}(n)}_{\subseteq \mathrm{HSBT}(n)} \text{ and } U \geq T \right\}$$
$$\left(\text{by the definition of } \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}\right)$$
$$\subseteq \operatorname{span}\left\{ \nabla^{\mathrm{Row}}_{V,U} \ \mid \ (\lambda, U, V) \in \mathrm{HSBT}(n) \text{ and } U \geq T \right\} = \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$$

(by the definition of $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$). Hence, Lemma 6.8.34 **(a)** (applied to $\mathbf{b}$ and $\mathbf{v}$ instead of $\mathbf{a}$ and $\mathbf{b}$) yields $\langle \mathbf{b}, \mathbf{v} \rangle = 0$. As we explained above, this completes the proof of Lemma 6.8.35 **(c)**.

**(d)** This is analogous to part **(c)**, but using Lemma 6.8.34 **(b)** instead of Lemma 6.8.34 **(a)**. □

**Lemma 6.8.36.** Let $T \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ be a standard $n$-tableau. Then:

**(a)** We have $\left(\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}\right)^{\perp} \subseteq \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}$.

**(b)** We have $\left(\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T}\right)^{\perp} \subseteq \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},>T}$.

**(c)** We have $\left(\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}\right)^{\perp} \subseteq \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}$.

**(d)** We have $\left(\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},>T}\right)^{\perp} \subseteq \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T}$.

*Proof.* **(a)** We must show that $\mathbf{a} \in \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}$ for each $\mathbf{a} \in \left(\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}\right)^{\perp}$.

So let $\mathbf{a} \in \left(\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}\right)^{\perp}$. We must show that $\mathbf{a} \in \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}$.

We have $\mathbf{a} \in \left(\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}\right)^{\perp} \subseteq \mathcal{A}$. Since the row Murphy basis $\left(\nabla^{\operatorname{Row}}_{V,U}\right)_{(\lambda,U,V)\in\operatorname{SBT}(n)}$ is a basis of $\mathcal{A}$ (by Theorem 6.8.14), we can thus write $\mathbf{a}$ as a $\mathbf{k}$-linear combination of its elements. In other words, we can write $\mathbf{a}$ as

$$\mathbf{a} = \sum_{(\lambda,U,V)\in\operatorname{SBT}(n)} \omega_{\lambda,U,V} \nabla^{\operatorname{Row}}_{V,U} \tag{527}$$

for some scalars $\omega_{\lambda,U,V} \in \mathbf{k}$. Consider these $\omega_{\lambda,U,V}$. We can rewrite the equality (527) as

$$\mathbf{a} = \sum_{(\mu,P,Q)\in\operatorname{SBT}(n)} \omega_{\mu,P,Q} \nabla^{\operatorname{Row}}_{Q,P} \tag{528}$$

(by renaming the summation index $(\lambda, U, V)$ as $(\mu, P, Q)$).

Now, we shall show the following:

*Claim 1:* We have

$$\omega_{\lambda,U,V} = 0 \qquad \text{for all } (\lambda, U, V) \in \operatorname{SBT}(n) \text{ satisfying } U < T.$$

*Proof of Claim 1.* We proceed by strong induction on $(\lambda, U, V)$ (using the total order on $\operatorname{SBT}(n)$ introduced in Definition 6.7.12). So we fix some $(\lambda, U, V) \in \operatorname{SBT}(n)$ satisfying $U < T$. We assume (as the induction hypothesis) that

$$\omega_{\mu,P,Q} = 0 \tag{529}$$

for all $(\mu, P, Q) \in \operatorname{SBT}(n)$ satisfying $P < T$ and $(\mu, P, Q) < (\lambda, U, V)$. Our goal is to prove that $\omega_{\lambda,U,V} = 0$.

Since $(\lambda, U, V) \in \operatorname{SBT}(n)$ and $U < T$, we have $\nabla^{-\operatorname{Col}}_{V,U} \in \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}$ by the definition of $\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}$ (in fact, $\nabla^{-\operatorname{Col}}_{V,U}$ is one of the spanning vectors in the definition of $\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}$).

Since $\mathbf{a} \in \left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right)^{\perp}$, we have $\langle \mathbf{a}, \mathbf{v} \rangle = 0$ for each $\mathbf{v} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ (by the definition of $\left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right)^{\perp}$). Applying this to $\mathbf{v} = \nabla^{-\mathrm{Col}}_{V,U}$, we obtain $\left\langle \mathbf{a},\ \nabla^{-\mathrm{Col}}_{V,U} \right\rangle = 0$ (since $\nabla^{-\mathrm{Col}}_{V,U} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$). But (513) shows that $\left\langle \mathbf{a},\ \nabla^{-\mathrm{Col}}_{V,U} \right\rangle = \left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \mathbf{a} \right\rangle$, hence $\left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \mathbf{a} \right\rangle = \left\langle \mathbf{a},\ \nabla^{-\mathrm{Col}}_{V,U} \right\rangle = 0$. Thus,

$$0 = \left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \mathbf{a} \right\rangle = \left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \sum_{(\mu,P,Q)\in \mathrm{SBT}(n)} \omega_{\mu,P,Q} \nabla^{\mathrm{Row}}_{Q,P} \right\rangle \qquad \text{(by (528))}$$

$$= \sum_{(\mu,P,Q)\in \mathrm{SBT}(n)} \omega_{\mu,P,Q} \left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{Q,P} \right\rangle \tag{530}$$

(since the form $\langle \cdot, \cdot \rangle$ is bilinear). Now, we observe the following:

- For each $(\mu, P, Q) \in \mathrm{SBT}(n)$ satisfying $(\mu, P, Q) < (\lambda, U, V)$, we have

$$\omega_{\mu,P,Q} = 0. \tag{531}$$

  [*Proof:* Let $(\mu, P, Q) \in \mathrm{SBT}(n)$ be such that $(\mu, P, Q) < (\lambda, U, V)$. Then, Lemma 6.7.14 **(a)** (applied to $(\mu, P, Q)$ and $(\lambda, U, V)$ instead of $(\lambda, A, B)$ and $(\mu, C, D)$) shows that $P \leq U$. Thus, $P \leq U < T$, so that $P < T$ (since the relation $<$ is a total order). Thus, (529) yields $\omega_{\mu,P,Q} = 0$. This proves (531).]

- For each $(\mu, P, Q) \in \mathrm{SBT}(n)$ satisfying $(\mu, P, Q) > (\lambda, U, V)$, we have

$$\left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{Q,P} \right\rangle = 0. \tag{532}$$

  [*Proof:* Let $(\mu, P, Q) \in \mathrm{SBT}(n)$ be such that $(\mu, P, Q) > (\lambda, U, V)$. Thus, $(\lambda, U, V) < (\mu, P, Q)$. Hence, Lemma 6.7.14 **(b)** (applied to $(\lambda, U, V)$ and $(\mu, P, Q)$ instead of $(\lambda, A, B)$ and $(\mu, C, D)$) yields $U < P$ or $V < Q$. Thus, Lemma 6.8.23 (applied to $A = U$, $B = V$, $C = P$ and $D = Q$) yields $\left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{Q,P} \right\rangle = 0$. This proves (532).]

Now, recall that the relation $<$ on $\mathrm{SBT}(n)$ is the smaller relation of a total order. Hence, each $(\mu, P, Q) \in \mathrm{SBT}(n)$ satisfies exactly one of the three statements $(\mu, P, Q) < (\lambda, U, V)$ and $(\mu, P, Q) > (\lambda, U, V)$ and $(\mu, P, Q) = (\lambda, U, V)$.

Accordingly, the sum on the right hand side of (530) can be split into three parts: the part comprising all $(\mu, P, Q) \in \mathrm{SBT}(n)$ that satisfy $(\mu, P, Q) < (\lambda, U, V)$; the part comprising all $(\mu, P, Q) \in \mathrm{SBT}(n)$ that satisfy $(\mu, P, Q) > (\lambda, U, V)$; and a single addend corresponding to $(\mu, P, Q) = (\lambda, U, V)$. Splitting

the sum in this way, we rewrite (530) as follows:

$$\begin{aligned}
0 &= \sum_{\substack{(\mu,P,Q)\in \mathrm{SBT}(n);\\ (\mu,P,Q)<(\lambda,U,V)}} \underbrace{\omega_{\mu,P,Q}}_{\substack{=0\\ \text{(by (531))}}} \left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{Q,P} \right\rangle \\
&\qquad + \sum_{\substack{(\mu,P,Q)\in \mathrm{SBT}(n);\\ (\mu,P,Q)>(\lambda,U,V)}} \omega_{\mu,P,Q} \underbrace{\left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{Q,P} \right\rangle}_{\substack{=0\\ \text{(by (532))}}} + \omega_{\lambda,U,V} \left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{V,U} \right\rangle \\
&= \underbrace{\sum_{\substack{(\mu,P,Q)\in \mathrm{SBT}(n);\\ (\mu,P,Q)<(\lambda,U,V)}} 0 \left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{Q,P} \right\rangle}_{=0} + \underbrace{\sum_{\substack{(\mu,P,Q)\in \mathrm{SBT}(n);\\ (\mu,P,Q)>(\lambda,U,V)}} \omega_{\mu,P,Q} 0}_{=0} + \omega_{\lambda,U,V} \left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{V,U} \right\rangle \\
&= \omega_{\lambda,U,V} \underbrace{\left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{V,U} \right\rangle}_{\substack{=\pm 1\\ \text{(by Lemma 6.8.22,}\\ \text{applied to } A=U \text{ and } B=V)}} = \pm\omega_{\lambda,U,V}.
\end{aligned}$$

Hence, $\pm\omega_{\lambda,U,V} = 0$. In other words, $\omega_{\lambda,U,V} = 0$. This completes the induction, and thus the proof of Claim 1. □

Now, (527) becomes

$$\begin{aligned}
\mathbf{a} &= \sum_{(\lambda,U,V)\in \mathrm{SBT}(n)} \omega_{\lambda,U,V} \nabla^{\mathrm{Row}}_{V,U} \\
&= \sum_{\substack{(\lambda,U,V)\in \mathrm{SBT}(n);\\ U<T}} \underbrace{\omega_{\lambda,U,V}}_{\substack{=0\\ \text{(by Claim 1)}}} \nabla^{\mathrm{Row}}_{V,U} + \sum_{\substack{(\lambda,U,V)\in \mathrm{SBT}(n);\\ U\geq T}} \omega_{\lambda,U,V} \nabla^{\mathrm{Row}}_{V,U} \\
&\qquad\qquad \left( \begin{array}{c} \text{since each } (\lambda,U,V)\in \mathrm{SBT}(n) \text{ satisfies} \\ \text{either } U<T \text{ or } U\geq T \text{ (but not both)} \end{array} \right) \\
&= \underbrace{\sum_{\substack{(\lambda,U,V)\in \mathrm{SBT}(n);\\ U<T}} 0 \nabla^{\mathrm{Row}}_{V,U}}_{=0} + \sum_{\substack{(\lambda,U,V)\in \mathrm{SBT}(n);\\ U\geq T}} \omega_{\lambda,U,V} \nabla^{\mathrm{Row}}_{V,U} \\
&= \sum_{\substack{(\lambda,U,V)\in \mathrm{SBT}(n);\\ U\geq T}} \omega_{\lambda,U,V} \nabla^{\mathrm{Row}}_{V,U} \\
&\in \operatorname{span}\left\{ \nabla^{\mathrm{Row}}_{V,U} \ \mid\ (\lambda,U,V)\in \mathrm{SBT}(n) \text{ and } U\geq T \right\} \\
&= \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} \qquad \left( \text{by the definition of } \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} \right).
\end{aligned}$$

This completes our proof of Lemma 6.8.36 **(a)**.

**(b)** This is analogous to the above proof of Lemma 6.8.36 **(a)**; the only difference is that some inequality signs (between standard $n$-tableaux) change from

strong to weak and vice versa. (For instance, the "$U < T$" in Claim 1 must be replaced by "$U \leq T$"). We leave it to the reader to make the necessary changes.

**(c)** This is sufficiently similar to part **(a)** that I am omitting the proof here, but still sufficiently different that I am nevertheless giving the proof in the appendix (Section A.6).

**(d)** Just as for part **(c)**, see the appendix (Section A.6). □

We are now ready to prove Theorem 6.8.29:

*Proof of Theorem 6.8.29.* **(a)** We have $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} \subseteq \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ (this follows from the definitions of these two sets[179]). Hence,

$$\begin{aligned}\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} \subseteq \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T} &\subseteq \left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right)^{\perp} \qquad \text{(by Lemma 6.8.35 \textbf{(a)})} \\ &\subseteq \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} \qquad \text{(by Lemma 6.8.36 \textbf{(a)})}.\end{aligned}$$

This is a chain of inclusions that starts and ends with the same set (namely, $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$). Thus, all inclusion signs in this chain must actually be equality signs (since any three sets $X, Y, Z$ satisfying $X \subseteq Y \subseteq Z \subseteq X$ must satisfy $X = Y = Z$). That is, we must have $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} = \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T} = \left(\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}\right)^{\perp}$.

It remains to show that $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$.

But Proposition 6.8.31 shows that $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$. In other words, $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$ (since $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} = \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$). Hence, the proof of Theorem 6.8.29 **(a)** is complete.

**(b)** The proof of part **(a)** applies to part **(b)** with minor changes (the inequality signs need to be changed, and we must use Lemma 6.8.35 **(b)** instead of Lemma 6.8.35 **(a)** and use Lemma 6.8.36 **(b)** instead of Lemma 6.8.36 **(a)**).

**(c)** The proof of part **(a)** applies to part **(c)** with minor changes (the inequality signs need to be changed, and we must use Lemma 6.8.35 **(c)** instead of Lemma 6.8.35 **(a)** and use Lemma 6.8.36 **(c)** instead of Lemma 6.8.36 **(a)**).

[179] *Proof.* The definition of $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$ yields

$$\begin{aligned}\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T} &= \operatorname{span}\left\{\nabla^{\mathrm{Row}}_{V,U} \;\mid\; (\lambda, U, V) \in \underbrace{\operatorname{SBT}(n)}_{\subseteq \operatorname{HSBT}(n)} \text{ and } U \geq T\right\} \\ &\subseteq \operatorname{span}\left\{\nabla^{\mathrm{Row}}_{V,U} \;\mid\; (\lambda, U, V) \in \operatorname{HSBT}(n) \text{ and } U \geq T\right\} = \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}\end{aligned}$$

(by the definition of $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$).

**(d)** The proof of part **(a)** applies to part **(d)** with minor changes (the inequality signs need to be changed, and we must use Lemma 6.8.35 **(d)** instead of Lemma 6.8.35 **(a)** and use Lemma 6.8.36 **(d)** instead of Lemma 6.8.36 **(a)**). □

### 6.8.9. The Murphy span filtrations

We shall now reformulate the results of the previous subsection (mainly Theorem 6.8.29) as two filtrations of the regular representation of $S_n$ (that is, the left $\mathcal{A}$-module $\mathcal{A}$) by left $\mathcal{A}$-submodules (the row- or column-Murphy spans). We will then recognize the subquotients of these filtrations as the Specht modules or their duals.

The two filtrations are defined in the following theorem:

**Theorem 6.8.37.** Let $T_1, T_2, \ldots, T_m$ be all the standard tableaux in the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, listed in increasing order (with respect to the total order defined in Definition 6.7.7). For each $i \in \{0, 1, \ldots, m\}$, we define the $\mathbf{k}$-submodules

$$\mathcal{F}_i^{\operatorname{Row}} := \begin{cases} \mathcal{F}^{\operatorname{Row}}_{\operatorname{std}, \geq T_{i+1}}, & \text{if } i < m; \\ 0, & \text{if } i = m \end{cases} \quad \text{and}$$

$$\mathcal{F}_i^{-\operatorname{Col}} := \begin{cases} \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std}, \leq T_i}, & \text{if } i > 0; \\ 0, & \text{if } i = 0 \end{cases} \quad \text{of } \mathcal{A}.$$

Then:

**(a)** We have $\mathcal{F}_i^{\operatorname{Row}} = \begin{cases} \mathcal{F}^{\operatorname{Row}}_{\operatorname{std}, > T_i}, & \text{if } i > 0; \\ \mathcal{A}, & \text{if } i = 0 \end{cases}$ for each $i \in \{0, 1, \ldots, m\}$.

**(b)** We have $\mathcal{F}_i^{-\operatorname{Col}} = \begin{cases} \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std}, < T_{i+1}}, & \text{if } i < m; \\ \mathcal{A}, & \text{if } i = m \end{cases}$ for each $i \in \{0, 1, \ldots, m\}$.

**(c)** For each $i \in \{0, 1, \ldots, m\}$, the set $\mathcal{F}_i^{\operatorname{Row}}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$.

**(d)** For each $i \in \{0, 1, \ldots, m\}$, the set $\mathcal{F}_i^{-\operatorname{Col}}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$.

**(e)** We have

$$0 = \mathcal{F}_m^{\operatorname{Row}} \subseteq \mathcal{F}_{m-1}^{\operatorname{Row}} \subseteq \cdots \subseteq \mathcal{F}_2^{\operatorname{Row}} \subseteq \mathcal{F}_1^{\operatorname{Row}} \subseteq \mathcal{F}_0^{\operatorname{Row}} = \mathcal{A}.$$

That is, $\left(\mathcal{F}_m^{\operatorname{Row}}, \mathcal{F}_{m-1}^{\operatorname{Row}}, \ldots, \mathcal{F}_2^{\operatorname{Row}}, \mathcal{F}_1^{\operatorname{Row}}, \mathcal{F}_0^{\operatorname{Row}}\right)$ is a filtration of $\mathcal{A}$ by left $\mathcal{A}$-submodules.

**(f)** We have

$$0 = \mathcal{F}_0^{-\operatorname{Col}} \subseteq \mathcal{F}_1^{-\operatorname{Col}} \subseteq \mathcal{F}_2^{-\operatorname{Col}} \subseteq \cdots \subseteq \mathcal{F}_{m-1}^{-\operatorname{Col}} \subseteq \mathcal{F}_m^{-\operatorname{Col}} = \mathcal{A}.$$

That is, $\left(\mathcal{F}_0^{-\mathrm{Col}}, \mathcal{F}_1^{-\mathrm{Col}}, \mathcal{F}_2^{-\mathrm{Col}}, \ldots, \mathcal{F}_{m-1}^{-\mathrm{Col}}, \mathcal{F}_m^{-\mathrm{Col}}\right)$ is a filtration of $\mathcal{A}$ by left $\mathcal{A}$-submodules.

**Example 6.8.38.** Let $n = 3$. Recall that the set $\bigcup_{\kappa \vdash n} \mathrm{SYT}(\kappa)$ consists of just four standard tableaux, which are ordered as follows:

$$1\backslash\backslash 2\backslash\backslash 3 < 13\backslash\backslash 2 < 12\backslash\backslash 3 < 123$$

(see Example 6.7.8). Thus, following the notations of Theorem 6.8.37, we have $m = 4$ and $T_1 = 1\backslash\backslash 2\backslash\backslash 3$ and $T_2 = 13\backslash\backslash 2$ and $T_3 = 12\backslash\backslash 3$ and $T_4 = 123$.

Recall also the six standard $n$-bitableaux in $\mathrm{SBT}(n)$ (see Example 6.7.13 for a list).

Now, the filtration $\left(\mathcal{F}_m^{\mathrm{Row}}, \mathcal{F}_{m-1}^{\mathrm{Row}}, \ldots, \mathcal{F}_2^{\mathrm{Row}}, \mathcal{F}_1^{\mathrm{Row}}, \mathcal{F}_0^{\mathrm{Row}}\right)$ in Theorem 6.8.37 **(e)** takes the form $\left(\mathcal{F}_4^{\mathrm{Row}}, \mathcal{F}_3^{\mathrm{Row}}, \mathcal{F}_2^{\mathrm{Row}}, \mathcal{F}_1^{\mathrm{Row}}, \mathcal{F}_0^{\mathrm{Row}}\right)$ with

$$\begin{aligned}\mathcal{F}_0^{\mathrm{Row}} &= \mathcal{F}_{\mathrm{std},\geq T_1}^{\mathrm{Row}} = \mathcal{A} \\ &= \mathrm{span}\left\{\nabla^{\mathrm{Row}}_{1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3}, \nabla^{\mathrm{Row}}_{13\backslash\backslash 2,\ 13\backslash\backslash 2}, \nabla^{\mathrm{Row}}_{13\backslash\backslash 2,\ 12\backslash\backslash 3}, \nabla^{\mathrm{Row}}_{12\backslash\backslash 3,\ 13\backslash\backslash 2}, \right. \\ &\qquad \left. \nabla^{\mathrm{Row}}_{12\backslash\backslash 3,\ 12\backslash\backslash 3}, \nabla^{\mathrm{Row}}_{123,\ 123}\right\};\end{aligned}$$

$$\begin{aligned}\mathcal{F}_1^{\mathrm{Row}} &= \mathcal{F}_{\mathrm{std},\geq T_2}^{\mathrm{Row}} = \mathcal{F}_{\mathrm{std},> T_1}^{\mathrm{Row}} \\ &= \mathrm{span}\left\{\nabla^{\mathrm{Row}}_{13\backslash\backslash 2,\ 13\backslash\backslash 2}, \nabla^{\mathrm{Row}}_{13\backslash\backslash 2,\ 12\backslash\backslash 3}, \nabla^{\mathrm{Row}}_{12\backslash\backslash 3,\ 13\backslash\backslash 2}, \nabla^{\mathrm{Row}}_{12\backslash\backslash 3,\ 12\backslash\backslash 3}, \right. \\ &\qquad \left. \nabla^{\mathrm{Row}}_{123,\ 123}\right\};\end{aligned}$$

$$\mathcal{F}_2^{\mathrm{Row}} = \mathcal{F}_{\mathrm{std},\geq T_3}^{\mathrm{Row}} = \mathcal{F}_{\mathrm{std},> T_2}^{\mathrm{Row}} = \mathrm{span}\left\{\nabla^{\mathrm{Row}}_{13\backslash\backslash 2,\ 12\backslash\backslash 3}, \nabla^{\mathrm{Row}}_{12\backslash\backslash 3,\ 12\backslash\backslash 3}, \nabla^{\mathrm{Row}}_{123,\ 123}\right\};$$

$$\mathcal{F}_3^{\mathrm{Row}} = \mathcal{F}_{\mathrm{std},\geq T_4}^{\mathrm{Row}} = \mathcal{F}_{\mathrm{std},> T_3}^{\mathrm{Row}} = \mathrm{span}\left\{\nabla^{\mathrm{Row}}_{123,\ 123}\right\};$$

$$\mathcal{F}_4^{\mathrm{Row}} = 0 = \mathcal{F}_{\mathrm{std},> T_4}^{\mathrm{Row}} = \mathrm{span}\left\{\right\}.$$

Thus, this filtration is precisely the filtration (117). We described its subquotients back in Subsection 4.4.4.

Likewise, the filtration $\left(\mathcal{F}_0^{-\mathrm{Col}}, \mathcal{F}_1^{-\mathrm{Col}}, \mathcal{F}_2^{-\mathrm{Col}}, \ldots, \mathcal{F}_{m-1}^{-\mathrm{Col}}, \mathcal{F}_m^{-\mathrm{Col}}\right)$ in Theorem 6.8.37 **(f)** takes the form $\left(\mathcal{F}_0^{-\mathrm{Col}}, \mathcal{F}_1^{-\mathrm{Col}}, \mathcal{F}_2^{-\mathrm{Col}}, \mathcal{F}_3^{-\mathrm{Col}}, \mathcal{F}_4^{-\mathrm{Col}}\right)$ with

$$\mathcal{F}_0^{-\mathrm{Col}} = 0 = \mathcal{F}_{\mathrm{std},< T_1}^{-\mathrm{Col}} = \mathrm{span}\left\{\right\};$$

$$\mathcal{F}_1^{-\mathrm{Col}} = \mathcal{F}_{\mathrm{std},\leq T_1}^{-\mathrm{Col}} = \mathcal{F}_{\mathrm{std},<T_2}^{-\mathrm{Col}} = \operatorname{span}\left\{\nabla^{-\mathrm{Col}}_{1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3}\right\};$$

$$\begin{aligned}\mathcal{F}_2^{-\mathrm{Col}} &= \mathcal{F}_{\mathrm{std},\leq T_2}^{-\mathrm{Col}} = \mathcal{F}_{\mathrm{std},<T_3}^{-\mathrm{Col}} \\ &= \operatorname{span}\left\{\nabla^{-\mathrm{Col}}_{1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3'},\ \nabla^{-\mathrm{Col}}_{13\backslash\backslash 2,\ 13\backslash\backslash 2'},\ \nabla^{-\mathrm{Col}}_{12\backslash\backslash 3,\ 13\backslash\backslash 2}\right\};\end{aligned}$$

$$\begin{aligned}\mathcal{F}_3^{-\mathrm{Col}} &= \mathcal{F}_{\mathrm{std},\leq T_3}^{-\mathrm{Col}} = \mathcal{F}_{\mathrm{std},<T_4}^{-\mathrm{Col}} \\ &= \operatorname{span}\left\{\nabla^{-\mathrm{Col}}_{1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3'},\ \nabla^{-\mathrm{Col}}_{13\backslash\backslash 2,\ 13\backslash\backslash 2'},\ \nabla^{-\mathrm{Col}}_{13\backslash\backslash 2,\ 12\backslash\backslash 3'},\ \nabla^{-\mathrm{Col}}_{12\backslash\backslash 3,\ 13\backslash\backslash 2'},\right. \\ &\qquad\qquad \left.\nabla^{-\mathrm{Col}}_{12\backslash\backslash 3,\ 12\backslash\backslash 3}\right\};\end{aligned}$$

$$\begin{aligned}\mathcal{F}_4^{-\mathrm{Col}} &= \mathcal{F}_{\mathrm{std},\leq T_4}^{-\mathrm{Col}} = \mathcal{A} \\ &= \operatorname{span}\left\{\nabla^{-\mathrm{Col}}_{1\backslash\backslash 2\backslash\backslash 3,\ 1\backslash\backslash 2\backslash\backslash 3'},\ \nabla^{-\mathrm{Col}}_{13\backslash\backslash 2,\ 13\backslash\backslash 2'},\ \nabla^{-\mathrm{Col}}_{13\backslash\backslash 2,\ 12\backslash\backslash 3'},\ \nabla^{-\mathrm{Col}}_{12\backslash\backslash 3,\ 13\backslash\backslash 2'},\right. \\ &\qquad\qquad \left.\nabla^{-\mathrm{Col}}_{12\backslash\backslash 3,\ 12\backslash\backslash 3'},\ \nabla^{-\mathrm{Col}}_{123,\ 123}\right\}.\end{aligned}$$

Anticipating Theorem 6.8.41, we observe that the subquotients of the latter filtration are isomorphic to Specht modules:

$$\mathcal{F}_4^{-\mathrm{Col}}/\mathcal{F}_3^{-\mathrm{Col}} \cong \mathcal{S}^{(3)};$$

$$\mathcal{F}_3^{-\mathrm{Col}}/\mathcal{F}_2^{-\mathrm{Col}} \cong \mathcal{S}^{(2,1)};$$

$$\mathcal{F}_2^{-\mathrm{Col}}/\mathcal{F}_1^{-\mathrm{Col}} \cong \mathcal{S}^{(2,1)};$$

$$\mathcal{F}_1^{-\mathrm{Col}}/\mathcal{F}_0^{-\mathrm{Col}} \cong \mathcal{S}^{(1,1,1)}.$$

*Proof of Theorem 6.8.37.* This is all just a matter of combining previously proved facts (Proposition 6.8.30, Theorem 6.8.29 and Proposition 6.8.32). See the appendix (Section A.6) for a detailed proof. □

As we said, the filtration of $\mathcal{A}$ constructed in Theorem 6.8.37 **(e)** is a filtration by left $\mathcal{A}$-submodules; i.e., each part $\mathcal{F}_i^{\mathrm{Row}}$ of this filtration is a left $\mathcal{A}$-submodule of $\mathcal{A}$. Thus, left multiplication by any given $\mathbf{a} \in \mathcal{A}$ preserves all the parts $\mathcal{F}_i^{\mathrm{Row}}$ of this filtration, and therefore is represented by a block-

upper-triangular matrix on the Murphy basis $\left(\nabla^{\mathrm{Row}}_{V,U}\right)_{(\lambda,U,V)\in \mathrm{SBT}(n)}$. The diagonal blocks have sizes $f^\lambda$ for the various partitions $\lambda$ of $n$, but each $f^\lambda$ appears $f^\lambda$ many times.

The filtration in Theorem 6.8.37 **(f)** also has the same block-triangularity property. But it also has the additional nice property that its subquotients are the Specht modules $\mathcal{S}^\lambda$ corresponding to the partitions $\lambda$ satisfying $T \in \mathrm{SYT}(\lambda)$. This follows from the following property of row-Murphy sums:

**Theorem 6.8.39.** Let $\lambda$ be a partition of $n$. Let $T$ be any standard $n$-tableau of shape $\lambda$. Then,

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} \cong \mathcal{A}\mathbf{E}_T \cong \mathcal{S}^\lambda \qquad \text{as } S_n\text{-representations.}$$

More concretely, the map

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} \to \mathcal{A}\mathbf{E}_T,$$
$$\overline{\mathbf{a}} \mapsto \mathbf{a}\nabla_{\mathrm{Row}\,T}$$

(where $\overline{\mathbf{a}}$ denotes the projection of a vector $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ on the quotient $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$) is a left $\mathcal{A}$-module isomorphism.

To prove this theorem, we need a lemma:

**Lemma 6.8.40.** Let $\lambda$ be a partition of $n$. Let $T$ be any standard $n$-tableau of shape $\lambda$. Then:

**(a)** For any $(\mu, U, V) \in \mathrm{SBT}(n)$ satisfying $U < T$, we have $\nabla^{-\mathrm{Col}}_{V,U}\nabla_{\mathrm{Row}\,T} = 0$.

**(b)** For any $n$-tableau $V$ of shape $\lambda$, we have $\nabla^{-\mathrm{Col}}_{V,T}\nabla_{\mathrm{Row}\,T} = \pm\mathbf{E}_{V,T}$. (See Definition 5.15.4 for the meaning of $\mathbf{E}_{V,T}$.)

**(c)** Each $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ satisfies $\mathbf{a}\nabla_{\mathrm{Row}\,T} = 0$.

**(d)** Each $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ satisfies $\mathbf{a}\nabla_{\mathrm{Row}\,T} \in \mathcal{A}\mathbf{E}_T$.

**(e)** We have $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T} = \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} + \mathrm{span}\left\{\nabla^{-\mathrm{Col}}_{V,T} \mid V \in \mathrm{SYT}(\lambda)\right\}$.

*Proof.* **(a)** Let $(\mu, U, V) \in \mathrm{SBT}(n)$ be such that $U < T$. Then, $U, V \in \mathrm{SYT}(\mu)$; thus, $V$ and $U$ are $n$-tableaux of the same shape (namely, of shape $Y(\mu)$). Hence, Definition 5.15.2 defines a permutation $w_{V,U} \in S_n$ that satisfies $w_{V,U}U = V$. Consider this $w_{V,U}$.

We have $U \in \mathrm{SYT}(\mu)$ and $T \in \mathrm{SYT}(\lambda)$ (since $T$ is a standard $n$-tableau of

shape $\lambda$). Thus, $U, T \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$. Moreover, $U < T$. Thus, Lemma 6.7.9 (applied to $P = U$ and $Q = T$) yields $\nabla^{-}_{\operatorname{Col} U} \nabla_{\operatorname{Row} T} = 0$ and $\nabla_{\operatorname{Row} T} \nabla^{-}_{\operatorname{Col} U} = 0$.

Set $u := w_{V,U}$; thus, $uU = w_{V,U}U = V$, so that $V = uU$. Hence, (499) (applied to $P = V$ and $Q = U$) yields $\nabla^{-\operatorname{Col}}_{V,U} = (-1)^u \nabla^{-}_{\operatorname{Col} V} u = (-1)^u u \nabla^{-}_{\operatorname{Col} U}$. Therefore,

$$\underbrace{\nabla^{-\operatorname{Col}}_{V,U}}_{=(-1)^u u \nabla^{-}_{\operatorname{Col} U}} \nabla_{\operatorname{Row} T} = (-1)^u u \underbrace{\nabla^{-}_{\operatorname{Col} U} \nabla_{\operatorname{Row} T}}_{=0} = 0.$$

This proves Lemma 6.8.40 **(a)**.

**(b)** Let $V$ be any $n$-tableau of shape $\lambda$. Then, $V$ and $T$ are $n$-tableaux of the same shape (namely, of shape $Y(\lambda)$). Hence, Definition 5.15.2 defines a permutation $w_{V,T} \in S_n$ that satisfies $w_{V,T} T = V$. Consider this $w_{V,T}$.

Set $u := w_{V,T}$; thus, $uT = w_{V,T}T = V$, so that $V = uT$. Hence, (499) (applied to $P = V$ and $Q = T$) yields $\nabla^{-\operatorname{Col}}_{V,T} = (-1)^u \nabla^{-}_{\operatorname{Col} V} u = (-1)^u u \nabla^{-}_{\operatorname{Col} T}$. Therefore,

$$\begin{aligned}
\underbrace{\nabla^{-\operatorname{Col}}_{V,T}}_{=(-1)^u u \nabla^{-}_{\operatorname{Col} T}} \nabla_{\operatorname{Row} T} &= (-1)^u \underbrace{u}_{=w_{V,T}} \underbrace{\nabla^{-}_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}}_{=\mathbf{E}_T \text{ (by the definition of } \mathbf{E}_T)} \\
&= (-1)^u \underbrace{w_{V,T} \mathbf{E}_T}_{=\mathbf{E}_{V,T} \text{ (since (270) yields } \mathbf{E}_{V,T} = w_{V,T}\mathbf{E}_T)} \\
&= \underbrace{(-1)^u}_{=\pm 1} \mathbf{E}_{V,T} = \pm \mathbf{E}_{V,T}.
\end{aligned} \tag{533}$$

This proves Lemma 6.8.40 **(b)**.

**(c)** The definition of $\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}$ says that

$$\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T} = \operatorname{span}\left\{ \nabla^{-\operatorname{Col}}_{V,U} \mid (\mu, U, V) \in \operatorname{SBT}(n) \text{ and } U < T \right\}$$

(note that we used the notation "$(\lambda, U, V)$" instead of "$(\mu, U, V)$" in the original Definition 6.8.26, but we have now renamed it as "$(\mu, U, V)$" because the letter $\lambda$ is already taken).

Now, let $\mathbf{a} \in \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}$. We must show that $\mathbf{a} \nabla_{\operatorname{Row} T} = 0$. This claim depends $\mathbf{k}$-linearly on $\mathbf{a}$. Thus, by linearity, we can WLOG assume that $\mathbf{a}$ is a vector of the form $\mathbf{a} = \nabla^{-\operatorname{Col}}_{V,U}$ for some $(\mu, U, V) \in \operatorname{SBT}(n)$ satisfying $U < T$ (because $\mathbf{a} \in \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T} = \operatorname{span}\left\{ \nabla^{-\operatorname{Col}}_{V,U} \mid (\mu, U, V) \in \operatorname{SBT}(n) \text{ and } U < T \right\}$ shows that $\mathbf{a}$ is always a $\mathbf{k}$-linear combination of vectors of this form). Assume this, and consider this $(\mu, U, V)$.

Now, from $\mathbf{a} = \nabla^{-\operatorname{Col}}_{V,U}$, we obtain $\mathbf{a} \nabla_{\operatorname{Row} T} = \nabla^{-\operatorname{Col}}_{V,U} \nabla_{\operatorname{Row} T} = 0$ (by Lemma 6.8.40 **(a)**). Thus, we have proved Lemma 6.8.40 **(c)**.

**(d)** The definition of $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ says that

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T} = \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{V,U} \;\mid\; (\mu, U, V) \in \mathrm{SBT}(n) \text{ and } U \leq T \right\}$$

(note that we used the notation "$(\lambda, U, V)$" instead of "$(\mu, U, V)$" in the original Definition 6.8.26, but we have now renamed it as "$(\mu, U, V)$" because the letter $\lambda$ is already taken).

Now, let $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$. We must show that $\mathbf{a}\nabla_{\mathrm{Row}\,T} \in \mathcal{A}\mathbf{E}_T$. This claim depends $\mathbf{k}$-linearly on $\mathbf{a}$ (since $\mathcal{A}\mathbf{E}_T$ is a $\mathbf{k}$-submodule of $\mathcal{A}$). Thus, by linearity, we can WLOG assume that $\mathbf{a}$ is a vector of the form $\mathbf{a} = \nabla^{-\mathrm{Col}}_{V,U}$ for some $(\mu, U, V) \in \mathrm{SBT}(n)$ satisfying $U \leq T$ (because $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T} = \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu, U, V) \in \mathrm{SBT}(n) \text{ and } U \leq T \right\}$ shows that $\mathbf{a}$ is always a $\mathbf{k}$-linear combination of vectors of this form). Assume this, and consider this $(\mu, U, V)$.

If $U < T$, then

$$\underbrace{\mathbf{a}}_{=\nabla^{-\mathrm{Col}}_{V,U}} \nabla_{\mathrm{Row}\,T} = \nabla^{-\mathrm{Col}}_{V,U}\nabla_{\mathrm{Row}\,T} = 0 \qquad \text{(by Lemma 6.8.40 \textbf{(a)})}$$
$$\in \mathcal{A}\mathbf{E}_T \qquad (\text{since } \mathcal{A}\mathbf{E}_T \text{ is a } \mathbf{k}\text{-module}).$$

Hence, $\mathbf{a}\nabla_{\mathrm{Row}\,T} \in \mathcal{A}\mathbf{E}_T$ is proved in the case when $U < T$. Thus, for the rest of this proof, we WLOG assume that we don't have $U < T$. Thus, we must have $U = T$ (since $U \leq T$). Thus, $U = T \in \mathrm{SYT}(\lambda)$ (since $T$ is a standard $n$-tableau of shape $\lambda$). But $U \in \mathrm{SYT}(\mu)$ (since $(\mu, U, V) \in \mathrm{SBT}(n)$). Hence, the sets $\mathrm{SYT}(\lambda)$ and $\mathrm{SYT}(\mu)$ have at least one element in common (namely, $U$). Thus, $\lambda = \mu$ (since the sets $\mathrm{SYT}(\kappa)$ for different partitions $\kappa$ are disjoint). Thus, $V$ is an $n$-tableau of shape $\lambda$ (since $V$ is an $n$-tableau of shape $\mu$ (because $(\mu, U, V) \in \mathrm{SBT}(n)$)).

Therefore, everything we wrote in our above proof of Lemma 6.8.40 **(b)** applies. In particular, (533) holds, where $u \in S_n$ is defined as in the proof just mentioned. Thus,

$$\nabla^{-\mathrm{Col}}_{V,T}\nabla_{\mathrm{Row}\,T} = \underbrace{(-1)^u w_{V,T}}_{\in\mathcal{A}} \mathbf{E}_T \in \mathcal{A}\mathbf{E}_T.$$

In other words, $\mathbf{a}\nabla_{\mathrm{Row}\,T} \in \mathcal{A}\mathbf{E}_T$ (since $\mathbf{a} = \nabla^{-\mathrm{Col}}_{V,U}$). This completes the proof of Lemma 6.8.40 **(d)**.

**(e)** This follows from the definitions of $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ and $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ – and from the fact that the only vectors $\nabla^{-\mathrm{Col}}_{V,U}$ in the definition of $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ that do not also appear in the definition of $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ are precisely the ones that satisfy $U = T$, that is, the ones of the form $\nabla^{-\mathrm{Col}}_{V,T}$ with $V \in \mathrm{SYT}(\lambda)$. See the appendix (Section A.6) for details. □

*Proof of Theorem 6.8.39.* We know that $T$ is an $n$-tableau of shape $Y(\lambda)$. Thus, $\mathcal{S}^{Y(\lambda)} \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules (by (228), applied to $D = Y(\lambda)$). Hence, $\mathcal{A}\mathbf{E}_T \cong \mathcal{S}^{Y(\lambda)} = \mathcal{S}^\lambda$ (since $\mathcal{S}^\lambda$ is defined as $\mathcal{S}^{Y(\lambda)}$).

Theorem 6.8.29 **(c)** shows that $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$. Theorem 6.8.29 **(d)** shows that $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$. Proposition 6.8.32 yields $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} \subseteq \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$. Hence, the quotient $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ is well-defined, and is a left $\mathcal{A}$-module, i.e., an $S_n$-representation.

Consider the $\mathbf{k}$-linear map

$$\widehat{\Phi} : \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T} \to \mathcal{A}\mathbf{E}_T,$$
$$\mathbf{a} \mapsto \mathbf{a}\nabla_{\mathrm{Row}\,T}$$

(this is well-defined, since Lemma 6.8.40 **(d)** shows that each $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ satisfies $\mathbf{a}\nabla_{\mathrm{Row}\,T} \in \mathcal{A}\mathbf{E}_T$). This map $\widehat{\Phi}$ is easily seen to be left $\mathcal{A}$-linear (because for any $\mathbf{b} \in \mathcal{A}$ and any $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$, we have $\Phi(\mathbf{ba}) = \mathbf{b}\underbrace{\mathbf{a}\nabla_{\mathrm{Row}\,T}}_{=\Phi(\mathbf{a})} = \mathbf{b}\Phi(\mathbf{a})$). In other words, $\widehat{\Phi}$ is a left $\mathcal{A}$-module morphism. Moreover, this morphism $\widehat{\Phi}$ sends the left $\mathcal{A}$-submodule $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ of $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ to 0 (because if $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$, then $\widehat{\Phi}(\mathbf{a}) = \mathbf{a}\nabla_{\mathrm{Row}\,T} = 0$ by Lemma 6.8.40 **(c)**). Hence, the universal property of quotient modules shows that this morphism $\widehat{\Phi}$ factors through the quotient $\mathcal{A}$-module $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$. In other words, $\widehat{\Phi} = \Phi \circ \pi$, where $\pi$ is the canonical projection from $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ to $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$, and where $\Phi : \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} \to \mathcal{A}\mathbf{E}_T$ is a left $\mathcal{A}$-module morphism that is explicitly given as follows:

$$\Phi : \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} \to \mathcal{A}\mathbf{E}_T,$$
$$\overline{\mathbf{a}} \mapsto \mathbf{a}\nabla_{\mathrm{Row}\,T}. \tag{534}$$

Consider this latter morphism $\Phi$. We shall now prove the following:

*Claim 1:* The map $\Phi$ is surjective.

*Claim 2:* The map $\Phi$ is injective.

*Proof of Claim 1.* Let $\mathbf{v} \in \mathcal{A}\mathbf{E}_T$. Thus, $\mathbf{v} = \mathbf{b}\mathbf{E}_T$ for some $\mathbf{b} \in \mathcal{A}$. Consider this $\mathbf{b}$. We shall show that $\mathbf{v}$ lies in the image of $\Phi$.

We have $\nabla^-_{\mathrm{Col}\,T} = \nabla^{-\mathrm{Col}}_{T,T}$ (by (493)). But $\nabla^{-\mathrm{Col}}_{T,T} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ (this follows from the definition of $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ pretty easily[180]). Therefore, $\nabla^-_{\mathrm{Col}\,T} = \nabla^{-\mathrm{Col}}_{T,T} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$,

[180] *Proof.* The definition of $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ says that

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T} = \mathrm{span}\left\{\nabla^{-\mathrm{Col}}_{V,U} \mid (\mu, U, V) \in \mathrm{SBT}(n) \text{ and } U \leq T\right\}$$

so that $\overline{\nabla^{-}_{\operatorname{Col} T}} \in \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T} / \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}$. Thus, $\Phi\left(\overline{\nabla^{-}_{\operatorname{Col} T}}\right)$ is well-defined. The explicit formula (534) for the map $\Phi$ shows that $\Phi\left(\overline{\nabla^{-}_{\operatorname{Col} T}}\right) = \nabla^{-}_{\operatorname{Col} T} \nabla_{\operatorname{Row} T} = \mathbf{E}_T$ (since $\mathbf{E}_T$ is defined as $\nabla^{-}_{\operatorname{Col} T} \nabla_{\operatorname{Row} T}$). Now, since $\Phi$ is a left $\mathcal{A}$-module morphism, we have

$$\Phi\left(\mathbf{b}\overline{\nabla^{-}_{\operatorname{Col} T}}\right) = \mathbf{b} \underbrace{\Phi\left(\overline{\nabla^{-}_{\operatorname{Col} T}}\right)}_{=\mathbf{E}_T} = \mathbf{b}\mathbf{E}_T = \mathbf{v} \qquad (\text{since } \mathbf{v} = \mathbf{b}\mathbf{E}_T),$$

and thus $\mathbf{v} = \Phi\left(\mathbf{b}\overline{\nabla^{-}_{\operatorname{Col} T}}\right)$. This shows that $\mathbf{v}$ lies in the image of $\Phi$.

Forget that we fixed $\mathbf{v}$. We thus have shown that each $\mathbf{v} \in \mathcal{A}\mathbf{E}_T$ lies in the image of $\Phi$. In other words, $\Phi$ is surjective. This proves Claim 1. □

*Proof of Claim 2.* We must prove that $\Phi$ is injective. Since $\Phi$ is $\mathbf{k}$-linear, it suffices to show that $\operatorname{Ker}\Phi = 0$ (because any $\mathbf{k}$-linear map whose kernel is 0 is automatically injective).

So let $\mathbf{v} \in \operatorname{Ker}\Phi$. We must show that $\mathbf{v} = 0$.

Clearly, $Y(\lambda)$ is a skew Young diagram, and we have $\operatorname{SYT}(Y(\lambda)) = \operatorname{SYT}(\lambda)$ (since tableaux of shape $Y(\lambda)$ are the same thing as tableaux of shape $\lambda$). Thus, Proposition 5.15.9 (applied to $D = Y(\lambda)$) shows that the family $\left(\mathbf{E}_{P,T}\right)_{P \in \operatorname{SYT}(Y(\lambda))}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_T$. Since $\operatorname{SYT}(Y(\lambda)) = \operatorname{SYT}(\lambda)$, we can rewrite this as follows: The family $\left(\mathbf{E}_{P,T}\right)_{P \in \operatorname{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_T$. In other words, the family $\left(\mathbf{E}_{V,T}\right)_{V \in \operatorname{SYT}(\lambda)}$ is a basis of the $\mathbf{k}$-module $\mathcal{A}\mathbf{E}_T$ (because this is the same family as $\left(\mathbf{E}_{P,T}\right)_{P \in \operatorname{SYT}(\lambda)}$; we just renamed the index $P$ as $V$). Hence, this family is $\mathbf{k}$-linearly independent.

We have $\mathbf{v} \in \operatorname{Ker}\Phi \subseteq \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T} / \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}$. Thus, $\mathbf{v} = \overline{\mathbf{c}}$ for some $\mathbf{c} \in \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T}$. Consider this $\mathbf{c}$. Then,

$$\mathbf{c} \in \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T} = \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T} + \operatorname{span}\left\{\nabla^{-\operatorname{Col}}_{V,T} \mid V \in \operatorname{SYT}(\lambda)\right\}$$

(by Lemma 6.8.40 **(e)**). In other words, $\mathbf{c}$ can be written as $\mathbf{c} = \mathbf{a} + \mathbf{b}$ for some $\mathbf{a} \in \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}$ and some $\mathbf{b} \in \operatorname{span}\left\{\nabla^{-\operatorname{Col}}_{V,T} \mid V \in \operatorname{SYT}(\lambda)\right\}$. Consider these $\mathbf{a}$ and

(note that we used the notation "$(\lambda, U, V)$" instead of "$(\mu, U, V)$" in the original Definition 6.8.26, but we have now renamed it as "$(\mu, U, V)$" because the letter $\lambda$ is already taken). But $(\lambda, T, T) \in \operatorname{SBT}(n)$ (since $T, T \in \operatorname{SYT}(\lambda)$). Hence, $(\lambda, T, T)$ is a $(\mu, U, V) \in \operatorname{SBT}(n)$ satisfying $U \leq T$ (since $T \leq T$). Therefore, the element $\nabla^{-\operatorname{Col}}_{T,T}$ has the form $\nabla^{-\operatorname{Col}}_{V,U}$ for some $(\mu, U, V) \in \operatorname{SBT}(n)$ satisfying $U \leq T$ (namely, for $(\mu, U, V) = (\lambda, T, T)$). That is,

$$\begin{aligned}\nabla^{-\operatorname{Col}}_{T,T} &\in \left\{\nabla^{-\operatorname{Col}}_{V,U} \mid (\mu, U, V) \in \operatorname{SBT}(n) \text{ and } U \leq T\right\} \\ &\subseteq \operatorname{span}\left\{\nabla^{-\operatorname{Col}}_{V,U} \mid (\mu, U, V) \in \operatorname{SBT}(n) \text{ and } U \leq T\right\} = \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T}.\end{aligned}$$

$\mathbf{b}$. Then, $\mathbf{c} = \mathbf{a} + \mathbf{b}$, so that $\mathbf{c} - \mathbf{b} = \mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$. Therefore, $\overline{\mathbf{c}} = \overline{\mathbf{b}}$ (because two vectors in $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}$ whose difference belongs to $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ must have the same projection onto the quotient $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$). Thus, $\mathbf{v} = \overline{\mathbf{c}} = \overline{\mathbf{b}}$, so that $\overline{\mathbf{b}} = \mathbf{v}$.

But $\mathbf{b} \in \operatorname{span}\left\{\nabla^{-\mathrm{Col}}_{V,T} \mid V \in \operatorname{SYT}(\lambda)\right\}$. Hence, we can write $\mathbf{b}$ as a $\mathbf{k}$-linear combination

$$\mathbf{b} = \sum_{V \in \operatorname{SYT}(\lambda)} \omega_V \nabla^{-\mathrm{Col}}_{V,T} \tag{535}$$

for some scalars $\omega_V \in \mathbf{k}$. Consider these $\omega_V$. But $\overline{\mathbf{b}} = \mathbf{v} \in \operatorname{Ker}\Phi$, so that $\Phi\left(\overline{\mathbf{b}}\right) = 0$. Hence,

$$\begin{aligned}
0 &= \Phi\left(\overline{\mathbf{b}}\right) = \mathbf{b}\nabla_{\mathrm{Row}\,T} \qquad \text{(by the explicit formula (534) for the map } \Phi\text{)} \\
&= \left(\sum_{V \in \operatorname{SYT}(\lambda)} \omega_V \nabla^{-\mathrm{Col}}_{V,T}\right) \nabla_{\mathrm{Row}\,T} \qquad \text{(by (535))} \\
&= \sum_{V \in \operatorname{SYT}(\lambda)} \omega_V \underbrace{\nabla^{-\mathrm{Col}}_{V,T}\nabla_{\mathrm{Row}\,T}}_{=\pm\mathbf{E}_{V,T}} = \sum_{V \in \operatorname{SYT}(\lambda)} \underbrace{\omega_V\left(\pm\mathbf{E}_{V,T}\right)}_{=(\pm\omega_V)\mathbf{E}_{V,T}} \\
&\qquad \text{(by Lemma 6.8.40 \textbf{(b)})} \\
&= \sum_{V \in \operatorname{SYT}(\lambda)} \left(\pm\omega_V\right)\mathbf{E}_{V,T}.
\end{aligned}$$

In other words,

$$\sum_{V \in \operatorname{SYT}(\lambda)} \left(\pm\omega_V\right)\mathbf{E}_{V,T} = 0 \qquad \text{(for an appropriate choice of } \pm \text{ signs)}.$$

Since the family $\left(\mathbf{E}_{V,T}\right)_{V \in \operatorname{SYT}(\lambda)}$ is $\mathbf{k}$-linearly independent, we thus conclude that all the coefficients $\pm\omega_V$ in this sum are 0. In other words, each $V \in \operatorname{SYT}(\lambda)$ satisfies $\pm\omega_V = 0$ and therefore $\omega_V = 0$. Thus, (535) simplifies to

$$\mathbf{b} = \sum_{V \in \operatorname{SYT}(\lambda)} \underbrace{\omega_V}_{=0} \nabla^{-\mathrm{Col}}_{V,T} = \sum_{V \in \operatorname{SYT}(\lambda)} 0\nabla^{-\mathrm{Col}}_{V,T} = 0.$$

Now recall that $\mathbf{v} = \overline{\mathbf{b}}$. In view of $\mathbf{b} = 0$, we can rewrite this as $\mathbf{v} = \overline{0} = 0$.

Thus we have shown that $\mathbf{v} = 0$ for each $\mathbf{v} \in \operatorname{Ker}\Phi$. In other words, $\operatorname{Ker}\Phi = 0$. As we said, this completes the proof of Claim 2. □

Now, the map $\Phi$ is injective (by Claim 2) and surjective (by Claim 1). Hence, $\Phi$ is bijective, i.e., invertible. Thus, $\Phi$ is a left $\mathcal{A}$-module isomorphism (since $\Phi$ is a left $\mathcal{A}$-module morphism). In other words, the map

$$\begin{aligned}
\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} &\to \mathcal{A}\mathbf{E}_T, \\
\overline{\mathbf{a}} &\mapsto \mathbf{a}\nabla_{\mathrm{Row}\,T}
\end{aligned}$$

is a left $\mathcal{A}$-module isomorphism (since this map is precisely $\Phi$, because of (534)). Therefore, $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} \cong \mathcal{A}\mathbf{E}_T$ as left $\mathcal{A}$-modules. Combining this with $\mathcal{A}\mathbf{E}_T \cong \mathcal{S}^\lambda$, we obtain $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T}/\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} \cong \mathcal{A}\mathbf{E}_T \cong \mathcal{S}^\lambda$ as left $\mathcal{A}$-modules, i.e., as $S_n$-representations. Thus, Theorem 6.8.39 is proved. □

The column-Murphy sums have a property analogous to Theorem 6.8.39, which we leave as an exercise:

**Exercise 6.8.7.** [3] Let $\lambda$ be a partition of $n$. Let $T$ be any standard $n$-tableau of shape $\lambda$. Then, prove that

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}/\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T} \cong \mathcal{A}\mathbf{F}_T \cong \left(\mathcal{S}^\lambda\right)^* \qquad \text{as } S_n\text{-representations}$$

(where $\mathbf{F}_T$ is as defined in Proposition 5.11.19, and where $\left(\mathcal{S}^\lambda\right)^*$ is the dual of the Specht module $\mathcal{S}^\lambda$). More concretely, prove that the map

$$\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}/\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T} \to \mathcal{A}\mathbf{F}_T,$$
$$\overline{\mathbf{a}} \mapsto \mathbf{a}\nabla^-_{\mathrm{Col}\,T}$$

(where $\overline{\mathbf{a}}$ denotes the projection of a vector $\mathbf{a} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}$ on the quotient $\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}/\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T}$) is a left $\mathcal{A}$-module isomorphism.

[**Hint:** From Lemma 5.18.15 and Theorem 5.19.34, we have $\mathcal{A}\mathbf{F}_T \cong \left(\mathcal{S}^{\mathbf{r}(D)}\right)^{\mathrm{sign}} \cong \left(\mathcal{S}^D\right)^*$ for $D = Y(\lambda)$. Also recall (280).]

Theorem 6.8.39 was stated in terms of a single tableau $T$, but we can easily restate it in terms of the entire filtration:

**Theorem 6.8.41.** Let $T_1, T_2, \ldots, T_m$ be all the standard tableaux in the set $\bigcup_{\kappa \vdash n} \mathrm{SYT}(\kappa)$, listed in increasing order (with respect to the total order defined in Definition 6.7.7). Recall from Theorem 6.8.37 **(f)** that $\mathcal{A}$ has a filtration

$$\left(\mathcal{F}_0^{-\mathrm{Col}}, \mathcal{F}_1^{-\mathrm{Col}}, \mathcal{F}_2^{-\mathrm{Col}}, \ldots, \mathcal{F}_{m-1}^{-\mathrm{Col}}, \mathcal{F}_m^{-\mathrm{Col}}\right) \tag{536}$$

by left $\mathcal{A}$-submodules.

Let $i \in [m]$. Let $\lambda^{(i)}$ be the partition of $n$ that satisfies $T_i \in \mathrm{SYT}\left(\lambda^{(i)}\right)$. Then, the subquotient $\mathcal{F}_i^{-\mathrm{Col}}/\mathcal{F}_{i-1}^{-\mathrm{Col}}$ of the filtration (536) is isomorphic (as a left $\mathcal{A}$-module) to the Specht module $\mathcal{S}^{\lambda^{(i)}}$.

*Proof.* Here is a sketch (see Section A.6 for more details). Set $T := T_i$ and $\lambda := \lambda^{(i)}$. Thus, $T \in \operatorname{SYT}(\lambda)$. Hence, Theorem 6.8.39 yields that $\mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T} / \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T} \cong \mathcal{S}^\lambda$ as $S_n$-representations, i.e., as left $\mathcal{A}$-modules.

But $\mathcal{F}_i^{-\operatorname{Col}} = \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T}$ (by the definition of $\mathcal{F}_i^{-\operatorname{Col}}$) and $\mathcal{F}_{i-1}^{-\operatorname{Col}} = \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T}$ (by Theorem 6.8.37 **(b)**, applied to $i-1$ instead of $i$). Hence,

$$\mathcal{F}_i^{-\operatorname{Col}} / \mathcal{F}_{i-1}^{-\operatorname{Col}} = \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},\leq T} / \mathcal{F}^{-\operatorname{Col}}_{\operatorname{std},<T} \cong \mathcal{S}^\lambda = \mathcal{S}^{\lambda^{(i)}} \qquad \left(\text{since } \lambda = \lambda^{(i)}\right)$$

as left $\mathcal{A}$-modules. This proves Theorem 6.8.41. □

Theorem 6.8.41 shows that the filtration (536) is a so-called *Specht filtration* – that is, a filtration whose each subquotient is isomorphic to a Specht module of partition shape (i.e., to a Specht module $\mathcal{S}^\lambda$ with $\lambda \vdash n$). When $\mathbf{k}$ is a field of characteristic 0, these subquotients are therefore irreducible as left $\mathcal{A}$-modules (since Corollary 5.13.3 shows that the Specht modules $\mathcal{S}^\lambda$ are irreducible), so that the filtration is a composition series. Thus we have proved the following:

**Corollary 6.8.42.** Assume that $\mathbf{k}$ is a field of characteristic 0. Then, the filtration (536) is a composition series.

**Exercise 6.8.8.** [1] Prove an analogous property of the filtration $\left(\mathcal{F}_m^{\operatorname{Row}}, \mathcal{F}_{m-1}^{\operatorname{Row}}, \ldots, \mathcal{F}_2^{\operatorname{Row}}, \mathcal{F}_1^{\operatorname{Row}}, \mathcal{F}_0^{\operatorname{Row}}\right)$.

Corollary 6.8.42 gives us a new proof of Corollary 5.13.8:

*Another proof of Corollary 5.13.8.* As we just said (Corollary 6.8.42), the filtration (536) of $\mathcal{A}$ is a composition series. Hence, Theorem 4.4.24 (applied to the ring $\mathcal{A}$, the number $m$ and the composition series $\left(\mathcal{F}_0^{-\operatorname{Col}}, \mathcal{F}_1^{-\operatorname{Col}}, \ldots, \mathcal{F}_m^{-\operatorname{Col}}\right)$ instead of the ring $R$, the number $k$, and the composition series $(W_0, W_1, \ldots, W_k)$) shows that each irreducible left $\mathcal{A}$-module is isomorphic to $\mathcal{F}_i^{-\operatorname{Col}} / \mathcal{F}_{i-1}^{-\operatorname{Col}}$ for some $i \in [m]$. But Theorem 6.8.41 shows that each such $\mathcal{F}_i^{-\operatorname{Col}} / \mathcal{F}_{i-1}^{-\operatorname{Col}}$ is isomorphic to $\mathcal{S}^\lambda$ for some partition $\lambda$ of $n$. Combining these two isomorphisms, we conclude that each irreducible left $\mathcal{A}$-module is isomorphic to $\mathcal{S}^\lambda$ for some partition $\lambda$ of $n$. This proves Corollary 5.13.8 again. □

**Remark 6.8.43.** The filtration (536) is not the only one that has the nice properties we proved (such as Theorem 6.8.41 and Corollary 6.8.42). One easy way to construct other such filtrations is by replacing our total order on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ by another that still satisfies Lemma 6.7.9. Such alternative orders are not unusual, although not every total order would work.

# A. Omitted details and proofs

This chapter contains some proofs (and parts of proofs) that have been omitted from the text above – usually because they are technical arguments or of tangential interest only.

Each of the proofs given below uses the notations and conventions of the chapter and section in which the respective claim appears.

## A.1. Tableaux

Here we supply detailed proofs for some claims made in Section 5.2.

*Proof of Proposition 5.2.7.* **(a)** Assume that every $(i,j) \in D$ satisfying $(i,j+1) \in D$ satisfies

$$T(i,j) < T(i,j+1). \tag{537}$$

We must show that $T$ is row-standard.

In other words, we must show that the entries of $T$ strictly increase (left to right) along each row. In other words, we must show that if $(i,j)$ and $(i,k)$ are two cells in the same row of $D$, with $(i,j)$ lying west of $(i,k)$ (that is, $j<k$), then $T(i,j) < T(i,k)$.

So let us show this. We fix the cell $(i,j) \in D$; then we must prove the following claim:

*Claim 1:* If $(i,k)$ is a cell of $D$ satisfying $j<k$, then $T(i,j) < T(i,k)$.

*Proof of Claim 1.* We shall prove Claim 1 by induction on $k$.

The *base case* is the case $k = j+1$ (since the assumption $j<k$ renders any smaller value of $k$ impossible). In this case, Claim 1 is saying that if $(i,j+1)$ is a cell of $D$, then $T(i,j) < T(i,j+1)$. But this is precisely our assumption (537). Thus, Claim 1 holds for $k = j+1$.

Now we come to the *induction step:* Fix some integer $p \geq j+1$. Assume (as the induction hypothesis) that Claim 1 holds for $k=p$. We must prove that Claim 1 holds for $k = p+1$. In other words, we must prove that if $(i,p+1)$ is a cell of $D$, then $T(i,j) < T(i,p+1)$.

So let us assume that $(i,p+1)$ is a cell of $D$. Thus, $(i,p+1) \in D$.

But $D$ is a skew Young diagram; that is, we have $D = Y(\lambda/\mu)$ for some skew partition $\lambda/\mu$. Consider this $\lambda/\mu$. We have $(i,j) \in D = Y(\lambda/\mu)$ and $(i,p+1) \in D = Y(\lambda/\mu)$. Furthermore, $p \geq j+1 > j$, so that $j<p$ and thus $j \leq p$, so that $(i,j) \leq (i,p)$ (since $i \leq i$ and $j \leq p$). Moreover, $(i,p) \leq (i,p+1)$ (since $i \leq i$ and $p \leq p+1$). Therefore, Proposition 5.1.6 (applied to $c=(i,j)$ and $d=(i,p)$ and $e=(i,p+1)$) yields $(i,p) \in Y(\lambda/\mu)$. In other words, $(i,p) \in D$ (since $D = Y(\lambda/\mu)$). Hence, (537) (applied to $p$ instead of $j$) yields $T(i,p) < T(i,p+1)$ (since $(i,p+1) \in D$).

But we assumed that Claim 1 holds for $k=p$. In other words, if $(i,p)$ is a cell of $D$ satisfying $j<p$, then $T(i,j) < T(i,p)$. Hence, we conclude that $T(i,j) < T(i,p)$ (since $(i,p) \in D$ and $j<p$). Altogether, we now have $T(i,j) < T(i,p) < T(i,p+1)$.

Thus, we have shown that $T(i,j) < T(i,p+1)$ under the assumption that $(i,p+1)$ is a cell of $D$. In other words, Claim 1 holds for $k = p+1$. This completes the induction step, so Claim 1 is proved. $\square$

As we explained, this completes the proof of Proposition 5.2.7 **(a)**.

**(b)** This is analogous to part **(a)**, which we have just proved. The only necessary change is to interchange the horizontal and vertical directions (i.e., rows become columns, x-coordinates become y-coordinates, etc.). □

## A.2. The Young natural basis of $\mathbf{k}[S_n]$

Here we supply detailed proofs for some claims made in Section 5.15.

*Proof of Proposition 5.15.12.* **(a)** Let $T$ be a standard $n$-tableau of shape $D$. Thus, $T$ is standard, i.e., both row-standard and column-standard. In particular, $T$ is row-standard. In other words, the entries of $T$ strictly increase (left to right) along each row.

However, the $n$-tableau $T\mathbf{r}$ is obtained by reflecting $T$ across the northwest-to-southeast diagonal. This reflection clearly turns the rows of $T$ into the columns of $T\mathbf{r}$. More precisely, if some row of $T$ has entries $a_1, a_2, \ldots, a_k$ from left to right, then the corresponding column of $T\mathbf{r}$ has entries $a_1, a_2, \ldots, a_k$ from top to bottom. Thus, the entries of $T\mathbf{r}$ strictly increase (top to bottom) along each column (since the entries of $T$ strictly increase (left to right) along each row). In other words, the tableau $T\mathbf{r}$ is column-standard. A similar argument (using the column-standardness of $T$) shows that the tableau $T\mathbf{r}$ is row-standard.

Now, we know that the $n$-tableau $T\mathbf{r}$ is both row-standard and column-standard. In other words, it is standard. This proves Proposition 5.15.12 **(a)**.

**(b)** Let $T$ be any $n$-tableau of shape $D$. We must prove that $(T\mathbf{r})\,\mathbf{r} = T$.

Visually, this is obvious (since $T\mathbf{r}$ is the reflection of $T$ across the diagonal, and $(T\mathbf{r})\,\mathbf{r}$ is the reflection of this reflection across the same diagonal). The rigorous proof is not much harder: Recall that the map $\mathbf{r} : \mathbb{Z}^2 \to \mathbb{Z}^2$ is an involution, and thus we have $\mathbf{r}(\mathbf{r}(D)) = D$. Moreover, the definition of $T\mathbf{r}$ yields $T\mathbf{r} = T \circ \left(\mathbf{r}\mid_{\mathbf{r}(D)}\right)$, so that

$$\begin{aligned}
(T\mathbf{r})\,\mathbf{r} &= \left(T \circ \left(\mathbf{r}\mid_{\mathbf{r}(D)}\right)\right)\mathbf{r} \\
&= T \circ \underbrace{\left(\mathbf{r}\mid_{\mathbf{r}(D)}\right) \circ \left(\mathbf{r}\mid_{\mathbf{r}(\mathbf{r}(D))}\right)}_{\substack{=\mathrm{id}_D \\ \text{(since } \mathbf{r} \text{ is an involution)}}} \qquad \left(\text{by the definition of } \left(T \circ \left(\mathbf{r}\mid_{\mathbf{r}(D)}\right)\right)\mathbf{r}\right) \\
&= T \circ \mathrm{id}_D = T.
\end{aligned}$$

Thus, Proposition 5.15.12 **(b)** is proved.

**(c)** Recall that the map $\mathbf{r} : \mathbb{Z}^2 \to \mathbb{Z}^2$ is an involution, and thus we have $\mathbf{r}(\mathbf{r}(D)) = D$. Furthermore, Proposition 5.15.12 **(b)** yields that $(T\mathbf{r})\,\mathbf{r} = T$ for any $T \in \mathrm{SYT}(D)$. Likewise, we have $(T\mathbf{r})\,\mathbf{r} = T$ for any $T \in \mathrm{SYT}(\mathbf{r}(D))$.

Proposition 5.15.12 **(a)** is saying that if $T$ is a standard $n$-tableau of shape $D$, then $T\mathbf{r}$ is a standard $n$-tableau of shape $\mathbf{r}(D)$. In other words, if $T \in \mathrm{SYT}(D)$, then $T\mathbf{r} \in \mathrm{SYT}(\mathbf{r}(D))$ (since $T \in \mathrm{SYT}(D)$ means that $T$ is a standard $n$-tableau of shape $D$, whereas $T\mathbf{r} \in \mathrm{SYT}(\mathbf{r}(D))$ means that $T\mathbf{r}$ is a standard $n$-tableau of shape $\mathbf{r}(D)$). Thus,

the map

$$\begin{aligned} \operatorname{SYT}(D) &\to \operatorname{SYT}(\mathbf{r}(D)), \\ T &\mapsto T\mathbf{r} \end{aligned} \tag{538}$$

is well-defined. The same argument (applied to $\mathbf{r}(D)$ instead of $D$) yields that the map

$$\begin{aligned} \operatorname{SYT}(\mathbf{r}(D)) &\to \operatorname{SYT}(\mathbf{r}(\mathbf{r}(D))), \\ T &\mapsto T\mathbf{r} \end{aligned}$$

is well-defined as well. In other words, the map

$$\begin{aligned} \operatorname{SYT}(\mathbf{r}(D)) &\to \operatorname{SYT}(D), \\ T &\mapsto T\mathbf{r} \end{aligned} \tag{539}$$

is well-defined (since $\mathbf{r}(\mathbf{r}(D)) = D$). These two maps (538) and (539) are mutually inverse (since $(T\mathbf{r})\,\mathbf{r} = T$ for any $T \in \operatorname{SYT}(D)$, and since $(T\mathbf{r})\,\mathbf{r} = T$ for any $T \in \operatorname{SYT}(\mathbf{r}(D))$). Hence, in particular, the former map is invertible, i.e., is a bijection. This proves Proposition 5.15.12 **(c)**.

**(d)** Let $T$ be an $n$-tableau of shape $D$. Let $w \in S_n$. Then, $T\mathbf{r} = T \circ \left(\mathbf{r}\mid_{\mathbf{r}(D)}\right)$ (by the definition of $T\mathbf{r}$) and $(w \circ T)\,\mathbf{r} = w \circ T \circ \left(\mathbf{r}\mid_{\mathbf{r}(D)}\right)$ (likewise). But $wT = w \rightharpoonup T = w \circ T$ (by Definition 5.3.9). Thus,

$$(wT)\,\mathbf{r} = (w \circ T)\,\mathbf{r} = w \circ T \circ \left(\mathbf{r}\mid_{\mathbf{r}(D)}\right). \tag{540}$$

On the other hand, $w(T\mathbf{r}) = w \rightharpoonup (T\mathbf{r}) = w \circ (T\mathbf{r})$ (again by Definition 5.3.9), and thus

$$w(T\mathbf{r}) = w \circ \underbrace{(T\mathbf{r})}_{=T\circ\left(\mathbf{r}\mid_{\mathbf{r}(D)}\right)} = w \circ T \circ \left(\mathbf{r}\mid_{\mathbf{r}(D)}\right).$$

Comparing this with (540), we obtain $(wT)\,\mathbf{r} = w(T\mathbf{r})$. This proves Proposition 5.15.12 **(d)**.

**(e)** Let $T$ be an $n$-tableau of shape $D$. We must prove that $\mathcal{R}(T) = \mathcal{C}(T\mathbf{r})$ and $\mathcal{C}(T) = \mathcal{R}(T\mathbf{r})$.

Recall once again that the $n$-tableau $T\mathbf{r}$ is obtained by reflecting $T$ across the northwest-to-southeast diagonal. This reflection turns the rows of $T$ into the columns of $T\mathbf{r}$. Thus, two numbers $i, j \in [n]$ lie in the same row of $T$ if and only if they lie in the same column of $T\mathbf{r}$. Hence, a permutation $w \in S_n$ is horizontal for $T$ if and only if it is vertical for $T\mathbf{r}$ (because "horizontal for $T$" means that for each $i \in [n]$, the number $w(i)$ lies in the same row of $T$ as $i$ does, whereas "vertical for $T\mathbf{r}$" means that for each $i \in [n]$, the number $w(i)$ lies in the same column of $T\mathbf{r}$ as $i$ does).

But $\mathcal{R}(T)$ is defined as the subgroup of $S_n$ consisting of all permutations $w \in S_n$ that are horizontal for $T$, whereas $\mathcal{C}(T\mathbf{r})$ is defined as the subgroup of $S_n$ consisting of all permutations $w \in S_n$ that are vertical for $T\mathbf{r}$. Thus, $\mathcal{R}(T) = \mathcal{C}(T\mathbf{r})$ (since a permutation $w \in S_n$ is horizontal for $T$ if and only if it is vertical for $T\mathbf{r}$).

An analogous argument (but with the roles of rows and columns interchanged) shows that $\mathcal{C}(T) = \mathcal{R}(T\mathbf{r})$. (Alternatively, we can derive $\mathcal{C}(T) = \mathcal{R}(T\mathbf{r})$ by applying the already proved equality $\mathcal{R}(T) = \mathcal{C}(T\mathbf{r})$ to $\mathbf{r}(D)$ and $T\mathbf{r}$ instead of $D$ and $T$, and then simplifying the result using Proposition 5.15.12 **(b)**.)

The proof of Proposition 5.15.12 **(e)** is now complete (since we have proved both $\mathcal{R}(T) = \mathcal{C}(T\mathbf{r})$ and $\mathcal{C}(T) = \mathcal{R}(T\mathbf{r})$). □

## A.3. The action of conjugacy class sums on Specht modules

Here we supply detailed proofs for some claims made in Section 5.16.

*Proof of Proposition 5.16.3.* Let us write the partition $\lambda$ as $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$. Then, the definition of $Y(\lambda)$ yields

$$Y(\lambda) = \{(i,j) \in \{1,2,\ldots,k\} \times \{1,2,3,\ldots\} \mid j \leq \lambda_i\}.$$

Thus, the summation sign $\sum_{(i,j)\in Y(\lambda)}$ can be rewritten as $\sum_{\substack{(i,j)\in\{1,2,\ldots,k\}\times\{1,2,3,\ldots\}; \\ j\leq\lambda_i}}$. Hence,

$$\sum_{(i,j)\in Y(\lambda)} (i-1) = \underbrace{\sum_{\substack{(i,j)\in\{1,2,\ldots,k\}\times\{1,2,3,\ldots\}; \\ j\leq\lambda_i}} (i-1)}_{= \sum\limits_{i\in\{1,2,\ldots,k\}} \ \sum\limits_{\substack{j\in\{1,2,3,\ldots\}; \\ j\leq\lambda_i}}} = \underbrace{\sum_{i\in\{1,2,\ldots,k\}}}_{=\sum\limits_{i=1}^{k}} \ \underbrace{\sum_{\substack{j\in\{1,2,3,\ldots\}; \\ j\leq\lambda_i}}}_{=\sum\limits_{j=1}^{\lambda_i}} (i-1)$$

$$= \sum_{i=1}^{k} \underbrace{\sum_{j=1}^{\lambda_i} (i-1)}_{\substack{=\lambda_i(i-1) \\ =(i-1)\lambda_i}} = \sum_{i=1}^{k} (i-1)\lambda_i = \mathrm{n}(\lambda)$$

(by the definition of $\mathrm{n}(\lambda)$). This proves Proposition 5.16.3. □

*Proof of Proposition 5.16.4.* Let us write the partition $\lambda$ itself as $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$. We shall first show that

$$\mathrm{n}(\lambda^t) = \sum_{i=1}^{k} \binom{\lambda_i}{2}.$$

Once this is proved, we will then apply this to $\lambda^t$ instead of $\lambda$ and easily derive the proposition.

Theorem 5.1.10 **(b)** shows that

$$\begin{aligned} \lambda_j^t &= (\#\text{ of positive integers } i \text{ such that } \lambda_i \geq j) \\ &= (\#\text{ of cells in the } j\text{-th column of } Y(\lambda)) \qquad \text{for all } j \geq 1. \end{aligned} \tag{541}$$

Thus, for all $j \geq 1$, we have

$$\lambda_j^t = \sum_{\substack{i\in[k]; \\ j\leq\lambda_i}} 1 \tag{542}$$

[181].

Moreover, for each $u \geq 1$, we have $\lambda_u \leq m$ [182] and thus $[\lambda_u] \subseteq [m]$, so that

$$[m] \cap [\lambda_u] = [\lambda_u]. \tag{545}$$

On the other hand, each $p \in \mathbb{N}$ satisfies

$$\begin{aligned}
\sum_{j \in [p]} (j-1) &= \sum_{j=1}^{p} (j-1) = 0 + 1 + \cdots + (p-1) \\
&= 1 + 2 + \cdots + (p-1) \\
&= \frac{(p-1)\left((p-1)+1\right)}{2} \qquad \text{(by the Little Gauss formula)} \\
&= \frac{p\,(p-1)}{2} = \binom{p}{2}.
\end{aligned} \tag{546}$$

---

[181] *Proof.* Let $j \geq 1$. Recall that $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$. Hence,

$$\lambda_i = 0 \qquad \text{for each } i > k. \tag{543}$$

Hence, if $i$ is a positive integer such that $\lambda_i \geq j$, then we must have $i \leq k$ (since otherwise, we would have $i > k$ and thus $\lambda_i = 0$ (by (543)), which would contradict $\lambda_i \geq j \geq 1 > 0$) and therefore $i \in [k]$. In other words, the positive integers $i$ that satisfy $\lambda_i \geq j$ all belong to $[k]$. Hence, these positive integers are precisely the $i \in [k]$ that satisfy $\lambda_i \geq j$. Therefore,

$$\begin{aligned}
&(\#\text{ of positive integers } i \text{ that satisfy } \lambda_i \geq j) \\
&= (\#\text{ of } i \in [k] \text{ that satisfy } \lambda_i \geq j).
\end{aligned} \tag{544}$$

Now,

$$\begin{aligned}
\sum_{\substack{i \in [k]; \\ j \leq \lambda_i}} 1 &= (\#\text{ of } i \in [k] \text{ that satisfy } j \leq \lambda_i) \cdot 1 \\
&= (\#\text{ of } i \in [k] \text{ that satisfy } j \leq \lambda_i) \\
&= (\#\text{ of } i \in [k] \text{ that satisfy } \lambda_i \geq j) \qquad \text{(here, we rewrote ``}j \leq \lambda_i\text{'' as ``}\lambda_i \geq j\text{''}) \\
&= (\#\text{ of positive integers } i \text{ that satisfy } \lambda_i \geq j) \qquad \text{(by (544))} \\
&= (\#\text{ of positive integers } i \text{ such that } \lambda_i \geq j) = \lambda_j^t \qquad \text{(by (541))}.
\end{aligned}$$

This proves (542).

[182] *Proof.* Let $u \geq 1$. We must prove that $\lambda_u \leq m$.

Let $j = \lambda_u$. Thus, $\lambda_u \geq j$ (since $\lambda_u = j$). Hence, there exists at least one positive integer $i$ such that $\lambda_i \geq j$ (namely, the integer $i = u$). In other words, $(\#\text{ of positive integers } i \text{ such that } \lambda_i \geq j) \geq 1$. Using (541), we can rewrite this as $\lambda_j^t \geq 1$.

However, $\lambda^t = \left(\lambda_1^t, \lambda_2^t, \ldots, \lambda_m^t\right)$, and thus $\lambda_p^t = 0$ for all $p > m$. Hence, if we had $j > m$, then we would have $\lambda_j^t = 0$, which would contradict $\lambda_j^t \geq 1 > 0$. Thus, we cannot have $j > m$. Hence, we have $j \leq m$, so that $\lambda_u = j \leq m$. Qed.

Hence, for each $i \geq 1$, we have

$$\underbrace{\sum_{\substack{j \in [m]; \\ j \leq \lambda_i}}}_{\substack{= \sum_{\substack{j \in [m]; \\ j \in [\lambda_i]}} \\ \text{(since the condition ``}j \leq \lambda_i\text{'' on a positive integer } j \text{ is equivalent to ``}j \in [\lambda_i]\text{'')}}} (j-1) = \underbrace{\sum_{\substack{j \in [m]; \\ j \in [\lambda_i]}}}_{\substack{= \sum_{j \in [m] \cap [\lambda_i]} = \sum_{j \in [\lambda_i]} \\ \text{(since (545) (applied to } u=i\text{)} \\ \text{yields } [m] \cap [\lambda_i] = [\lambda_i]\text{)}}} (j-1) = \sum_{j \in [\lambda_i]} (j-1)$$

$$= \binom{\lambda_i}{2} \tag{547}$$

(by (546), applied to $p = \lambda_i$).

Now, $\lambda^t = \left(\lambda_1^t, \lambda_2^t, \ldots, \lambda_m^t\right)$. Thus, the definition of $\mathrm{n}\left(\lambda^t\right)$ yields

$$\mathrm{n}\left(\lambda^t\right) = \sum_{i=1}^{m} (i-1)\, \lambda_i^t = \underbrace{\sum_{j=1}^{m}}_{= \sum_{j \in [m]}} (j-1) \underbrace{\lambda_j^t}_{\substack{= \sum_{\substack{i \in [k]; \\ j \leq \lambda_i}} 1 \\ \text{(by (542))}}} = \sum_{j \in [m]} (j-1) \sum_{\substack{i \in [k]; \\ j \leq \lambda_i}} 1$$

$$= \underbrace{\sum_{j \in [m]} \sum_{\substack{i \in [k]; \\ j \leq \lambda_i}}}_{= \sum_{i \in [k]} \sum_{\substack{j \in [m]; \\ j \leq \lambda_i}}} \underbrace{(j-1) \cdot 1}_{=j-1} = \underbrace{\sum_{i \in [k]}}_{= \sum_{i=1}^{k}} \underbrace{\sum_{\substack{j \in [m]; \\ j \leq \lambda_i}} (j-1)}_{\substack{= \binom{\lambda_i}{2} \\ \text{(by (547))}}} = \sum_{i=1}^{k} \binom{\lambda_i}{2}.$$

The same reasoning (applied to $\lambda^t$, $k$ and $m$ instead of $\lambda$, $m$ and $k$) shows that

$$\mathrm{n}\left(\left(\lambda^t\right)^t\right) = \sum_{i=1}^{m} \binom{\lambda_i^t}{2}. \tag{548}$$

However, Theorem 5.1.10 **(d)** says that $\left(\lambda^t\right)^t = \lambda$. Hence, we can rewrite (548) as

$$\mathrm{n}\left(\lambda\right) = \sum_{i=1}^{m} \binom{\lambda_i^t}{2}.$$

This proves Proposition 5.16.4. □

*Proof of Proposition 5.16.5.* We know that $T$ is an $n$-tableau of shape $\lambda$. That is, $T$ is an $n$-tableau of shape $Y\left(\lambda\right)$. In particular, $T$ is injective (like any $n$-tableau), i.e., all entries of $T$ are distinct.

Write the conjugate $\lambda^t$ of $\lambda$ as $\lambda^t = \left(\lambda_1^t, \lambda_2^t, \ldots, \lambda_m^t\right)$. Thus,

$$\lambda_k^t = 0 \qquad \text{for all } k > m. \tag{549}$$

Any transposition in $S_n$ has the form $t_{i,j}$ for some 2-element subset $\{i, j\}$ of $[n]$; moreover, the subset $\{i, j\}$ is uniquely determined by this transposition. Moreover, a

transposition $t_{i,j} \in S_n$ belongs to $\mathcal{C}(T)$ if and only if the numbers $i$ and $j$ lie in the same column of $T$ [183]. Hence, the transpositions in $\mathcal{C}(T)$ are precisely the transpositions of the form $t_{i,j}$, where $i$ and $j$ are two distinct numbers lying in the same column of $T$. Therefore,

$$\begin{aligned}
&(\#\text{ of transpositions in } \mathcal{C}(T)) \\
&= (\#\text{ of transpositions of the form } t_{i,j}\text{, where } i \text{ and } j \text{ are} \\
&\qquad \text{two distinct numbers lying in the same column of } T) \\
&= \sum_{k \geq 1} (\#\text{ of transpositions of the form } t_{i,j}\text{, where } i \text{ and } j \text{ are} \\
&\qquad \text{two distinct numbers lying in the } k\text{-th column of } T) \qquad (550)
\end{aligned}$$

(since all entries of $T$ are distinct, and thus a given number $i$ cannot lie in two different columns of $T$ at once).

However, we can prove the following:

*Claim 1:* Each $k \geq 1$ satisfies

$$\begin{aligned}
&(\#\text{ of transpositions of the form } t_{i,j}\text{, where } i \text{ and } j \text{ are} \\
&\qquad \text{two distinct numbers lying in the } k\text{-th column of } T) \\
&= \binom{\lambda_k^t}{2}. \qquad (551)
\end{aligned}$$

*Proof of Claim 1.* Let $k \geq 1$. From Theorem 5.1.10 **(b)** (applied to $j = k$), we obtain

$$\begin{aligned}
\lambda_k^t &= (\#\text{ of positive integers } i \text{ such that } \lambda_i \geq k) \\
&= (\#\text{ of cells in the } k\text{-th column of } Y(\lambda)).
\end{aligned}$$

Hence, the $k$-th column of $Y(\lambda)$ has $\lambda_k^t$ cells. Thus, the $k$-th column of $T$ has $\lambda_k^t$ entries (since $T$ is an $n$-tableau of shape $Y(\lambda)$). All these $\lambda_k^t$ entries are distinct (since all entries of $T$ are distinct). Hence, there are $\lambda_k^t$ distinct entries in the $k$-th column of $T$. Thus, there are $\binom{\lambda_k^t}{2}$ ways to choose two distinct entries $i$ and $j$ from the $k$-th column of $T$, provided that the order does not matter (i.e., we do not distinguish between choosing $i$ and $j$ and choosing $j$ and $i$). Each of these $\binom{\lambda_k^t}{2}$ ways yields a transposition $t_{i,j}$, and all these $\binom{\lambda_k^t}{2}$ transpositions are distinct (since a transposition $t_{i,j}$ uniquely determines

[183] *Proof.* Let $t_{i,j} \in S_n$ be a transposition. If the numbers $i$ and $j$ lie in the same column of $T$, then the transposition $t_{i,j}$ is vertical for $T$ and thus belongs to $\mathcal{C}(T)$. On the other hand, if the numbers $i$ and $j$ do not lie in the same column of $T$, then the transposition $t_{i,j}$ is not vertical for $T$ (since $t_{i,j}(i) = j$ does not lie in the same column of $T$ as $i$ does), and thus does not belong to $\mathcal{C}(T)$. Combining the preceding two sentences, we conclude that $t_{i,j}$ belongs to $\mathcal{C}(T)$ if and only if the numbers $i$ and $j$ lie in the same column of $T$. Qed.

the 2-element set $\{i,j\}$). Hence,

$$\begin{aligned}&\left(\begin{array}{c}\text{\# of transpositions of the form } t_{i,j}\text{, where } i \text{ and } j \text{ are}\\ \text{two distinct numbers lying in the } k\text{-th column of } T\end{array}\right)\\ &= \binom{\lambda_k^t}{2}.\end{aligned}$$

Thus, Claim 1 is proved. □

Now, using (551), we can rewrite (550) as

$$\begin{aligned}\left(\text{\# of transpositions in } \mathcal{C}(T)\right) &= \sum_{k\geq 1}\binom{\lambda_k^t}{2} = \sum_{k=1}^{m}\binom{\lambda_k^t}{2} + \sum_{k=m+1}^{\infty}\underbrace{\binom{\lambda_k^t}{2}}_{\substack{=0\\ \text{(since } k\geq m+1>m\text{,}\\ \text{so that (549) yields } \lambda_k^t=0\\ \text{and thus } \binom{\lambda_k^t}{2}=\binom{0}{2}=0\text{)}}}\\ &= \sum_{k=1}^{m}\binom{\lambda_k^t}{2} + \underbrace{\sum_{k=m+1}^{\infty}0}_{=0} = \sum_{k=1}^{m}\binom{\lambda_k^t}{2} = \sum_{i=1}^{m}\binom{\lambda_i^t}{2}\\ &= \mathrm{n}(\lambda) \qquad \text{(by Proposition 5.16.4)}.\end{aligned}$$

In other words, $\mathrm{n}(\lambda)$ is the number of transpositions in $\mathcal{C}(T)$. This proves Proposition 5.16.5. □

## A.4. What to do with a basis

Here we supply detailed proofs for some claims made in Section 6.2.

*Detailed proof of Lemma 6.2.1.* For each partition $\lambda$ of $n$, we have

$$\begin{aligned}&\left(\text{\# of standard } n\text{-tableaux of shape } \lambda\right)\\ &= \left(\text{\# of standard } n\text{-tableaux of shape } Y(\lambda)\right)\\ &= \left(\text{\# of standard tableaux of shape } Y(\lambda)\right)\\ &\qquad \left(\begin{array}{c}\text{since Proposition 5.2.6 (applied to } D = Y(\lambda)\text{)}\\ \text{shows that the standard tableaux of shape } Y(\lambda)\\ \text{are precisely the standard } n\text{-tableaux of shape } Y(\lambda)\\ \text{(because } |Y(\lambda)| = |\lambda| = n\text{)}\end{array}\right)\\ &= f^{\lambda} \qquad \left(\text{by the definition of } f^{\lambda}\right). \end{aligned} \tag{552}$$

Recall that $\mathrm{SBT}(n)$ is the set of all standard $n$-bitableaux, i.e., of all triples $(\lambda, U, V)$ consisting of a partition $\lambda$ of $n$ and two standard $n$-tableaux $U, V$ of shape $\lambda$. In other words,

$$\mathrm{SBT}(n) = \bigcup_{\lambda \text{ is a partition of } n} \{(\lambda, U, V) \mid U \text{ and } V \text{ are standard } n\text{-tableaux of shape } \lambda\}.$$

This union is a disjoint union (since $\lambda$ can be uniquely recovered from the triple $(\lambda, U, V)$). Since the size of a disjoint union equals the sum of the sizes of the sets involved, we thus obtain

$$\begin{aligned}
&|\mathrm{SBT}(n)| \\
&= \sum_{\lambda \text{ is a partition of } n} \underbrace{\left|\left\{(\lambda, U, V) \mid U \text{ and } V \text{ are standard } n\text{-tableaux of shape } \lambda\right\}\right|}_{\substack{=(\#\text{ of pairs } (U,V) \text{ where } U \text{ and } V \text{ are standard } n\text{-tableaux of shape } \lambda) \\ =(\#\text{ of standard } n\text{-tableaux of shape } \lambda)^2 \\ =\left(f^\lambda\right)^2 \\ \text{(by (552))}}} \\
&= \sum_{\lambda \text{ is a partition of } n} \left(f^\lambda\right)^2 = n! \qquad \text{(by Corollary 5.12.18)}.
\end{aligned}$$

This proves Lemma 6.2.1. □

## A.5. Combinatorial interlude: Lexicographic orders

Here we supply detailed proofs for some claims made in Section 6.7.

*Detailed proof of Proposition 6.7.2.* First, we shall show that the relation $<$ is the smaller relation of a partial order on the set of all partitions. In order to prove this, we need to verify three claims:

> *Claim 1:* The relation $<$ is irreflexive (i.e., there exists no partition $\lambda$ such that $\lambda < \lambda$).
>
> *Claim 2:* The relation $<$ is transitive (i.e., if three partitions $\lambda$, $\mu$ and $\nu$ satisfy $\lambda < \mu$ and $\mu < \nu$, then $\lambda < \nu$).
>
> *Claim 3:* The relation $<$ is asymmetric (i.e., no two partitions $\lambda$ and $\mu$ satisfy $\lambda < \mu$ and $\mu < \lambda$ at the same time).

*Proof of Claim 1.* Let $\lambda$ be a partition such that $\lambda < \lambda$. We shall derive a contradiction.

Indeed, we have $\lambda < \lambda$. By the definition of the relation $<$, this means that

- there exists at least one $i \geq 1$ such that $\lambda_i \neq \lambda_i$, and
- the **smallest** such $i$ satisfies $\lambda_i < \lambda_i$.

But the first of these two bullet points is clearly absurd (since $\lambda_i \neq \lambda_i$ can never happen). Thus, we have found a contradiction.

Forget that we fixed $\lambda$. We thus have obtained a contradiction for each partition $\lambda$ such that $\lambda < \lambda$. Hence, there exists no such $\lambda$. In other words, the relation $<$ is irreflexive. This proves Claim 1. □

*Proof of Claim 2.* Let $\lambda$, $\mu$ and $\nu$ be three partitions that satisfy $\lambda < \mu$ and $\mu < \nu$. We must prove that $\lambda < \nu$.

Indeed, we have $\lambda < \mu$. By the definition of the relation $<$, this means that

- there exists at least one $i \geq 1$ such that $\lambda_i \neq \mu_i$, and
- the **smallest** such $i$ satisfies $\lambda_i < \mu_i$.

Let us denote this smallest $i$ by $p$. Thus, we have $\lambda_p < \mu_p$, whereas

$$\text{each } i < p \text{ satisfies } \lambda_i = \mu_i \tag{553}$$

(since $p$ is the **smallest** $i \geq 1$ that satisfies $\lambda_i \neq \mu_i$).

Furthermore, we have $\mu < \nu$. By the definition of the relation $<$, this means that

- there exists at least one $i \geq 1$ such that $\mu_i \neq \nu_i$, and
- the **smallest** such $i$ satisfies $\mu_i < \nu_i$.

Let us denote this smallest $i$ by $q$. Thus, we have $\mu_q < \nu_q$, whereas

$$\text{each } i < q \text{ satisfies } \mu_i = \nu_i \tag{554}$$

(since $q$ is the **smallest** $i \geq 1$ that satisfies $\mu_i \neq \nu_i$). Hence, it is easy to see that

$$\text{each } i \leq q \text{ satisfies } \mu_i \leq \nu_i \tag{555}$$

[184].

Now, let $w := \min\{p, q\}$. Then, $w \leq p$ and $w \leq q$. Hence,

$$\text{each } i < w \text{ satisfies } \lambda_i = \nu_i \tag{556}$$

[185]. Furthermore,

$$\lambda_w < \nu_w \tag{557}$$

[186] and therefore $\lambda_w \neq \nu_w$. Hence, $w$ is an $i \geq 1$ such that $\lambda_i \neq \nu_i$. Moreover, $w$ is the **smallest** such $i$ (since (556) shows that all smaller $i$'s satisfy $\lambda_i = \nu_i$). Hence, we have now shown that

[184] *Proof of (555):* Let $i$ be an integer such that $i \leq q$. We must show that $\mu_i \leq \nu_i$. If $i < q$, then this follows from (554) (since (554) says $\mu_i = \nu_i$ in this case, and hence $\mu_i \leq \nu_i$). Hence, we WLOG assume that we don't have $i < q$. Thus, $i \geq q$. Combining this with $i \leq q$, we obtain $i = q$. Hence, $\mu_i = \mu_q < \nu_q = \nu_i$ (since $q = i$). Therefore, $\mu_i \leq \nu_i$. Thus, (555) is proved.

[185] *Proof.* Let $i \geq 1$ be such that $i < w$. Then,

$$\begin{aligned} \lambda_i &= \mu_i && \text{(by (553), since } i < w \leq p\text{)} \\ &= \nu_i && \text{(by (554), since } i < w \leq q\text{)}. \end{aligned}$$

This proves (556).

[186] *Proof of (557):* We are in one of the following two cases:

*Case 1:* We have $p \leq q$.

*Case 2:* We have $p > q$.

Let us first consider Case 1. In this case, we have $p \leq q$. Hence, $w = \min\{p, q\} = p$ (since $p \leq q$). Applying (555) to $i = w$, we obtain $\mu_w \leq \nu_w$ (since $w \leq q$). Recall that $\lambda_p < \mu_p$. Since $w = p$, we can rewrite this as $\lambda_w < \mu_w$. Hence, $\lambda_w < \mu_w \leq \nu_w$. Hence, (557) is proved in Case 1.

Let us next consider Case 2. In this case, we have $p > q$. Hence, $w = \min\{p, q\} = q$ (since $p > q$). Thus, $w = q < p$ (since $p > q$) and therefore $\lambda_w = \mu_w$ (by (553), applied to $i = w$). Recall that $\mu_q < \nu_q$. Since $w = q$, we can rewrite this as $\mu_w < \nu_w$. Hence, $\lambda_w = \mu_w < \nu_w$. Hence, (557) is proved in Case 2.

Thus, (557) is proved in both Cases 1 and 2.

- there exists at least one $i \geq 1$ such that $\lambda_i \neq \nu_i$ (namely, $i = w$), and
- the **smallest** such $i$ satisfies $\lambda_i < \nu_i$ (because the smallest such $i$ is $w$, and we know that $w$ satisfies $\lambda_w < \nu_w$).

In other words, $\lambda < \nu$ (by the definition of the relation $<$). This proves Claim 2. $\square$

*Proof of Claim 3.* Any binary relation that is irreflexive and transitive is automatically asymmetric (since $\lambda < \mu$ and $\mu < \lambda$ would entail $\lambda < \lambda$ by transitivity, but this would contradict irreflexivity). Since the relation $<$ is irreflexive (by Claim 1) and transitive (by Claim 2), we thus conclude that it is also asymmetric. This proves Claim 3. $\square$

Having proved Claims 1, 2 and 3, we thus conclude that the relation $<$ is the smaller relation of a partial order. It remains to show that this partial order is a total order. For this purpose, we need to prove the following last claim:

> *Claim 4:* Let $\lambda$ and $\mu$ be two distinct partitions. Then, we have $\lambda < \mu$ or $\mu < \lambda$.

*Proof of Claim 4.* We shall first show that there exists some $i \geq 1$ such that $\lambda_i \neq \mu_i$. Indeed, we must be in one of the three cases $\ell(\lambda) > \ell(\mu)$ and $\ell(\mu) > \ell(\lambda)$ and $\ell(\lambda) = \ell(\mu)$, but we can easily find such an $i$ in each of these three cases:

- If $\ell(\lambda) > \ell(\mu)$, then it is easy to see that $\ell(\lambda)$ is such an $i$ [187].
- If $\ell(\mu) > \ell(\lambda)$, then it is easy to see that $\ell(\mu)$ is such an $i$ [188].
- If $\ell(\lambda) = \ell(\mu)$, then such an $i$ also exists[189].

Hence, we have shown that there exists some $i \geq 1$ such that $\lambda_i \neq \mu_i$.

Consider the **smallest** such $i$. Then, this $i$ satisfies $\lambda_i \neq \mu_i$. Hence, either $\lambda_i < \mu_i$ or $\lambda_i > \mu_i$. In the former case, we have $\lambda < \mu$ (by the definition of the relation $<$). In the

[187] *Proof.* Assume that $\ell(\lambda) > \ell(\mu)$. We must show that $\ell(\lambda)$ is an $i \geq 1$ such that $\lambda_i \neq \mu_i$. In other words, we must prove that $\ell(\lambda) \geq 1$ and $\lambda_{\ell(\lambda)} \neq \mu_{\ell(\lambda)}$.

Indeed, we have $\ell(\lambda) > \ell(\mu) \geq 0$ and thus $\ell(\lambda) \geq 1$ (since $\ell(\lambda)$ is an integer). Moreover, from $\ell(\lambda) > \ell(\mu)$, we obtain $\mu_{\ell(\lambda)} = 0$ (since each integer $i > \ell(\mu)$ satisfies $\mu_i = 0$ according to Definition 5.1.1 **(h)**). But $\lambda_{\ell(\lambda)} > 0$ (since $\lambda_{\ell(\lambda)}$ is the last entry of the partition $\lambda$, and thus is a positive integer). Hence, $\lambda_{\ell(\lambda)} > 0 = \mu_{\ell(\lambda)}$, so that $\lambda_{\ell(\lambda)} \neq \mu_{\ell(\lambda)}$. This completes our proof.

[188] *Proof.* The argument is the same as for the previous sentence, but the roles of $\lambda$ and $\mu$ are switched.

[189] *Proof.* Assume that $\ell(\lambda) = \ell(\mu)$. Thus, $\ell(\mu) = \ell(\lambda)$. But $\lambda \neq \mu$ (since $\lambda$ and $\mu$ are distinct). In light of $\lambda = \left(\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}\right)$ and $\mu = \left(\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}\right) = \left(\mu_1, \mu_2, \ldots, \mu_{\ell(\lambda)}\right)$ (since $\ell(\mu) = \ell(\lambda)$), we can rewrite this as

$$\left(\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}\right) \neq \left(\mu_1, \mu_2, \ldots, \mu_{\ell(\lambda)}\right).$$

Hence, there exists some $i \in [\ell(\lambda)]$ such that $\lambda_i \neq \mu_i$ (because two distinct $\ell(\lambda)$-tuples must differ in at least one position). Therefore, there exists some $i \geq 1$ such that $\lambda_i \neq \mu_i$ (because $i \in [\ell(\lambda)]$ entails $i \geq 1$).

latter case, we have $\mu < \lambda$ (by the definition of the relation $<$ again[190]). Thus, we have found that $\lambda < \mu$ or $\mu < \lambda$. This proves Claim 4. $\square$

Claim 4 shows that the partial order whose smaller relation is $<$ is actually a total order. The proof of Proposition 6.7.2 is thus complete. $\square$

*Detailed proof of Proposition 6.7.6.* First, we shall show that the relation $<$ is the smaller relation of a partial order on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$. In order to prove this, we need to verify three claims:

> *Claim 1:* The relation $<$ is irreflexive (i.e., there exists no $P \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ such that $P < P$).
>
> *Claim 2:* The relation $<$ is transitive (i.e., if three tableaux $P, Q, R \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ satisfy $P < Q$ and $Q < R$, then $P < R$).
>
> *Claim 3:* The relation $<$ is asymmetric (i.e., no two tableaux $P$ and $Q$ satisfy $P < Q$ and $Q < P$ at the same time).

*Proof of Claim 1.* Let $P \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ be a tableau such that $P < P$. We shall derive a contradiction.

Indeed, we have $P \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$. Hence, $P \in \operatorname{SYT}(\lambda)$ for some partition $\lambda$. Consider this $\lambda$. Now, we have assumed that $P < P$. By the definition of the relation $<$, this means that

$$\left(\lambda < \lambda \text{ or } \left(\lambda = \lambda \text{ and } \overline{P} < \overline{P}\right)\right)$$

(since $P \in \operatorname{SYT}(\lambda)$). But this is clearly false (since neither $\lambda < \lambda$ nor $\overline{P} < \overline{P}$ can hold (because both relations "$<$" in these two inequalities are total orders)). Thus, we have found a contradiction.

Forget that we fixed $P$. We thus have obtained a contradiction for each tableau $P$ such that $P < P$. Hence, there exists no such $P$. In other words, the relation $<$ is irreflexive. This proves Claim 1. $\square$

*Proof of Claim 2.* Let $P, Q, R \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ be three tableaux that satisfy $P < Q$ and $Q < R$. We must prove that $P < R$.

Indeed, we have $P \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$. Hence, we have $P \in \operatorname{SYT}(\lambda)$ for some partition $\lambda$. Similarly, we have $Q \in \operatorname{SYT}(\mu)$ for some partition $\mu$. Likewise, we have $R \in \operatorname{SYT}(\nu)$ for some partition $\nu$. Consider these three partitions $\lambda, \mu, \nu$.

Now, we have assumed that $P < Q$. By the definition of the relation $<$ on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, this means that

$$\left(\lambda < \mu \text{ or } \left(\lambda = \mu \text{ and } \overline{P} < \overline{Q}\right)\right)$$

[190] *Proof.* Assume that $\lambda_i > \mu_i$. Thus, $\mu_i < \lambda_i$. But recall that our $i$ is the **smallest** $i \geq 1$ such that $\lambda_i \neq \mu_i$. In other words, our $i$ is the **smallest** $i \geq 1$ such that $\mu_i \neq \lambda_i$. Hence, there exists an $i \geq 1$ such that $\mu_i \neq \lambda_i$, and the **smallest** such $i$ satisfies $\mu_i < \lambda_i$. In other words, we have $\mu < \lambda$ (by the definition of the relation $<$), qed.

(since $P \in \operatorname{SYT}(\lambda)$ and $Q \in \operatorname{SYT}(\mu)$). Thus, we must have either $\lambda < \mu$ or $\left(\lambda = \mu \text{ and } \overline{P} < \overline{Q}\right)$. In both of these cases, we have $\lambda \leq \mu$ (since $\lambda \leq \mu$ follows both from $\lambda < \mu$ and from $\lambda = \mu$).

So we have derived $\lambda \leq \mu$ from $P < Q$. Likewise, from $Q < R$, we can derive $\mu \leq \nu$. But the lexicographic order on partitions is a total order; thus, the relation $\leq$ on the set of all partitions is transitive. Hence, from $\lambda \leq \mu$ and $\mu \leq \nu$, we obtain $\lambda \leq \nu$. In other words, we have either $\lambda = \nu$ or $\lambda < \nu$. Hence, we are in one of the following two cases:

*Case 1:* We have $\lambda = \nu$.

*Case 2:* We have $\lambda < \nu$.

Let us first consider Case 1. In this case, we have $\lambda = \nu$. Thus, $\nu = \lambda \leq \mu$. Combining this with $\mu \leq \nu$, we obtain $\mu = \nu$ (since the lexicographic order on partitions is a total order, so that the relation $\leq$ on the set of all partitions is antisymmetric). Hence, $\mu = \nu = \lambda$, so that $\lambda = \mu$.

Now, recall that we have either $\lambda < \mu$ or $\left(\lambda = \mu \text{ and } \overline{P} < \overline{Q}\right)$. Since $\lambda < \mu$ is impossible (because $\lambda = \mu$), we thus conclude that $\left(\lambda = \mu \text{ and } \overline{P} < \overline{Q}\right)$ must hold. Hence, $\overline{P} < \overline{Q}$.

Moreover, we have assumed that $Q < R$. By the definition of the relation $<$ on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, this means that

$$\left(\mu < \nu \text{ or } \left(\mu = \nu \text{ and } \overline{Q} < \overline{R}\right)\right)$$

(since $Q \in \operatorname{SYT}(\mu)$ and $R \in \operatorname{SYT}(\nu)$). Since $\mu < \nu$ is impossible (because $\mu = \nu$), we thus must have $\left(\mu = \nu \text{ and } \overline{Q} < \overline{R}\right)$. Hence, $\overline{Q} < \overline{R}$.

But the Young last letter order on the $n$-tabloids is a total order, thus transitive. Hence, from $\overline{P} < \overline{Q}$ and $\overline{Q} < \overline{R}$, we obtain $\overline{P} < \overline{R}$. Combining this with $\lambda = \nu$, we obtain $\left(\lambda = \nu \text{ and } \overline{P} < \overline{R}\right)$. Hence, we have

$$\left(\lambda < \nu \text{ or } \left(\lambda = \nu \text{ and } \overline{P} < \overline{R}\right)\right).$$

But this is equivalent to saying that $P < R$ (by the definition of the relation $<$ on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, because $P \in \operatorname{SYT}(\lambda)$ and $R \in \operatorname{SYT}(\nu)$). Thus, $P < R$ is proved in Case 1.

Let us now consider Case 2. In this case, we have $\lambda < \nu$. Hence, we have

$$\left(\lambda < \nu \text{ or } \left(\lambda = \nu \text{ and } \overline{P} < \overline{R}\right)\right).$$

But this is equivalent to saying that $P < R$ (by the definition of the relation $<$ on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, because $P \in \operatorname{SYT}(\lambda)$ and $R \in \operatorname{SYT}(\nu)$). Thus, $P < R$ is proved in Case 2.

We have now proved $P < R$ in both Cases 1 and 2. Hence, $P < R$ always holds. This proves Claim 2. □

*Proof of Claim 3.* Any binary relation that is irreflexive and transitive is automatically asymmetric (since $P < Q$ and $Q < P$ would entail $P < P$ by transitivity, but this would contradict irreflexivity). Since the relation $<$ is irreflexive (by Claim 1) and transitive (by Claim 2), we thus conclude that it is also asymmetric. This proves Claim 3. □

Having proved Claims 1, 2 and 3, we thus conclude that the relation $<$ is the smaller relation of a partial order. It remains to show that this partial order is a total order. For this purpose, we need to prove the following last claim:

*Claim 4:* Let $P$ and $Q$ be two distinct tableaux in $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$. Then, we have $P < Q$ or $Q < P$.

*Proof of Claim 4.* We have $P \in \bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$. Hence, we have $P \in \operatorname{SYT}(\lambda)$ for some partition $\lambda$. Similarly, we have $Q \in \operatorname{SYT}(\mu)$ for some partition $\mu$. Consider these two partitions $\lambda$ and $\mu$.

The lexicographic order on partitions is a total order. Hence, we have either $\lambda < \mu$ or $\lambda = \mu$ or $\lambda > \mu$. In other words, we are in one of the following three cases:

*Case 1:* We have $\lambda < \mu$.

*Case 2:* We have $\lambda = \mu$.

*Case 3:* We have $\lambda > \mu$.

Let us consider Case 1 first. In this case, we have $\lambda < \mu$. Thus,

$$\left(\lambda < \mu \text{ or } \left(\lambda = \mu \text{ and } \overline{P} < \overline{Q}\right)\right).$$

But this is equivalent to saying that $P < Q$ (by the definition of the relation $<$ on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, since $P \in \operatorname{SYT}(\lambda)$ and $Q \in \operatorname{SYT}(\mu)$). Thus, we have $P < Q$, and therefore ($P < Q$ or $Q < P$). Thus, Claim 4 is proved in Case 1.

Now consider Case 2. In this case, we have $\lambda = \mu$. Hence, $\operatorname{SYT}(\lambda) = \operatorname{SYT}(\mu)$. Thus, both $P$ and $Q$ are standard $n$-tableaux of shape $Y(\lambda)$ (since $P \in \operatorname{SYT}(\lambda)$ and $Q \in \operatorname{SYT}(\mu) = \operatorname{SYT}(\lambda)$). Thus, $P$ and $Q$ are row-standard $n$-tableaux of shape $Y(\lambda)$ (since standard tableaux are always row-standard). Proposition 5.3.8 (applied to $D = Y(\lambda)$) shows that there is a bijection

$$\begin{aligned} &\text{from } \{\text{row-standard } n\text{-tableaux of shape } Y(\lambda)\} \\ &\text{to } \{n\text{-tabloids of shape } Y(\lambda)\} \end{aligned}$$

that sends each row-standard $n$-tableau $T$ to its tabloid $\overline{T}$. Hence, in particular, this bijection is injective. In other words, if $T_1$ and $T_2$ are two distinct row-standard $n$-tableaux of shape $Y(\lambda)$, then the $n$-tabloids $\overline{T_1}$ and $\overline{T_2}$ are also distinct. Applying this to $T_1 = P$ and $T_2 = Q$, we conclude that the $n$-tabloids $\overline{P}$ and $\overline{Q}$ are distinct (since the $n$-tableaux $P$ and $Q$ are row-standard and distinct). Since the Young last letter order on $\{n\text{-tabloids of shape } Y(\lambda)\}$ is a total order, we thus conclude that either $\overline{P} < \overline{Q}$ or $\overline{Q} < \overline{P}$. We WLOG assume that $\overline{P} < \overline{Q}$ holds, since in the other case ($\overline{Q} < \overline{P}$) we can achieve this by swapping $P$ with $Q$. Thus, we have $\lambda = \mu$ and $\overline{P} < \overline{Q}$. Thus,

$$\left(\lambda < \mu \text{ or } \left(\lambda = \mu \text{ and } \overline{P} < \overline{Q}\right)\right).$$

But this is equivalent to saying that $P < Q$ (by the definition of the relation $<$ on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$, since $P \in \operatorname{SYT}(\lambda)$ and $Q \in \operatorname{SYT}(\mu)$). Thus, we have $P < Q$, and therefore ($P < Q$ or $Q < P$). Thus, Claim 4 is proved in Case 2.

It remains to handle Case 3. In this case, we have $\lambda > \mu$. This is equivalent to $\mu < \lambda$. Thus, Case 3 is the same case as Case 1, except that the roles of $P$ and $Q$ (and thus the roles of $\lambda$ and $\mu$) are switched. Hence, the proof of Claim 4 in Case 3 is analogous to the proof in Case 1 (since Claim 4 is symmetric in $P$ and $Q$).

We have now proved Claim 4 in all three Cases 1, 2 and 3; thus, its proof is complete. □

Claim 4 shows that the partial order whose smaller relation is $<$ is actually a total order. The proof of Proposition 6.7.6 is thus complete. $\square$

*Proof of Proposition 6.7.11.* First, we shall show that the relation $<$ is the smaller relation of a partial order on $\operatorname{SBT}(n)$. In order to prove this, we need to verify three claims:[191]

> *Claim 1:* The relation $<$ is irreflexive (i.e., there exists no $\mathfrak{P} \in \operatorname{SBT}(n)$ such that $\mathfrak{P} < \mathfrak{P}$).
>
> *Claim 2:* The relation $<$ is transitive (i.e., if three $n$-bitableaux $\mathfrak{P}, \mathfrak{Q}, \mathfrak{R} \in \operatorname{SBT}(n)$ satisfy $\mathfrak{P} < \mathfrak{Q}$ and $\mathfrak{Q} < \mathfrak{R}$, then $\mathfrak{P} < \mathfrak{R}$).
>
> *Claim 3:* The relation $<$ is asymmetric (i.e., no two $n$-bitableaux $\mathfrak{P}$ and $\mathfrak{Q}$ satisfy $\mathfrak{P} < \mathfrak{Q}$ and $\mathfrak{Q} < \mathfrak{P}$ at the same time).

*Proof of Claim 1.* Let $\mathfrak{P} \in \operatorname{SBT}(n)$ be an $n$-bitableau such that $\mathfrak{P} < \mathfrak{P}$. We shall derive a contradiction.

Indeed, let us write the $n$-bitableau $\mathfrak{P}$ as $\mathfrak{P} = (\lambda, U, V)$. Now, we have assumed that $\mathfrak{P} < \mathfrak{P}$. In other words, $(\lambda, U, V) < (\lambda, U, V)$ (since $\mathfrak{P} = (\lambda, U, V)$). By the definition of the relation $<$, this means that

$$(U < U \text{ or } (U = U \text{ and } V < V)).$$

But this is clearly false (since neither $U < U$ nor $V < V$ is true (because the relation $<$ on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ is a total order)). Thus, we have found a contradiction.

Forget that we fixed $\mathfrak{P}$. We thus have obtained a contradiction for each $n$-bitableau $\mathfrak{P}$ such that $\mathfrak{P} < \mathfrak{P}$. Hence, there exists no such $\mathfrak{P}$. In other words, the relation $<$ is irreflexive. This proves Claim 1. $\square$

*Proof of Claim 2.* Let $\mathfrak{P}, \mathfrak{Q}, \mathfrak{R} \in \operatorname{SBT}(n)$ be three $n$-bitableaux that satisfy $\mathfrak{P} < \mathfrak{Q}$ and $\mathfrak{Q} < \mathfrak{R}$. We must prove that $\mathfrak{P} < \mathfrak{R}$.

Indeed, let us write the $n$-bitableaux $\mathfrak{P}, \mathfrak{Q}, \mathfrak{R}$ as

$$\mathfrak{P} = (\lambda, A, B), \qquad \mathfrak{Q} = (\mu, C, D), \qquad \mathfrak{R} = (\nu, E, F),$$

respectively. Now, recall that $\mathfrak{P} < \mathfrak{Q}$. In other words, $(\lambda, A, B) < (\mu, C, D)$ (since $\mathfrak{P} = (\lambda, A, B)$ and $\mathfrak{Q} = (\mu, C, D)$). By the definition of the relation $<$, this means that

$$(A < C \text{ or } (A = C \text{ and } B < D)).$$

So we must have $A < C$ or $(A = C$ and $B < D)$. Likewise, from $\mathfrak{Q} < \mathfrak{R}$, we conclude that we must have $C < E$ or $(C = E$ and $D < F)$.

We just have showed that we have $A < C$ or $(A = C$ and $B < D)$. In both of these cases, we have $A \leq C$ (since $A \leq C$ follows both from $A < C$ and from $A = C$).

So we have derived $A \leq C$ from $\mathfrak{P} < \mathfrak{Q}$. Likewise, from $\mathfrak{Q} < \mathfrak{R}$, we can derive $C \leq E$. But the relation $\leq$ on the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ is transitive (since it belongs to a total order). Hence, from $A \leq C$ and $C \leq E$, we obtain $A \leq E$. In other words, we have either $A = E$ or $A < E$. Hence, we are in one of the following two cases:

[191] We are using capital Fraktur letters like $\mathfrak{P}, \mathfrak{Q}, \mathfrak{R}$ for standard $n$-bitableaux.

*Case 1:* We have $A = E$.

*Case 2:* We have $A < E$.

Let us first consider Case 1. In this case, we have $A = E$. Thus, $E = A \leq C$. Combining this with $C \leq E$, we obtain $C = E$ (since the relation $\leq$ on $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ belongs to a total order and thus is antisymmetric). Hence, $C = E = A$, so that $A = C$.

Now, recall that we have either $A < C$ or $(A = C \text{ and } B < D)$. Since $A < C$ is impossible (because $A = C$), we thus conclude that $(A = C \text{ and } B < D)$ must hold. Hence, $B < D$.

Recall that we showed that we must have $C < E$ or $(C = E \text{ and } D < F)$. Since $C < E$ is impossible (because $C = E$), we thus must have $(C = E \text{ and } D < F)$. Hence, $D < F$.

But the relation $<$ on the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ is a total order, thus transitive. Hence, from $B < D$ and $D < F$, we obtain $B < F$. Combining this with $A = E$, we obtain $(A = E \text{ and } B < F)$. Hence, we have

$$(A < E \text{ or } (A = E \text{ and } B < F)).$$

But this is equivalent to saying that $(\lambda, A, B) < (\nu, E, F)$ (by the definition of the relation $<$ on $\operatorname{SBT}(n)$). In other words, $\mathfrak{P} < \mathfrak{R}$ (since $\mathfrak{P} = (\lambda, A, B)$ and $\mathfrak{R} = (\nu, E, F)$). Thus, Claim 2 is proved in Case 1.

Let us now consider Case 2. In this case, we have $A < E$. Hence, we have

$$(A < E \text{ or } (A = E \text{ and } B < F)).$$

But this is equivalent to saying that $(\lambda, A, B) < (\nu, E, F)$ (by the definition of the relation $<$ on $\operatorname{SBT}(n)$). In other words, $\mathfrak{P} < \mathfrak{R}$ (since $\mathfrak{P} = (\lambda, A, B)$ and $\mathfrak{R} = (\nu, E, F)$). Thus, Claim 2 is proved in Case 2.

We have now proved Claim 2 in both Cases 1 and 2. Hence, the proof of Claim 2 is complete. □

*Proof of Claim 3.* Any binary relation that is irreflexive and transitive is automatically asymmetric (since $\mathfrak{P} < \mathfrak{Q}$ and $\mathfrak{Q} < \mathfrak{P}$ would entail $\mathfrak{P} < \mathfrak{P}$ by transitivity, but this would contradict irreflexivity). Since the relation $<$ is irreflexive (by Claim 1) and transitive (by Claim 2), we thus conclude that it is also asymmetric. This proves Claim 3. □

Having proved Claims 1, 2 and 3, we thus conclude that the relation $<$ is the smaller relation of a partial order. It remains to show that this partial order is a total order. For this purpose, we need to prove the following last claim:

> *Claim 4:* Let $\mathfrak{P}$ and $\mathfrak{Q}$ be two distinct $n$-bitableaux in $\operatorname{SBT}(n)$. Then, we have $\mathfrak{P} < \mathfrak{Q}$ or $\mathfrak{Q} < \mathfrak{P}$.

*Proof of Claim 4.* Let us write the $n$-bitableaux $\mathfrak{P}$ and $\mathfrak{Q}$ as

$$\mathfrak{P} = (\lambda, A, B) \qquad \text{and} \qquad \mathfrak{Q} = (\mu, C, D),$$

respectively. We are clearly in one of the following two cases:

*Case 1:* We have $A = C$.

*Case 2:* We have $A \neq C$.

Let us consider Case 1 first. In this case, we have $A = C$. But $A = C \in \operatorname{SYT}(\mu)$ (since $(\mu, C, D)$ is a standard $n$-bitableau) and $A \in \operatorname{SYT}(\lambda)$ (since $(\lambda, A, B)$ is a standard $n$-bitableau). Hence, the sets $\operatorname{SYT}(\lambda)$ and $\operatorname{SYT}(\mu)$ have at least one element in common (namely, $A$). Since the sets $\operatorname{SYT}(\kappa)$ for different partitions $\kappa$ are disjoint, we thus conclude that the partitions $\lambda$ and $\mu$ must be identical. In other words, $\lambda = \mu$.

If we also had $B = D$, then we would have $(\lambda, A, B) = (\mu, C, D)$ (since $\lambda = \mu$ and $A = C$ and $B = D$), so that $\mathfrak{P} = (\lambda, A, B) = (\mu, C, D) = \mathfrak{Q}$; but this would contradict the fact that $\mathfrak{P}$ and $\mathfrak{Q}$ are distinct (by assumption). Hence, we cannot have $B = D$. Thus, $B \neq D$.

But the relation $<$ on the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ is a total order. Hence, from $B \neq D$, we conclude that either $B < D$ or $D < B$. These two cases are completely analogous (indeed, we can switch between them by swapping $\mathfrak{P}$ with $\mathfrak{Q}$; note that this switch does not affect the equality $A = C$), so we can WLOG assume that we are in the first of them – i.e., that we have $B < D$. Now, we have

$$(A < C \text{ or } (A = C \text{ and } B < D))$$

(since we have $A = C$ and $B < D$). But this is equivalent to saying that $(\lambda, A, B) < (\mu, C, D)$ (by the definition of the relation $<$ on $\operatorname{SBT}(n)$). Thus, we have $(\lambda, A, B) < (\mu, C, D)$. In other words, $\mathfrak{P} < \mathfrak{Q}$ (since $\mathfrak{P} = (\lambda, A, B)$ and $\mathfrak{Q} = (\mu, C, D)$). Therefore, $(\mathfrak{P} < \mathfrak{Q}$ or $\mathfrak{Q} < \mathfrak{P})$. Thus, Claim 4 is proved in Case 1.

Now consider Case 2. In this case, we have $A \neq C$. But the relation $<$ on the set $\bigcup_{\kappa \vdash n} \operatorname{SYT}(\kappa)$ is a total order. Hence, from $A \neq C$, we conclude that either $A < C$ or $C < A$. These two cases are completely analogous (indeed, we can switch between them by swapping $\mathfrak{P}$ with $\mathfrak{Q}$), so we can WLOG assume that we are in the first of them – i.e., that we have $A < C$. Thus, we have

$$(A < C \text{ or } (A = C \text{ and } B < D)).$$

But this is equivalent to saying that $(\lambda, A, B) < (\mu, C, D)$ (by the definition of the relation $<$ on $\operatorname{SBT}(n)$). Thus, we have $(\lambda, A, B) < (\mu, C, D)$. In other words, $\mathfrak{P} < \mathfrak{Q}$ (since $\mathfrak{P} = (\lambda, A, B)$ and $\mathfrak{Q} = (\mu, C, D)$). Therefore, $(\mathfrak{P} < \mathfrak{Q}$ or $\mathfrak{Q} < \mathfrak{P})$. Thus, Claim 4 is proved in Case 2.

We have now proved Claim 4 in both Cases 1 and 2; thus, its proof is complete. □

Claim 4 shows that the partial order whose smaller relation is $<$ is actually a total order. The proof of Proposition 6.7.11 is thus complete. □

*Proof of Lemma 6.7.14.* We have $(\lambda, A, B) < (\mu, C, D)$. By the definition of the relation $<$, this means that

$$(A < C \text{ or } (A = C \text{ and } B < D)).$$

So we must have $A < C$ or $(A = C$ and $B < D)$. Thus, we must have $A < C$ or $A = C$ (since the statement "$(A = C$ and $B < D)$" implies the statement "$A = C$"). In other words, $A \leq C$. This proves Lemma 6.7.14 **(a)**.

**(b)** Recall that we must have $A < C$ or $(A = C$ and $B < D)$. Hence, we must have $A < C$ or $B < D$ (since the statement "$(A = C$ and $B < D)$" implies the statement "$B < D$"). This proves Lemma 6.7.14 **(b)**. □

## A.6. The Murphy bases

Here we supply detailed proofs for some claims made in Section 6.8.

*Proof of Lemma 6.8.6.* Both $U$ and $V$ are $n$-tableaux. Thus, all entries of $U$ are distinct, and all entries of $V$ are distinct.

**(a)** Assume that we have $\left(\operatorname{Row}\left(i,U\right)=\operatorname{Row}\left(i,V\right)\ \text{for all } i\in\mathbb{Z}\right)$.

Fix $i\in\mathbb{Z}$. Then, $\operatorname{Row}\left(i,U\right)=\operatorname{Row}\left(i,V\right)$ (by the assumption we just made). In other words, the set of entries in the $i$-th row of $U$ equals the set of entries in the $i$-th row of $V$ (since these two sets are precisely $\operatorname{Row}\left(i,U\right)$ and $\operatorname{Row}\left(i,V\right)$). We can furthermore replace the word "set" by "multiset" here (since the entries in the $i$-th row of $U$ are distinct (because all entries of $U$ are distinct), and so are the entries in the $i$-th row of $V$, and therefore the multisets of these entries do not have any nontrivial multiplicities). Thus, we conclude that the multiset of entries in the $i$-th row of $U$ equals the multiset of entries in the $i$-th row of $V$.

Forget that we fixed $i$. We thus have shown that for each $i\in\mathbb{Z}$, the multiset of entries in the $i$-th row of $U$ equals the multiset of entries in the $i$-th row of $V$. In other words, the $n$-tableaux $U$ and $V$ are row-equivalent (by the definition of "row-equivalent"). In other words, $\overline{U}=\overline{V}$ (since an $n$-tabloid is an equivalence class with respect to row-equivalence).

Forget our assumption that $\left(\operatorname{Row}\left(i,U\right)=\operatorname{Row}\left(i,V\right)\ \text{for all } i\in\mathbb{Z}\right)$. We thus have proved that $\overline{U}=\overline{V}$ if $\left(\operatorname{Row}\left(i,U\right)=\operatorname{Row}\left(i,V\right)\ \text{for all } i\in\mathbb{Z}\right)$. In other words, we have proved the implication

$$\left(\operatorname{Row}\left(i,U\right)=\operatorname{Row}\left(i,V\right)\ \text{for all } i\in\mathbb{Z}\right)\implies\left(\overline{U}=\overline{V}\right).$$

The same argument (but read in reverse) shows the converse implication

$$\left(\overline{U}=\overline{V}\right)\implies\left(\operatorname{Row}\left(i,U\right)=\operatorname{Row}\left(i,V\right)\ \text{for all } i\in\mathbb{Z}\right)$$

(indeed, this is even easier, since the equality of two multisets automatically implies the equality of the corresponding sets). Combining these two implications, we conclude the equivalence

$$\left(\overline{U}=\overline{V}\right)\iff\left(\operatorname{Row}\left(i,U\right)=\operatorname{Row}\left(i,V\right)\ \text{for all } i\in\mathbb{Z}\right).$$

This proves Lemma 6.8.6 **(a)**.

**(b)** This is entirely analogous to part **(a)**. □

*Detailed proof of Proposition 6.8.31.* Recall that we defined $\mathcal{F}^{\operatorname{Row}}_{\text{all},\geq T}$ as

$$\mathcal{F}^{\operatorname{Row}}_{\text{all},\geq T}=\operatorname{span}\left\{\nabla^{\operatorname{Row}}_{V,U}\ \mid\ \left(\lambda,U,V\right)\in\operatorname{HSBT}\left(n\right)\ \text{and}\ U\geq T\right\}.$$

In other words, $\mathcal{F}^{\operatorname{Row}}_{\text{all},\geq T}$ is the span of all the vectors of the form $\nabla^{\operatorname{Row}}_{V,U}$ with $\left(\lambda,U,V\right)\in\operatorname{HSBT}\left(n\right)$ satisfying $U\geq T$. We shall refer to such vectors as the "good nablas". Thus, $\mathcal{F}^{\operatorname{Row}}_{\text{all},\geq T}$ is the span of all the good nablas.

Let us now prove that $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$. For this, it suffices to show that $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is an $S_n$-subset of $\mathcal{A}$ (since $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is clearly a $\mathbf{k}$-submodule of $\mathcal{A}$). In other words, it suffices to show that $w\mathbf{a} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ for each $w \in S_n$ and each $\mathbf{a} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$.

So let us show this. Let $w \in S_n$ and $\mathbf{a} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$. We must prove that $w\mathbf{a} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$. Since this claim depends $\mathbf{k}$-linearly on $\mathbf{a}$ (because $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is a $\mathbf{k}$-submodule of $\mathcal{A}$), we can WLOG assume (by linearity) that $\mathbf{a}$ is one of the good nablas (since $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is the span of all the good nablas). In other words, we can WLOG assume that $\mathbf{a} = \nabla^{\mathrm{Row}}_{V,U}$ for some $(\lambda, U, V) \in \mathrm{HSBT}(n)$ satisfying $U \geq T$. Assume this, and consider this $(\lambda, U, V) \in \mathrm{HSBT}(n)$. Then, $(\lambda, U, V) \in \mathrm{HSBT}(n)$ shows that $(\lambda, U, V)$ is a half-standard $n$-bitableau. In other words, $\lambda$ is a partition of $n$ and $U$ and $V$ are two $n$-tableaux of shape $\lambda$ such that $U$ is standard. Hence, of course, $wV$ is an $n$-tableau of shape $\lambda$ as well. Hence, $(\lambda, U, wV)$ is again a half-standard $n$-bitableau, i.e., an element of $\mathrm{HSBT}(n)$. Therefore, $\nabla^{\mathrm{Row}}_{wV,U}$ is one of the good nablas as well (since $U \geq T$), and hence belongs to $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ (since $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is the span of all the good nablas). In other words, $\nabla^{\mathrm{Row}}_{wV,U} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$.

Now, (494) (applied to $P = V$ and $Q = U$ and $g = w$ and $h = \mathrm{id}$) yields $\nabla^{\mathrm{Row}}_{wV,\mathrm{id}\,U} = w\nabla^{\mathrm{Row}}_{V,U} \underbrace{\mathrm{id}^{-1}}_{=\mathrm{id}=1} = w\nabla^{\mathrm{Row}}_{V,U}$. In other words, $\nabla^{\mathrm{Row}}_{wV,U} = w\nabla^{\mathrm{Row}}_{V,U}$ (since $\mathrm{id}\,U = U$). Hence, $w\nabla^{\mathrm{Row}}_{V,U} = \nabla^{\mathrm{Row}}_{wV,U} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$. In other words, $w\mathbf{a} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ (since $\mathbf{a} = \nabla^{\mathrm{Row}}_{V,U}$).

Forget that we fixed $w$ and $\mathbf{a}$. We thus have shown that $w\mathbf{a} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ for each $w \in S_n$ and each $\mathbf{a} \in \mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$. In other words, $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is an $S_n$-subset of $\mathcal{A}$. Hence, $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is a subrepresentation of the $S_n$-representation $\mathcal{A}$ (since $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is a $\mathbf{k}$-submodule of $\mathcal{A}$). In other words, $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is a left $\mathbf{k}[S_n]$-submodule of $\mathcal{A}$ (since Proposition 4.3.17 shows that the subrepresentations of the $S_n$-representation $\mathcal{A}$ are precisely the left $\mathbf{k}[S_n]$-submodules of $\mathcal{A}$). In other words, $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},\geq T}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$ (since $\mathbf{k}[S_n] = \mathcal{A}$). Likewise, we can show that $\mathcal{F}^{\mathrm{Row}}_{\mathrm{all},>T}$, $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},\leq T}$ and $\mathcal{F}^{-\mathrm{Col}}_{\mathrm{all},<T}$ are left $\mathcal{A}$-submodules of $\mathcal{A}$. □

*Proof of Lemma 6.8.36, parts **(c)** and **(d)**.* We shall now prove parts **(c)** and **(d)** of Lemma 6.8.36. (Parts **(a)** and **(b)** were already proved in Section 6.8.)

**(c)** We must show that $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$ for each $\mathbf{a} \in \left(\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}\right)^{\perp}$.

So let $\mathbf{a} \in \left(\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}\right)^{\perp}$. We must show that $\mathbf{a} \in \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T}$.

We have $\mathbf{a} \in \left(\mathcal{F}^{\mathrm{Row}}_{\mathrm{std},\geq T}\right)^{\perp} \subseteq \mathcal{A}$. Since the column Murphy basis $\left(\nabla^{-\mathrm{Col}}_{V,U}\right)_{(\lambda,U,V)\in\mathrm{SBT}(n)}$ is a basis of $\mathcal{A}$ (by Theorem 6.8.14), we can thus write $\mathbf{a}$ as a $\mathbf{k}$-linear combination of its elements. In other words, we can write $\mathbf{a}$ as

$$\mathbf{a} = \sum_{(\lambda,U,V)\in\mathrm{SBT}(n)} \omega_{\lambda,U,V} \nabla^{-\mathrm{Col}}_{V,U} \tag{558}$$

for some scalars $\omega_{\lambda,U,V} \in \mathbf{k}$. Consider these $\omega_{\lambda,U,V}$. We can rewrite the equality (527) as

$$\mathbf{a} = \sum_{(\mu,P,Q)\in\mathrm{SBT}(n)} \omega_{\mu,P,Q} \nabla^{-\mathrm{Col}}_{Q,P} \tag{559}$$

(by renaming the summation index $(\lambda, U, V)$ as $(\mu, P, Q)$).

Now, we shall show the following:

*Claim 2:* We have

$$\omega_{\lambda,U,V} = 0 \qquad \text{for all } (\lambda, U, V) \in \operatorname{SBT}(n) \text{ satisfying } U \geq T.$$

*Proof of Claim 2.* We proceed by strong induction on $(\lambda, U, V)$ using the **reverse** of the total order on $\operatorname{SBT}(n)$ introduced in Definition 6.7.12. So we fix some $(\lambda, U, V) \in \operatorname{SBT}(n)$ satisfying $U \geq T$. We assume (as the induction hypothesis) that

$$\omega_{\mu,P,Q} = 0 \tag{560}$$

for all $(\mu, P, Q) \in \operatorname{SBT}(n)$ satisfying $P \geq T$ and $(\mu, P, Q) > (\lambda, U, V)$. Our goal is to prove that $\omega_{\lambda,U,V} = 0$.

Since $(\lambda, U, V) \in \operatorname{SBT}(n)$ and $U \geq T$, we have $\nabla^{\operatorname{Row}}_{V,U} \in \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}$ by the definition of $\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}$ (in fact, $\nabla^{\operatorname{Row}}_{V,U}$ is one of the spanning vectors in the definition of $\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}$).

Since $\mathbf{a} \in \left(\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}\right)^{\perp}$, we have $\langle \mathbf{a}, \mathbf{v} \rangle = 0$ for each $\mathbf{v} \in \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}$ (by the definition of $\left(\mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}\right)^{\perp}$). Applying this to $\mathbf{v} = \nabla^{\operatorname{Row}}_{V,U}$, we obtain $\left\langle \mathbf{a},\ \nabla^{\operatorname{Row}}_{V,U} \right\rangle = 0$ (since $\nabla^{\operatorname{Row}}_{V,U} \in \mathcal{F}^{\operatorname{Row}}_{\operatorname{std},\geq T}$). Thus,

$$\begin{aligned} 0 &= \left\langle \mathbf{a},\ \nabla^{\operatorname{Row}}_{V,U} \right\rangle = \left\langle \sum_{(\mu,P,Q)\in\operatorname{SBT}(n)} \omega_{\mu,P,Q} \nabla^{-\operatorname{Col}}_{Q,P},\ \nabla^{\operatorname{Row}}_{V,U} \right\rangle \qquad \text{(by (559))} \\ &= \sum_{(\mu,P,Q)\in\operatorname{SBT}(n)} \omega_{\mu,P,Q} \left\langle \nabla^{-\operatorname{Col}}_{Q,P},\ \nabla^{\operatorname{Row}}_{V,U} \right\rangle \end{aligned} \tag{561}$$

(since the form $\langle \cdot, \cdot \rangle$ is bilinear). Now, we observe the following:

- For each $(\mu, P, Q) \in \operatorname{SBT}(n)$ satisfying $(\mu, P, Q) > (\lambda, U, V)$, we have

$$\omega_{\mu,P,Q} = 0. \tag{562}$$

[*Proof:* Let $(\mu, P, Q) \in \operatorname{SBT}(n)$ be such that $(\mu, P, Q) > (\lambda, U, V)$. Thus, $(\lambda, U, V) < (\mu, P, Q)$. Hence, Lemma 6.7.14 **(a)** (applied to $(\lambda, U, V)$ and $(\mu, P, Q)$ instead of $(\lambda, A, B)$ and $(\mu, C, D)$) shows that $U \leq P$. Thus, $P \geq U \geq T$, so that $P \geq T$ (since the relation $<$ is a total order). Thus, (560) yields $\omega_{\mu,P,Q} = 0$. This proves (562).]

- For each $(\mu, P, Q) \in \operatorname{SBT}(n)$ satisfying $(\mu, P, Q) < (\lambda, U, V)$, we have

$$\left\langle \nabla^{-\operatorname{Col}}_{Q,P},\ \nabla^{\operatorname{Row}}_{V,U} \right\rangle = 0. \tag{563}$$

[*Proof:* Let $(\mu, P, Q) \in \operatorname{SBT}(n)$ be such that $(\mu, P, Q) < (\lambda, U, V)$. Thus, Lemma 6.7.14 **(b)** (applied to $(\mu, P, Q)$ and $(\lambda, U, V)$ instead of $(\lambda, A, B)$ and $(\mu, C, D)$) yields $P < U$ or $Q < V$. Thus, Lemma 6.8.23 (applied to $A = P$, $B = Q$, $C = U$ and $D = V$) yields $\left\langle \nabla^{-\operatorname{Col}}_{Q,P},\ \nabla^{\operatorname{Row}}_{V,U} \right\rangle = 0$. This proves (563).]

Now, recall that the relation $<$ on $\operatorname{SBT}(n)$ is the smaller relation of a total order. Hence, each $(\mu, P, Q) \in \operatorname{SBT}(n)$ satisfies exactly one of the three statements $(\mu, P, Q) < (\lambda, U, V)$ and $(\mu, P, Q) > (\lambda, U, V)$ and $(\mu, P, Q) = (\lambda, U, V)$.

Accordingly, the sum on the right hand side of (561) can be split into three parts: the part comprising all $(\mu, P, Q) \in \operatorname{SBT}(n)$ that satisfy $(\mu, P, Q) < (\lambda, U, V)$; the part comprising all $(\mu, P, Q) \in \operatorname{SBT}(n)$ that satisfy $(\mu, P, Q) > (\lambda, U, V)$; and a single addend corresponding to $(\mu, P, Q) = (\lambda, U, V)$. Splitting the sum in this way, we rewrite (561) as follows:

$$\begin{aligned}
0 &= \sum_{\substack{(\mu,P,Q)\in\operatorname{SBT}(n);\\ (\mu,P,Q)<(\lambda,U,V)}} \omega_{\mu,P,Q} \underbrace{\left\langle \nabla^{-\mathrm{Col}}_{Q,P},\ \nabla^{\mathrm{Row}}_{V,U} \right\rangle}_{\substack{=0\\ \text{(by (563))}}} \\
&\qquad + \sum_{\substack{(\mu,P,Q)\in\operatorname{SBT}(n);\\ (\mu,P,Q)>(\lambda,U,V)}} \underbrace{\omega_{\mu,P,Q}}_{\substack{=0\\ \text{(by (562))}}} \left\langle \nabla^{-\mathrm{Col}}_{Q,P},\ \nabla^{\mathrm{Row}}_{V,U} \right\rangle + \omega_{\lambda,U,V} \left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{V,U} \right\rangle \\
&= \underbrace{\sum_{\substack{(\mu,P,Q)\in\operatorname{SBT}(n);\\ (\mu,P,Q)<(\lambda,U,V)}} \omega_{\mu,P,Q} 0}_{=0} + \underbrace{\sum_{\substack{(\mu,P,Q)\in\operatorname{SBT}(n);\\ (\mu,P,Q)>(\lambda,U,V)}} 0 \left\langle \nabla^{-\mathrm{Col}}_{Q,P},\ \nabla^{\mathrm{Row}}_{V,U} \right\rangle}_{=0} + \omega_{\lambda,U,V} \left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{V,U} \right\rangle \\
&= \omega_{\lambda,U,V} \underbrace{\left\langle \nabla^{-\mathrm{Col}}_{V,U},\ \nabla^{\mathrm{Row}}_{V,U} \right\rangle}_{\substack{=\pm 1\\ \text{(by Lemma 6.8.22,}\\ \text{applied to } A=U \text{ and } B=V)}} = \pm\omega_{\lambda,U,V}.
\end{aligned}$$

Hence, $\pm\omega_{\lambda,U,V} = 0$. In other words, $\omega_{\lambda,U,V} = 0$. This completes the induction, and thus the proof of Claim 2. $\square$

Now, (558) becomes

$$\begin{aligned}
\mathbf{a} &= \sum_{(\lambda,U,V)\in\operatorname{SBT}(n)} \omega_{\lambda,U,V} \nabla^{-\mathrm{Col}}_{V,U} \\
&= \sum_{\substack{(\lambda,U,V)\in\operatorname{SBT}(n);\\ U\geq T}} \underbrace{\omega_{\lambda,U,V}}_{\substack{=0\\ \text{(by Claim 2)}}} \nabla^{-\mathrm{Col}}_{V,U} + \sum_{\substack{(\lambda,U,V)\in\operatorname{SBT}(n);\\ U<T}} \omega_{\lambda,U,V} \nabla^{-\mathrm{Col}}_{V,U} \\
&\qquad\qquad \left( \begin{array}{l} \text{since each } (\lambda, U, V) \in \operatorname{SBT}(n) \text{ satisfies} \\ \text{either } U \geq T \text{ or } U \leq < T \text{ (but not both)} \end{array} \right) \\
&= \underbrace{\sum_{\substack{(\lambda,U,V)\in\operatorname{SBT}(n);\\ U\geq T}} 0 \nabla^{-\mathrm{Col}}_{V,U}}_{=0} + \sum_{\substack{(\lambda,U,V)\in\operatorname{SBT}(n);\\ U<T}} \omega_{\lambda,U,V} \nabla^{-\mathrm{Col}}_{V,U} \\
&= \sum_{\substack{(\lambda,U,V)\in\operatorname{SBT}(n);\\ U<T}} \omega_{\lambda,U,V} \nabla^{-\mathrm{Col}}_{V,U} \\
&\in \operatorname{span}\left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\lambda, U, V) \in \operatorname{SBT}(n) \text{ and } U < T \right\} \\
&= \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} \qquad \left( \text{by the definition of } \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} \right).
\end{aligned}$$

This completes our proof of Lemma 6.8.36 **(c)**.

**(d)** This is analogous to the above proof of Lemma 6.8.36 **(c)**; the only difference is that some inequality signs (between standard $n$-tableaux) change from strong to weak and vice versa. (For instance, the "$U \geq T$" in Claim 2 must be replaced by "$U > T$"). We leave it to the reader to make the necessary changes. □

*Detailed proof of Theorem 6.8.37.* As in the proof of Proposition 6.8.30, we can see that $m > 0$. Hence, $0 < m$.

**(a)** Let $i \in \{0, 1, \ldots, m\}$. We must prove that

$$\mathcal{F}_i^{\mathrm{Row}} = \begin{cases} \mathcal{F}_{\mathrm{std},>T_i}^{\mathrm{Row}}, & \text{if } i > 0; \\ \mathcal{A}, & \text{if } i = 0. \end{cases}$$

We have $i \in \{0, 1, \ldots, m\}$. Thus, we are in one of the following three cases:
*Case 1:* We have $i = 0$.
*Case 2:* We have $0 < i < m$.
*Case 3:* We have $i = m$.

First, let us consider Case 1. In this case, we have $i = 0$. Hence, $i + 1 = 1$ and $i = 0 < m$. The definition of $\mathcal{F}_i^{\mathrm{Row}}$ yields

$$\begin{aligned} \mathcal{F}_i^{\mathrm{Row}} &= \begin{cases} \mathcal{F}_{\mathrm{std},\geq T_{i+1}}^{\mathrm{Row}}, & \text{if } i < m; \\ 0, & \text{if } i = m \end{cases} = \mathcal{F}_{\mathrm{std},\geq T_{i+1}}^{\mathrm{Row}} \qquad (\text{since } i < m) \\ &= \mathcal{F}_{\mathrm{std},\geq T_1}^{\mathrm{Row}} \qquad (\text{since } i + 1 = 1) \\ &= \mathcal{A} \qquad (\text{by Proposition 6.8.30 } \textbf{(a)}) \\ &= \begin{cases} \mathcal{F}_{\mathrm{std},>T_i}^{\mathrm{Row}}, & \text{if } i > 0; \\ \mathcal{A}, & \text{if } i = 0 \end{cases} \qquad (\text{since } i = 0)\,. \end{aligned}$$

Thus, Theorem 6.8.37 **(a)** is proved in Case 1.

Let us next consider Case 2. In this case, we have $0 < i < m$. Thus, $i \in [m-1]$ and $i > 0$. The definition of $\mathcal{F}_i^{\mathrm{Row}}$ yields

$$\begin{aligned} \mathcal{F}_i^{\mathrm{Row}} &= \begin{cases} \mathcal{F}_{\mathrm{std},\geq T_{i+1}}^{\mathrm{Row}}, & \text{if } i < m; \\ 0, & \text{if } i = m \end{cases} = \mathcal{F}_{\mathrm{std},\geq T_{i+1}}^{\mathrm{Row}} \qquad (\text{since } i < m) \\ &= \mathcal{F}_{\mathrm{std},>T_i}^{\mathrm{Row}} \qquad (\text{by Proposition 6.8.30 } \textbf{(c)}) \\ &= \begin{cases} \mathcal{F}_{\mathrm{std},>T_i}^{\mathrm{Row}}, & \text{if } i > 0; \\ \mathcal{A}, & \text{if } i = 0 \end{cases} \qquad (\text{since } i > 0)\,. \end{aligned}$$

Thus, Theorem 6.8.37 **(a)** is proved in Case 2.

Finally, let us consider Case 3. In this case, we have $i = m$. Hence, $i = m > 0$. The definition of $\mathcal{F}_i^{\mathrm{Row}}$ yields

$$\mathcal{F}_i^{\mathrm{Row}} = \begin{cases} \mathcal{F}_{\mathrm{std},\geq T_{i+1}}^{\mathrm{Row}}, & \text{if } i < m; \\ 0, & \text{if } i = m \end{cases} = 0 \qquad (\text{since } i = m)\,.$$

Comparing this with

$$\begin{cases} \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T_i}, & \text{if } i>0; \\ \mathcal{A}, & \text{if } i=0 \end{cases} = \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T_i} \qquad (\text{since } i>0)$$

$$= \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T_m} \qquad (\text{since } i=m)$$

$$= 0 \qquad (\text{by Proposition 6.8.30 } \textbf{(b)}),$$

we obtain

$$\mathcal{F}^{\mathrm{Row}}_i = \begin{cases} \mathcal{F}^{\mathrm{Row}}_{\mathrm{std},>T_i}, & \text{if } i>0; \\ \mathcal{A}, & \text{if } i=0. \end{cases}$$

Thus, Theorem 6.8.37 **(a)** is proved in Case 3.

We have now proved Theorem 6.8.37 **(a)** in each of the three Cases 1, 2 and 3. Hence, Theorem 6.8.37 **(a)** always holds.

**(b)** Let $i \in \{0,1,\ldots,m\}$. We must prove that

$$\mathcal{F}^{-\mathrm{Col}}_i = \begin{cases} \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T_{i+1}}, & \text{if } i<m; \\ \mathcal{A}, & \text{if } i=m. \end{cases}$$

We have $i \in \{0,1,\ldots,m\}$. Thus, we are in one of the following three cases:

*Case 1:* We have $i=0$.

*Case 2:* We have $0<i<m$.

*Case 3:* We have $i=m$.

First, let us consider Case 1. In this case, we have $i=0$. Hence, $i+1=1$ and $i=0<m$. The definition of $\mathcal{F}^{-\mathrm{Col}}_i$ yields

$$\mathcal{F}^{-\mathrm{Col}}_i = \begin{cases} \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T_i}, & \text{if } i>0; \\ 0, & \text{if } i=0 \end{cases} = 0 \qquad (\text{since } i=0).$$

Comparing this with

$$\begin{cases} \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T_{i+1}}, & \text{if } i<m; \\ \mathcal{A}, & \text{if } i=m \end{cases} = \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T_{i+1}} \qquad (\text{since } i<m)$$

$$= \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T_1} \qquad (\text{since } i+1=1)$$

$$= 0 \qquad (\text{by Proposition 6.8.30 } \textbf{(b)}),$$

we obtain

$$\mathcal{F}^{-\mathrm{Col}}_i = \begin{cases} \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T_{i+1}}, & \text{if } i<m; \\ \mathcal{A}, & \text{if } i=m. \end{cases}$$

Thus, Theorem 6.8.37 **(b)** is proved in Case 1.

Let us next consider Case 2. In this case, we have $0 < i < m$. Thus, $i \in [m-1]$ and $i > 0$. The definition of $\mathcal{F}_i^{-\mathrm{Col}}$ yields

$$\begin{aligned}
\mathcal{F}_i^{-\mathrm{Col}} &= \begin{cases} \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}}, & \text{if } i > 0; \\ 0, & \text{if } i = 0 \end{cases} = \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}} \qquad (\text{since } i > 0) \\
&= \mathcal{F}_{\mathrm{std},<T_{i+1}}^{-\mathrm{Col}} \qquad (\text{by Proposition 6.8.30 } \textbf{(c)}) \\
&= \begin{cases} \mathcal{F}_{\mathrm{std},<T_{i+1}}^{-\mathrm{Col}}, & \text{if } i < m; \\ \mathcal{A}, & \text{if } i = m \end{cases} \qquad (\text{since } i < m).
\end{aligned}$$

Thus, Theorem 6.8.37 **(b)** is proved in Case 2.

Finally, let us consider Case 3. In this case, we have $i = m$. Hence, $i = m > 0$. The definition of $\mathcal{F}_i^{-\mathrm{Col}}$ yields

$$\begin{aligned}
\mathcal{F}_i^{-\mathrm{Col}} &= \begin{cases} \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}}, & \text{if } i > 0; \\ 0, & \text{if } i = 0 \end{cases} = \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}} \qquad (\text{since } i > 0) \\
&= \mathcal{F}_{\mathrm{std},\leq T_m}^{-\mathrm{Col}} \qquad (\text{since } i = m) \\
&= \mathcal{A} \qquad (\text{by Proposition 6.8.30 } \textbf{(a)}) \\
&= \begin{cases} \mathcal{F}_{\mathrm{std},<T_{i+1}}^{-\mathrm{Col}}, & \text{if } i < m; \\ \mathcal{A}, & \text{if } i = m \end{cases} \qquad (\text{since } i = m).
\end{aligned}$$

Thus, Theorem 6.8.37 **(b)** is proved in Case 3.

We have now proved Theorem 6.8.37 **(b)** in each of the three Cases 1, 2 and 3. Hence, Theorem 6.8.37 **(b)** always holds.

**(c)** Let $i \in \{0, 1, \ldots, m\}$. We must prove that the set $\mathcal{F}_i^{\mathrm{Row}}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$.

If $i = m$, then this is obvious (because in this case, $\mathcal{F}_i^{\mathrm{Row}}$ is defined to be 0, which is clearly a left $\mathcal{A}$-submodule of $\mathcal{A}$). Thus, we WLOG assume that $i \neq m$. Hence, $i < m$. Thus, the definition of $\mathcal{F}_i^{\mathrm{Row}}$ shows that $\mathcal{F}_i^{\mathrm{Row}} = \mathcal{F}_{\mathrm{std},\geq T_{i+1}}^{\mathrm{Row}}$. But Theorem 6.8.29 **(a)** (applied to $T = T_{i+1}$) shows that the set $\mathcal{F}_{\mathrm{std},\geq T_{i+1}}^{\mathrm{Row}}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$. In other words, $\mathcal{F}_i^{\mathrm{Row}}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$ (since $\mathcal{F}_i^{\mathrm{Row}} = \mathcal{F}_{\mathrm{std},\geq T_{i+1}}^{\mathrm{Row}}$). Thus, Theorem 6.8.37 **(c)** is proved.

**(d)** Let $i \in \{0, 1, \ldots, m\}$. We must prove that the set $\mathcal{F}_i^{-\mathrm{Col}}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$.

If $i = 0$, then this is obvious (because in this case, $\mathcal{F}_i^{-\mathrm{Col}}$ is defined to be 0, which is clearly a left $\mathcal{A}$-submodule of $\mathcal{A}$). Thus, we WLOG assume that $i \neq 0$. Hence, $i > 0$. Thus, the definition of $\mathcal{F}_i^{-\mathrm{Col}}$ shows that $\mathcal{F}_i^{-\mathrm{Col}} = \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}}$. But Theorem 6.8.29 **(d)** (applied to $T = T_i$) shows that the set $\mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$. In other words, $\mathcal{F}_i^{-\mathrm{Col}}$ is a left $\mathcal{A}$-submodule of $\mathcal{A}$ (since $\mathcal{F}_i^{-\mathrm{Col}} = \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}}$). Thus, Theorem 6.8.37 **(d)** is proved.

**(e)** We already know that the sets $\mathcal{F}_i^{\mathrm{Row}}$ are left $\mathcal{A}$-submodules of $\mathcal{A}$ (by Theorem 6.8.37 **(c)**). Thus, it remains to prove that

$$0 = \mathcal{F}_m^{\mathrm{Row}} \subseteq \mathcal{F}_{m-1}^{\mathrm{Row}} \subseteq \cdots \subseteq \mathcal{F}_2^{\mathrm{Row}} \subseteq \mathcal{F}_1^{\mathrm{Row}} \subseteq \mathcal{F}_0^{\mathrm{Row}} = \mathcal{A}.$$

In other words, it remains to prove that $\mathcal{F}_m^{\mathrm{Row}} = 0$ and $\mathcal{F}_0^{\mathrm{Row}} = \mathcal{A}$ and that each $i \in [m]$ satisfies $\mathcal{F}_i^{\mathrm{Row}} \subseteq \mathcal{F}_{i-1}^{\mathrm{Row}}$.

But $\mathcal{F}_m^{\mathrm{Row}} = 0$ is obvious (an immediate consequence of the definition of $\mathcal{F}_m^{\mathrm{Row}}$), and $\mathcal{F}_0^{\mathrm{Row}} = \mathcal{A}$ is also easy to check (since Theorem 6.8.37 **(a)** (applied to $i = 0$) shows that $\mathcal{F}_0^{\mathrm{Row}} = \begin{cases} \mathcal{F}_{\mathrm{std},>T_0}^{\mathrm{Row}}, & \text{if } 0 > 0; \\ \mathcal{A}, & \text{if } 0 = 0 \end{cases} = \mathcal{A}$ (since $0 = 0$)). Hence, it remains to prove that each $i \in [m]$ satisfies $\mathcal{F}_i^{\mathrm{Row}} \subseteq \mathcal{F}_{i-1}^{\mathrm{Row}}$.

So let $i \in [m]$. We must show that $\mathcal{F}_i^{\mathrm{Row}} \subseteq \mathcal{F}_{i-1}^{\mathrm{Row}}$.

We have $i \in [m]$, so that $i > 0$ and $i \leq m$ and therefore $i - 1 < i \leq m$. The definition of $\mathcal{F}_{i-1}^{\mathrm{Row}}$ yields

$$\mathcal{F}_{i-1}^{\mathrm{Row}} = \begin{cases} \mathcal{F}_{\mathrm{std},\geq T_{(i-1)+1}}^{\mathrm{Row}}, & \text{if } i-1 < m; \\ 0, & \text{if } i-1 = m \end{cases} = \mathcal{F}_{\mathrm{std},\geq T_{(i-1)+1}}^{\mathrm{Row}} \qquad (\text{since } i-1 < m)$$

$$= \mathcal{F}_{\mathrm{std},\geq T_i}^{\mathrm{Row}} \qquad (\text{since } (i-1)+1 = i).$$

On the other hand, Theorem 6.8.37 **(a)** yields

$$\mathcal{F}_i^{\mathrm{Row}} = \begin{cases} \mathcal{F}_{\mathrm{std},>T_i}^{\mathrm{Row}}, & \text{if } i > 0; \\ \mathcal{A}, & \text{if } i = 0 \end{cases} = \mathcal{F}_{\mathrm{std},>T_i}^{\mathrm{Row}} \qquad (\text{since } i > 0).$$

But Proposition 6.8.32 (applied to $T = T_i$) yields

$$\mathcal{F}_{\mathrm{std},>T_i}^{\mathrm{Row}} \subseteq \mathcal{F}_{\mathrm{std},\geq T_i}^{\mathrm{Row}} \qquad \text{and} \qquad \mathcal{F}_{\mathrm{std},<T_i}^{-\mathrm{Col}} \subseteq \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}}.$$

Hence, $\mathcal{F}_i^{\mathrm{Row}} = \mathcal{F}_{\mathrm{std},>T_i}^{\mathrm{Row}} \subseteq \mathcal{F}_{\mathrm{std},\geq T_i}^{\mathrm{Row}} = \mathcal{F}_{i-1}^{\mathrm{Row}}$ (since $\mathcal{F}_{i-1}^{\mathrm{Row}} = \mathcal{F}_{\mathrm{std},\geq T_i}^{\mathrm{Row}}$).

Thus we have shown that each $i \in [m]$ satisfies $\mathcal{F}_i^{\mathrm{Row}} \subseteq \mathcal{F}_{i-1}^{\mathrm{Row}}$. As we said, this completes the proof of Theorem 6.8.37 **(e)**.

**(f)** We already know that the sets $\mathcal{F}_i^{-\mathrm{Col}}$ are left $\mathcal{A}$-submodules of $\mathcal{A}$ (by Theorem 6.8.37 **(d)**). Thus, it remains to prove that

$$0 = \mathcal{F}_0^{-\mathrm{Col}} \subseteq \mathcal{F}_1^{-\mathrm{Col}} \subseteq \mathcal{F}_2^{-\mathrm{Col}} \subseteq \cdots \subseteq \mathcal{F}_{m-1}^{-\mathrm{Col}} \subseteq \mathcal{F}_m^{-\mathrm{Col}} = \mathcal{A}.$$

In other words, it remains to prove that $\mathcal{F}_0^{-\mathrm{Col}} = 0$ and $\mathcal{F}_m^{-\mathrm{Col}} = \mathcal{A}$ and that each $i \in [m]$ satisfies $\mathcal{F}_{i-1}^{-\mathrm{Col}} \subseteq \mathcal{F}_i^{-\mathrm{Col}}$.

But $\mathcal{F}_0^{-\mathrm{Col}} = 0$ is obvious (an immediate consequence of the definition of $\mathcal{F}_0^{-\mathrm{Col}}$), and $\mathcal{F}_m^{-\mathrm{Col}} = \mathcal{A}$ is also easy to check (since Theorem 6.8.37 **(b)** (applied to $i = m$) shows

that $\mathcal{F}_m^{-\mathrm{Col}} = \begin{cases} \mathcal{F}_{\mathrm{std},<T_{m+1}}^{-\mathrm{Col}}, & \text{if } m < m; \\ \mathcal{A}, & \text{if } m = m \end{cases} = \mathcal{A}$ (since $m = m$)). Hence, it remains to prove that each $i \in [m]$ satisfies $\mathcal{F}_{i-1}^{-\mathrm{Col}} \subseteq \mathcal{F}_i^{-\mathrm{Col}}$.

So let $i \in [m]$. We must show that $\mathcal{F}_{i-1}^{-\mathrm{Col}} \subseteq \mathcal{F}_i^{-\mathrm{Col}}$.

We have $i \in [m]$, so that $i > 0$ and $i \leq m$ and therefore $i - 1 < i \leq m$. Theorem 6.8.37 **(b)** (applied to $i-1$ instead of $i$) yields

$$\mathcal{F}_{i-1}^{-\mathrm{Col}} = \begin{cases} \mathcal{F}_{\mathrm{std},<T_{(i-1)+1}}^{-\mathrm{Col}}, & \text{if } i-1 < m; \\ \mathcal{A}, & \text{if } i-1 = m \end{cases} = \mathcal{F}_{\mathrm{std},<T_{(i-1)+1}}^{-\mathrm{Col}} \qquad (\text{since } i-1 < m)$$
$$= \mathcal{F}_{\mathrm{std},<T_i}^{-\mathrm{Col}} \qquad (\text{since } (i-1)+1 = i).$$

On the other hand, the definition of $\mathcal{F}_i^{-\mathrm{Col}}$ yields

$$\mathcal{F}_i^{-\mathrm{Col}} = \begin{cases} \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}}, & \text{if } i > 0; \\ 0, & \text{if } i = 0 \end{cases} = \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}} \qquad (\text{since } i > 0).$$

But Proposition 6.8.32 (applied to $T = T_i$) yields

$$\mathcal{F}_{\mathrm{std},>T_i}^{\mathrm{Row}} \subseteq \mathcal{F}_{\mathrm{std},\geq T_i}^{\mathrm{Row}} \qquad \text{and} \qquad \mathcal{F}_{\mathrm{std},<T_i}^{-\mathrm{Col}} \subseteq \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}}.$$

Hence, $\mathcal{F}_{i-1}^{-\mathrm{Col}} = \mathcal{F}_{\mathrm{std},<T_i}^{-\mathrm{Col}} \subseteq \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}} = \mathcal{F}_i^{-\mathrm{Col}}$ (since $\mathcal{F}_i^{-\mathrm{Col}} = \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}}$).

Thus we have shown that each $i \in [m]$ satisfies $\mathcal{F}_{i-1}^{-\mathrm{Col}} \subseteq \mathcal{F}_i^{-\mathrm{Col}}$. As we said, this completes the proof of Theorem 6.8.37 **(f)**. □

*Detailed proof of Lemma 6.8.40 **(e)**.* We observe that $T \in \mathrm{SYT}(\lambda)$ (since $T$ is a standard $n$-tableau of shape $\lambda$).

As we saw in the proof of Lemma 6.8.40 **(c)**, we have

$$\mathcal{F}_{\mathrm{std},<T}^{-\mathrm{Col}} = \mathrm{span}\left\{ \nabla_{V,U}^{-\mathrm{Col}} \mid (\mu, U, V) \in \mathrm{SBT}(n) \text{ and } U < T \right\}.$$

As we saw in the proof of Lemma 6.8.40 **(d)**, we have

$$\begin{aligned}
\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T} &= \operatorname{span} \underbrace{\left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U \leq T \right\}}_{\substack{= \left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U<T \text{ or } U=T \right\} \\ \text{(since the condition ``}U \leq T\text{'' is equivalent to ``}U<T \text{ or } U=T\text{'')}}} \\
&= \operatorname{span} \underbrace{\left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U < T \text{ or } U = T \right\}}_{\substack{= \left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U<T \right\} \\ \cup \left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U=T \right\}}} \\
&= \operatorname{span} \left( \left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U < T \right\} \right. \\
&\qquad \left. \cup \left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U = T \right\} \right) \\
&= \operatorname{span} \left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U < T \right\} \\
&\qquad + \operatorname{span} \left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U = T \right\}
\end{aligned}$$

(since any two subsets $X$ and $Y$ of a $\mathbf{k}$-module satisfy $\operatorname{span}(X \cup Y) = \operatorname{span} X + \operatorname{span} Y$). In view of

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} = \operatorname{span} \left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U < T \right\},$$

we can rewrite this as

$$\mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},\leq T} = \mathcal{F}^{-\mathrm{Col}}_{\mathrm{std},<T} + \operatorname{span} \left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U = T \right\}.$$

However, we have the two inclusions

$$\left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U = T \right\} \subseteq \left\{ \nabla^{-\mathrm{Col}}_{W,T} \mid W \in \mathrm{SYT}(\lambda) \right\}$$

[192] and

$$\left\{ \nabla^{-\mathrm{Col}}_{W,T} \mid W \in \mathrm{SYT}(\lambda) \right\} \subseteq \left\{ \nabla^{-\mathrm{Col}}_{V,U} \mid (\mu,U,V) \in \mathrm{SBT}(n) \text{ and } U = T \right\}$$

[192] *Proof.* This inclusion is just saying that each element of the form $\nabla^{-\mathrm{Col}}_{V,U}$ with $(\mu,U,V) \in \mathrm{SBT}(n)$ satisfying $U = T$ can also be written in the form $\nabla^{-\mathrm{Col}}_{W,T}$ for some $W \in \mathrm{SYT}(\lambda)$. So let us prove this.

Let $(\mu,U,V) \in \mathrm{SBT}(n)$ be such that $U = T$. We must show that $\nabla^{-\mathrm{Col}}_{V,U}$ can be written in the form $\nabla^{-\mathrm{Col}}_{W,T}$ for some $W \in \mathrm{SYT}(\lambda)$.

We have $U = T \in \mathrm{SYT}(\lambda)$ and $U \in \mathrm{SYT}(\mu)$ (since $(\mu,U,V) \in \mathrm{SBT}(n)$). Hence, the sets $\mathrm{SYT}(\lambda)$ and $\mathrm{SYT}(\mu)$ have at least one element in common (namely, $U$). Thus, $\lambda = \mu$ (since the sets $\mathrm{SYT}(\kappa)$ for different partitions $\kappa$ are disjoint), so that $\mu = \lambda$. However, from $(\mu,U,V) \in \mathrm{SBT}(n)$, we also obtain $V \in \mathrm{SYT}(\mu) = \mathrm{SYT}(\lambda)$ (since $\mu = \lambda$).

Now we know that $V \in \mathrm{SYT}(\lambda)$ and $\nabla^{-\mathrm{Col}}_{V,U} = \nabla^{-\mathrm{Col}}_{V,T}$ (since $U = T$). Therefore, $\nabla^{-\mathrm{Col}}_{V,U}$ can be written in the form $\nabla^{-\mathrm{Col}}_{W,T}$ for some $W \in \mathrm{SYT}(\lambda)$ (namely, for $W = V$). This is precisely what we wanted to prove.

[193]. Combining these two inclusions, we obtain the equality

$$\begin{aligned}\left\{\nabla_{V,U}^{-\mathrm{Col}} \mid (\mu,U,V)\in \mathrm{SBT}(n) \text{ and } U=T\right\} &= \left\{\nabla_{W,T}^{-\mathrm{Col}} \mid W\in \mathrm{SYT}(\lambda)\right\}\\ &= \left\{\nabla_{V,T}^{-\mathrm{Col}} \mid V\in \mathrm{SYT}(\lambda)\right\}\end{aligned}$$

(here, we have renamed the index $W$ as $V$).

Now, recall that

$$\begin{aligned}\mathcal{F}_{\mathrm{std},\leq T}^{-\mathrm{Col}} &= \mathcal{F}_{\mathrm{std},<T}^{-\mathrm{Col}} + \mathrm{span}\underbrace{\left\{\nabla_{V,U}^{-\mathrm{Col}} \mid (\mu,U,V)\in \mathrm{SBT}(n) \text{ and } U=T\right\}}_{=\left\{\nabla_{V,T}^{-\mathrm{Col}} \mid V\in \mathrm{SYT}(\lambda)\right\}}\\ &= \mathcal{F}_{\mathrm{std},<T}^{-\mathrm{Col}} + \mathrm{span}\left\{\nabla_{V,T}^{-\mathrm{Col}} \mid V\in \mathrm{SYT}(\lambda)\right\}.\end{aligned}$$

This proves Lemma 6.8.40 **(e)**. □

*Detailed proof of Theorem 6.8.41.* Set $T := T_i$ and $\lambda := \lambda^{(i)}$. Thus, $T \in \mathrm{SYT}(\lambda)$ (since $T_i \in \mathrm{SYT}\left(\lambda^{(i)}\right)$). In other words, $T$ is a standard $n$-tableau of shape $\lambda$. Hence, Theorem 6.8.39 yields that $\mathcal{F}_{\mathrm{std},\leq T}^{-\mathrm{Col}}/\mathcal{F}_{\mathrm{std},<T}^{-\mathrm{Col}} \cong \mathcal{A}\mathbf{E}_T \cong \mathcal{S}^\lambda$ as $S_n$-representations, i.e., as left $\mathcal{A}$-modules.

But the definition of $\mathcal{F}_i^{-\mathrm{Col}}$ yields

$$\begin{aligned}\mathcal{F}_i^{-\mathrm{Col}} &= \begin{cases}\mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}}, & \text{if } i>0;\\ 0, & \text{if } i=0\end{cases} = \mathcal{F}_{\mathrm{std},\leq T_i}^{-\mathrm{Col}} \qquad (\text{since } i>0)\\ &= \mathcal{F}_{\mathrm{std},\leq T}^{-\mathrm{Col}} \qquad (\text{since } T_i = T).\end{aligned}$$

Moreover, $i-1 < i \leq m$. Now, Theorem 6.8.37 **(b)** (applied to $i-1$ instead of $i$) yields

$$\begin{aligned}\mathcal{F}_{i-1}^{-\mathrm{Col}} &= \begin{cases}\mathcal{F}_{\mathrm{std},<T_{(i-1)+1}}^{-\mathrm{Col}}, & \text{if } i-1<m;\\ \mathcal{A}, & \text{if } i-1=m\end{cases} = \mathcal{F}_{\mathrm{std},<T_{(i-1)+1}}^{-\mathrm{Col}} \qquad (\text{since } i-1<m)\\ &= \mathcal{F}_{\mathrm{std},<T_i}^{-\mathrm{Col}} \qquad (\text{since } (i-1)+1=i)\\ &= \mathcal{F}_{\mathrm{std},<T}^{-\mathrm{Col}} \qquad (\text{since } T_i = T).\end{aligned}$$

[193] *Proof.* This inclusion is just saying that each element of the form $\nabla_{W,T}^{-\mathrm{Col}}$ with $W \in \mathrm{SYT}(\lambda)$ can also be written in the form $\nabla_{V,U}^{-\mathrm{Col}}$ for some $(\mu,U,V) \in \mathrm{SBT}(n)$ satisfying $U = T$. But this is clear: We just need to take $(\mu,U,V) = (\lambda,T,W)$ (this is an element of $\mathrm{SBT}(n)$ because $T \in \mathrm{SYT}(\lambda)$ and $W \in \mathrm{SYT}(\lambda)$; and it satisfies $U = T$ by its definition), and then we obviously have $\nabla_{V,U}^{-\mathrm{Col}} = \nabla_{W,T}^{-\mathrm{Col}}$, so that we have indeed written the element $\nabla_{W,T}^{-\mathrm{Col}}$ in the form $\nabla_{V,U}^{-\mathrm{Col}}$ for some $(\mu,U,V) \in \mathrm{SBT}(n)$ satisfying $U = T$.

Hence,

$$\underbrace{\mathcal{F}_i^{-\mathrm{Col}}}_{=\mathcal{F}_{\mathrm{std},\leq T}^{-\mathrm{Col}}} / \underbrace{\mathcal{F}_{i-1}^{-\mathrm{Col}}}_{=\mathcal{F}_{\mathrm{std},<T}^{-\mathrm{Col}}} = \mathcal{F}_{\mathrm{std},\leq T}^{-\mathrm{Col}} / \mathcal{F}_{\mathrm{std},<T}^{-\mathrm{Col}} \cong \mathcal{S}^{\lambda} \qquad \text{(as we proved above)}$$

$$= \mathcal{S}^{\lambda^{(i)}} \qquad \left(\text{since } \lambda = \lambda^{(i)}\right)$$

as left $\mathcal{A}$-modules. This proves Theorem 6.8.41. □

# References


[17f-hw7s] Darij Grinberg, *UMN Fall 2017 Math 4990 homework set #7 with solutions*, `http://www.cip.ifi.lmu.de/~grinberg/t/17f/hw7os.pdf`

[19fco] Darij Grinberg, *Enumerative Combinatorics: class notes (Drexel Fall 2019 Math 222 notes)*, 11 September 2022.
`http://www.cip.ifi.lmu.de/~grinberg/t/19fco/n/n.pdf`

[21s] Darij Grinberg, *An Introduction to Algebraic Combinatorics (Math 531, Winter 2024 lecture notes)*, 8 February 2025.
`http://www.cip.ifi.lmu.de/~grinberg/t/21s/lecs.pdf`

[22fco] Darij Grinberg, *Math 220: Enumerative Combinatorics, Fall 2022*.
`https://www.cip.ifi.lmu.de/~grinberg/t/22fco/`

[23wa] Darij Grinberg, *An introduction to the algebra of rings and fields*, 10 March 2025.
`https://www.cip.ifi.lmu.de/~grinberg/t/23wa/23wa.pdf`

[AFBCCL24] Ilani Axelrod-Freed, Sarah Brauner, Judy Hsin-Hui Chiang, Patricia Commins, Veronica Lang, *Spectrum of random-to-random shuffling in the Hecke algebra*, arXiv:2407.08644v2.

[Aigner07] Martin Aigner, *A Course in Enumeration*, Graduate Texts in Mathematics #238, Springer 2007.

[AlFais25] Faisal Al-Faisal, *PMATH 445/745 Representations of Finite Groups*, 21 April 2025.
`https://www.math.uwaterloo.ca/~f2alfais/notes/pm445-notes.pdf`

[Artin10] Michael Artin, *Algebra*, 2nd edition, Pearson 2010.
See `https://math.mit.edu/~poonen/701/comments_algebra.pdf` for unofficial errata.

[Axler24] Sheldon Axler, *Linear Algebra Done Right*, 4th edition, corrected version 9 March 2025.
`https://linear.axler.net/`

[BCGS25] Sarah Brauner, Patricia Commins, Darij Grinberg, Franco Saliola, *The q-deformed random-to-random family in the Hecke algebra*, arXiv:2503.17580v1.

[Beissi87] Janet Simpson Beissinger, *Similar constructions for Young tableaux and involutions, and their application to shiftable tableaux*, Discrete Mathematics **67** (1987), Issue 2, pp. 149–163.

[BilPaw13] Sara Billey, Brendan Pawlowski, *Permutation patterns, Stanley symmetric functions and generalized Specht modules*, Journal of Combinatorial Theory, Series A **127** (2014), pp. 85–120.

[BjoBre05] Anders Björner, Francesco Brenti, *Combinatorics of Coxeter Groups*, Springer 2005.
See `https://www.mat.uniroma2.it/~brenti/correct.ps` for errata.

[BleLau01] D. Blessenohl and H. Laue, *Algebraic combinatorics related to the free Lie algebra*, Séminaire Lotharingien de Combinatoire **29** (1992).

[BleSch05] Dieter Blessenohl, Manfred Schocker, *Noncommutative Character Theory of the Symmetric Group*, Imperial College Press 2005.

[BluCos16] Ben Blum-Smith, Samuel Coskey, *The Fundamental Theorem on Symmetric Polynomials: History's First Whiff of Galois Theory*, arXiv:1301.7116v5.

[Bona22] Miklós Bóna, *Combinatorics of Permutations*, 3rd edition, Taylor&Francis 2022.
`https://doi.org/10.1201/9780429274107`

[BorOls16] Alexei Borodin, Grigori Olshanski, *Representations of the infinite symmetric group*, Cambridge University Press 2016.
See `http://iitp.ru/upload/userpage/52/Book2016Feb26.pdf` for a preprint.

[Bourba03] Nicolas Bourbaki, *Algebra II: Chapters 4–7*, Springer 2003.

[Bourba74] Nicolas Bourbaki, *Algebra I: Chapters 1–3*, Addison-Wesley 1974.

[BrMaPe16] Murray Bremner, Sara Madariaga, Luiz A. Peresi, *Structure theory for the group algebra of the symmetric group, with applications to polynomial identities for the octonions*, Commentationes Mathematicae Universitatis Carolinae **57** (2016), issue 4, pp. 413–452.

See `https://www.cip.ifi.lmu.de/~grinberg/algebra/bmp-sn-errata.pdf` for unofficial errata.

[CaLaLe92] Christophe Carre, Alain Lascoux, Bernard Leclerc, *Turbo-straightening for decomposition into standard bases*, International Journal of Algebra and Computation **2** (1992), No. 03, pp. 275–290.

[CanWil89] E. R. Canfield, S. G. Williamson, *A Combinatorial Study of Straightening Bases*, Linear and Multilinear Algebra, **24** (1989), pp. 241–264.

[Carter09] Nathan Carter, *Visual Group Theory*, MAA 2009.
See `https://web.bentley.edu/empl/c/ncarter/vgt/errata.html` for errata.

[Clause91] Michael Clausen, *Multivariate Polynomials, Standard Tableaux and Representations of Symmetric Groups*, J. Symbolic Computation **11** (1991), issues 5–6, pp. 483–522.

[Conrad] Keith Conrad, *Expository papers*, `https://kconrad.math.uconn.edu/blurbs/` .

[CurRei62] Charles W. Curtis, Irving Reiner, *Representation Theory of Finite Groups and Associative Algebras*, Wiley 1962.

[CSScTo10] Tullio Ceccherini-Silberstein, Fabio Scarabotti, Filippo Tolli, *Representation Theory of the Symmetric Groups*, Cambridge University Press 2010.
See `https://mathoverflow.net/questions/485958/` for errata.

[Davis23] Tom Davis, *Pólya's Counting Theory*, 14 September 2023.
`http://www.geometer.org/mathcircles/polya.pdf`

[DiaFul23] Persi Diaconis, Jason Fulman, *The Mathematics of Shuffling Cards*, AMS 2023.

[DieSal18] Antonius B. Dieker, Franco V. Saliola, *Spectral analysis of random-to-random Markov chains*, Advances in Mathematics **323** (2018), pp. 427–485.

[DipJam85] Richard Dipper, Gordon James, *Blocks and idempotents of Hecke algebras of general linear groups*, Proceedings of the London Mathematical Society **s3-54** (1987), issue 1, pp. 57–82.

[DraGij09] Jan Draisma and Dion Gijswijt, *Invariant Theory with Applications*, 8 October 2009.
`https://web.archive.org/web/20240430053931/https://mathsites.unibe.ch/jdraisma/teaching/invtheory0910/lecturenotes12.pdf`

See https://www.cip.ifi.lmu.de/~grinberg/algebra/dgi-errata1.pdf for an inofficial list of errata.

[Dumas08] François Dumas, *An introduction to noncommutative polynomial invariants*, lecture notes (Cimpa-Unesco-Argentina "Homological methods and representations of non-commutative algebras", Mar del Plata, Argentina March 6 - 17, 2006).
https://lmbp.uca.fr/~fdumas/fichiers/CIMPA.pdf

[DumFoo04] David S. Dummit, Richard M. Foote, *Abstract Algebra*, 3rd edition, Wiley 2004. ISBN: 978-0-471-43334-7.
See https://site.uvm.edu/rfoote/files/2022/06/errata_3rd_edition.pdf for errata.

[Egge19] Eric S. Egge, *An Introduction to Symmetric Functions and Their Combinatorics*, AMS 2019.
See https://www.ericegge.net/cofsf/index.html for corrections and addenda.

[EGHetc11] Pavel Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, Elena Yudovina, *Introduction to Representation Theory*, with historical interludes by Slava Gerovitch, Student Mathematical Library **59**, AMS 2011, updated version 2024.

[Elman22] Richard Elman, *Lectures on Abstract Algebra*, 21 September 2024.
https://www.math.ucla.edu/~rse/algebra_book.pdf

[Fayers03] Matthew Fayers, *On the structure of Specht modules*, Journal of the London Mathematical Society (2) **67** (2003), pp. 85–102.
See https://webspace.maths.qmul.ac.uk/m.fayers/papers/specht3.pdf for a preprint.

[Feray16] Valentin Féray, *Approche Duale des représentations du groupe symétrique*, Cours Peccot, Collège de France, Spartacus 2016.
https://vferay.perso.math.cnrs.fr/Peccot/ApprocheDualeVFeray.pdf

[FrRoTh54] J. S. Frame, G. d. B. Robinson, and R. M. Thrall, *The hook graphs of the symmetric groups*, Canadian J. Math. **6** (1954), pp. 316–324.

[Fulton97] William Fulton, *Young Tableaux, With Applications to Representation Theory and Geometry*, London Mathematical Society Student Texts **35**, Cambridge University Press 1997.
See https://mathoverflow.net/questions/456463/ for errata.

[GarEge20] Adriano M. Garsia, Ömer Eğecioğlu, *Lectures in Algebraic Combinatorics*, Springer 2020.

[Garnir50] H. G. Garnir, *Théorie de la représentation linéaire des groupes symétriques*, Mem. Soc. Roy. Sci. Liege (4) **10**, No. 2 (1950), pp. 5–100.

[Garsia90] Adriano M. Garsia, *Combinatorics of the Free Lie Algebra and the Symmetric Group*, Analysis, et cetera, Res. Pap. in Honor of J. Moser's 60th Birthd. (1990), pp. 309–382, 1990.

[Goodma16] Frederick M. Goodman, *Algebra: Abstract and Concrete*, edition 2.6, 12 October 2016.
`https://homepage.divms.uiowa.edu/~goodman/algebrabook.dir/algebrabook.html`

[GriLaf22] Darij Grinberg, Nadia Lafrenière, *The one-sided cycle shuffles in the symmetric group algebra*, arXiv:2212.06274v4.
A shortened version has appeared in: Algebraic Combinatorics **7** (2024) no. 2, pp. 275–326.

[Grinbe15] Darij Grinberg, *Notes on the combinatorial fundamentals of algebra*, 15 September 2022.
`http://www.cip.ifi.lmu.de/~grinberg/primes2015/sols.pdf`
The numbering of theorems and formulas in this link might shift when the project gets updated; for a "frozen" version whose numbering is guaranteed to match that in the citations above, see `https://github.com/darijgr/detnotes/releases/tag/2022-09-15c` or arXiv:2008.09862v3.

[Grinbe18] Darij Grinberg, *Is this sum of cycles invertible in* $\mathbb{Q}[S_n]$*?*, answers to MathOverflow question #308536.
`https://mathoverflow.net/questions/308536/is-this-sum-of-cycles-invertible-in-mathbb-qs-n/308600#308600`

[Grinbe23] Darij Grinberg, *Commutator nilpotency for somewhere-to-below shuffles*, arXiv:2309.05340v2.

[Grinbe25] Darij Grinberg, *Rook sums in the symmetric group algebra*, draft, 17 May 2025.
`https://www.cip.ifi.lmu.de/~grinberg/algebra/rooksn.pdf`

[GriRei14] Darij Grinberg, Victor Reiner, *Hopf Algebras in Combinatorics*, arXiv:1409.8356v7.

[Hall08] Matthew John Hall, *Representations of the symmetric groups and combinatorics of the Frobenius-Young correspondence*, honors thesis at University of Arizona, 2008.

[Hodges23] Reuven Hodges, *A non-iterative formula for straightening fillings of Young diagrams*, arXiv:1710.05214v2, Sel. Math. New Ser. **30**, 37 (2024).

[Hopkin22] Sam Hopkins, *RSK via local transformations*, 13 September 2022. `https://www.samuelfhopkins.com/docs/rsk.pdf`

[Howe22] R. Michael Howe, *An Invitation to Representation Theory: Polynomial Representations of the Symmetric Group*, Springer 2022.

[James78] G. D. James, *The Representation Theory of the Symmetric Groups*, Lecture Notes in Mathematics **682**, Springer 1978.

[JamKer81] Gordon James, Adalbert Kerber, *The representation theory of the symmetric group*, Encyclopedia of Mathematics and its Applications **16**, Addison-Wesley 1981.

[JamMat98] Gordon James, Andrew Mathas, *The irreducible Specht modules in characteristic* 2, Bulletin of the London Mathematical Society **31**, Issue 4, July 1999, pp. 457–462.

[JamPee79] G.D. James and M.H. Peel, *Specht series for skew representations of symmetric groups*, Journal of Algebra **56** (1979), pp. 343–364.

[Jozefi89] Tadeusz Józefiak, *Characters of projective representations of symmetric groups*, Expo. Math. **7** (1989), pp. 193–247.

[Jucys74] A.-A. A. Jucys, *Symmetric polynomials and the center of the symmetric group ring*, Reports on Mathematical Physics **5** (1974), pp. 107–112.

[KarPur23] Steven N. Karp, Kevin Purbhoo, *Universal Plücker coordinates for the Wronski map and positivity in real Schubert calculus*, arXiv:2309.04645v1.

[Kerber99] Adalbert Kerber, *Applied Finite Group Actions*, 2nd edition, Springer 1999.

[Kleshc05] Alexander Kleshchev, *Linear and Projective Representations of Symmetric Groups*, Cambridge University Press 2005.
See `https://darkwing.uoregon.edu/~klesh/CorrBook.pdf` for errata.

[Knuth98] Donald Ervin Knuth, *The Art of Computer Programming, volume 3: Sorting and Searching*, 2nd edition, Addison-Wesley 1998.
See `https://www-cs-faculty.stanford.edu/~knuth/taocp.html` for errata.

[Knutso73] Donald Knutson, *λ-Rings and the Representation Theory of the Symmetric Group*, Lecture Notes in Mathematics **308**, Springer, New York 1973.

[KraPro10] Hanspeter Kraft, Claudio Procesi, *Classical invariant theory: A primer*, July 1996.
`https://kraftadmin.wixsite.com/hpkraft`
See `http://www.cip.ifi.lmu.de/~grinberg/algebra/KP-errata-web.pdf` for unofficial errata.

[Krishn86] V. Krishnamurthy, *Combinatorics: Theory and Applications*, Ellis Horwood Ltd. 1986.

[Krob95] Daniel Krob, *Eléments de combinatoire*, ENS 1995.
`https://web.archive.org/web/20220119141152/http://krob.cesames.net/IMG/ps/combi.ps`

[Kuenze15] Matthias Künzer, *Ties for the integral group ring of the symmetric group*, thesis at Universität Ulm 1998, version 2015.
`https://pnp.mathematik.uni-stuttgart.de/lexmath/kuenzer/thesis.pdf`

[Ladisc11] Frieder Ladisch, *Answer to MathOverflow question #58776 ("Pseudo-idempotent matrix generating a free module")*,
`https://mathoverflow.net/questions/58776`

[Lafren19] Nadia Lafrenière, *Valeurs propres des opérateurs de mélanges symétrisés*, PhD thesis, 2019, arXiv:1912.07718v1.

[LakBro18] V. Lakshmibai, Justin Brown, *Flag Varieties: An Interplay of Geometry, Combinatorics, and Representation Theory*, 2nd edition, Springer 2018.

[Lassue23] Caroline Lassueur, *Modular Representation Theory of Finite Groups: An Introduction*, 2 May 2023.
`https://seafile.rlp.net/f/9c6d8cbf1d06458d9b47/`

[Leeuwe11] Marc A. A. van Leeuwen, *The Robinson-Schensted and Schützenberger algorithms, an elementary approach*, version 25 Nov 2011.
`http://www-math.univ-poitiers.fr/~maavl/`

[Liu10] Ricky Ini Liu, *Specht Modules and Schubert Varieties for General Diagrams*, PhD thesis at MIT 2010.
`https://dspace.mit.edu/handle/1721.1/60196`

[Liu16] Ricky Ini Liu, *Complete branching rules for Specht modules*, Journal of Algebra **446** (2016), pp. 77–102.

[Liulev80] Arunas Liulevicius, *Arrows, symmetries and representation rings*, J. Pure Appl. Algebra **19** (1980), pp. 259–273.

[Loehr11] Nicholas A. Loehr, *Bijective Combinatorics*, Chapman & Hall/CRC 2011.

[Lorenz18] Martin Lorenz, *A Tour of Representation Theory*, Graduate Studies in Mathematics **193**, AMS 2018.
See `https://web.archive.org/web/20250123115734/https://www.math.temple.edu/~lorenz/ToRCorrections.pdf` for errata.

[Macdon95] Ian G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford Mathematical Monographs, 2nd edition, Oxford Science Publications 1995.
`https://math.berkeley.edu/~corteel/MATH249/macdonald.pdf`

[Makhli11] Igor Makhlin, *Answer to MathOverflow question #83150*,
`https://mathoverflow.net/questions/83150/why-are-jucys-murphy-elements-eigenvalues-whole-numbers/83493#83493`

[McClea23] John McCleary, *Actions of Groups: A Second Course in Algebra*, Cambridge University Press 2023.

[Meliot17] Pierre-Loic Meliot, *Representation Theory of Symmetric Groups*, CRC Press 2017.

[MenRem15] Anthony Mendes, Jeffrey Remmel, *Counting with Symmetric Functions*, Springer 2015.

[Milion24] Christoforos Athanasios Milionis, *Different approaches to Young-Jucys-Murphy elements*, master's thesis at RFWU Bonn, 2024.
`https://www.math.uni-bonn.de/ag/stroppel/Thesis_Milionis.pdf`

[Moran92] Gadi Moran, *The center of $\mathbb{Z}[S_{n+1}]$ is the set of symmetric polynomials in n commuting transposition-sums*, Trans. Amer. Math. Soc. **332** (1992), pp. 167–180.

[Morel19] Sophie Morel, *Representation theory*, lecture notes, 26 March 2019.
`https://perso.ens-lyon.fr/sophie.morel/rep_theory_notes.pdf`

[Murphy81] G. E. Murphy, *A new construction of Young's Seminormal Representation of the Symmetric Groups*, Journal of Algebra **69** (1981), pp. 287–297.

[Murphy83] G. E. Murphy, *The idempotents of the symmetric group and Nakayama's conjecture*, Journal of Algebra **81**, Issue 1, March 1983, pp. 258–265.

[Murphy92] G. E. Murphy, *On the representation theory of the symmetric groups and associated Hecke algebras*, J. Algebra **152** (1992), no. 2, pp. 492–513.

[Murphy95] G. E. Murphy, *The representations of Hecke algebras of type $A_n$*, J. Algebra **173** (1995), no. 1, pp. 97–121.

[Newman12] Stephen C. Newman, *A classical introduction to Galois theory*, Wiley 2012.

[Pak22] Igor Pak, *What is a combinatorial interpretation?*, arXiv:2209.06142v1.

[Peel71] M. H. Peel, *Hook representations of the symmetric groups*, Glasgow Mathematical Journal **12** (1971), issue 2, pp. 136–149.

[Prasad15] Amritanshu Prasad, *Representation Theory: A Combinatorial Viewpoint*, Cambridge University Press 2015.

[Proces07] Claudio Procesi, *Lie Groups: An Approach through Invariants and Representations*, Springer 2007.

[Purbho22] Kevin Purbhoo, *An identity in the Bethe subalgebra of $\mathbb{C}[\mathfrak{S}_n]$*, arXiv:2207.05743v2, published in: Proceedings of the London Mathematical Society **127** (2023), issue 5, pp. 1247–1267.

[ReiShi95a] Victor Reiner, Mark Shimozono, *Specht Series for Column-Convex Diagrams*, Journal of Algebra **174** (1995), pp. 489–522.

[ReiShi95b] Victor Reiner, Mark Shimozono, *Plactification*, Journal of Algebraic Combinatorics **4** (1995), pp. 331–351.

[Rentel23] Paul Renteln, *A natural idempotent in the descent algebra of a finite Coxeter group*, Algebraic Combinatorics **6**, issue 5 (2023), pp. 1177–1188.

[ReSaWe11] Victor Reiner, Franco Saliola, Volkmar Welker, *Spectra of Symmetrized Shuffling Operators*, arXiv:1102.2460v2. Published in: Memoirs of the AMS **228** (2014).

[Reuten93] Christophe Reutenauer, *Free Lie Algebras*, Oxford University Press 1993.

[Romik15] Dan Romik, *The Surprising Mathematics of Longest Increasing Subsequences*, Cambridge University Press 2015.
`https://web.archive.org/web/20240818002142/https:`

//www.math.ucdavis.edu/~romik/download-book.php
See https://www.math.ucdavis.edu/~romik/data/uploads/misc/lisbook-errata.pdf for errata.

[Roseng15] Hjalmar Rosengren, *Lectures on Representation Theory*, 17 October 2019.
https://gu.instructure.com/courses/27395

[RotGui71] Gian-Carlo Rota, *Combinatorial Theory and Invariant Theory*, lecture notes taken by L. Guibas, 1971.

[Ruther48] Daniel Edwin Rutherford, *Substitutional Analysis*, Edinburgh University Press, Edinburgh 1948.

[Sagan01] Bruce Sagan, *The Symmetric Group*, Graduate Texts in Mathematics **203**, 2nd edition 2001.
https://doi.org/10.1007/978-1-4757-6804-6
See https://users.math.msu.edu/users/bsagan/Books/Sym/errata.pdf for errata.

[Sagan19] Bruce Sagan, *Combinatorics: The Art of Counting*, Graduate Studies in Mathematics **210**, AMS 2020.
https://users.math.msu.edu/users/bsagan/Books/Aoc/final.pdf
See https://users.math.msu.edu/users/bsagan/Books/Aoc/errata.pdf for errata.

[Sam20] Steven V. Sam, *Notes for Math 202B*, 28 February 2020.
https://mathweb.ucsd.edu/~ssam/old/20W-202B/notes.pdf

[Savage22] Alistair Savage, *Symmetric Functions*, lecture notes, 2 May 2022.
https://alistairsavage.ca/symfunc/notes/Savage-SymmetricFunctions.pdf

[Schnei13] Peter Schneider, *Modular Representation Theory of Finite Groups*, Springer 2013.

[Serre77] Jean-Pierre Serre, *Linear Representations of Finite Groups*, Springer 1977.
See https://www.college-de-france.fr/media/jean-pierre-serre/UPL5190088847250963678_Corrections_to_Serre_books_2019.pdf for errata.

[Smith95] Larry Smith, *Polynomial Invariants of Finite Groups*, A K Peters 1995.

[Specht35] Wilhelm Specht, *Die irreduziblen Darstellungen der symmetrischen Gruppe*, Mathematische Zeitschrift **39** (1935), pp. 696–711.

[Stanle11] Richard P. Stanley, *Enumerative Combinatorics, volume 1*, Second edition, Cambridge University Press 2012.
See `http://math.mit.edu/~rstan/ec/` for errata.

[Stanle18] Richard P. Stanley, *Algebraic Combinatorics: Walks, Trees, Tableaux, and More*, Undergraduate Texts in Mathematics, 2nd edition, Springer 2018.
`http://www-math.mit.edu/~rstan/algcomb/index.html`

[Stanle23] Richard P. Stanley, *Enumerative Combinatorics, volume 2*, Second edition, Cambridge University Press 2023.
See `http://math.mit.edu/~rstan/ec/` for errata.

[Sternb95] Shlomo Sternberg, *Group Theory and Physics*, Cambridge University Press 1995.

[StoLui19] Manfred Stoll, Ronald von Luijk, *Linear algebra II*, 2019.
`https://pub.math.leidenuniv.nl/~luijkrmvan/linalg2/2019/`
(scroll down to "Literatuur").

[TauZas59] Olga Taussky, Hans Zassenhaus, *On the similarity transformation between a matrix and its transpose*, Pacific J. Math. **9**(3) (1959), pp. 893–896.

[Taylor01] Brian D. Taylor, *A Straightening Algorithm for Row-Convex Tableaux*, Journal of Algebra **236** (2001), pp. 155–191.

[VerTsi16] N. V. Tsilevich, A. M. Vershik, *On the relation of some combinatorial functions to representation theory*, arXiv:1612.09088v1.

[Waerde91] B. L. van der Waerden, *Algebra: Volume I and Volume II*, based in part on lectures by E. Artin and E. Noether, 7th edition (volume I) and 5th edition (volume II), Springer 1991.

[Wallac88] Nolan R. Wallach, *Lie Algebra Cohomology and Holomorphic Continuation of Generalized Jacquet Integrals*, Advanced Studies in Pure Mathematics **14** (1988), Representations of Lie Groups, Kyoto, Hiroshima, 1986, pp. 123–151.

[Webb16] Peter Webb, *A Course in Finite Group Representation Theory*, Cambridge University Press 2016.
See `https://www-users.cse.umn.edu/~webb/RepBook/typos` for errata.

[Welsh92] Trevor Alan Welsh, *Young Tableaux and Modules of Groups and Lie Algebras*, PhD thesis at University of Southampton, 1992.
`https://eprints.soton.ac.uk/462077/`

[Wever47] Franz Wever, *Über Invarianten in Lie'schen Ringen*, Math. Annalen **120** (1947), pp. 563–580.

[Wildon18] Mark Wildon, *Representation theory of the symmetric group*, 5 April 2018.
`https://www.ma.rhul.ac.uk/~uvah099/Maths/Sym/SymGroup2014.pdf`

[Wildon19] Mark Wildon, *On Bidigare's proof of Solomon's theorem*, 6 February 2019.
`http://www.ma.rhul.ac.uk/~uvah099/Maths/Bidigare4.pdf`
See `https://www.cip.ifi.lmu.de/~grinberg/algebra/bidigare4-2019-errata.pdf` for an inofficial list of errata.

[Wildon20] Mark Wildon, *An involutive introduction to symmetric functions*, 8 May 2020.
`http://www.ma.rhul.ac.uk/~uvah099/Maths/Sym/SymFuncs2020.pdf`
See `http://www.cip.ifi.lmu.de/~grinberg/algebra/symfuncs2017-2020-05-08-errata.pdf` for an inofficial list of errata.

[Willia03] Geordie Williamson, *Mind your P and Q-symbols: Why the Kazhdan-Lusztig basis of the Hecke algebra of type A is cellular*, Honours thesis at the University of Sydney, October 2003.
See `https://www.maths.usyd.edu.au/u/geordie/pandq-errata.pdf` for an inofficial list of errata.

[Young77] Alfred Young, *The Collected Papers of Alfred Young 1873–1940*, University of Toronto Press 1977.

[Zelevi81a] A. V. Zelevinsky, *A generalization of the Littlewood–Richardson rule and the Robinson–Schensted–Knuth correspondence*, J. Algebra **69** (1981), pp. 82–94.

[Zelevi81b] Andrey V. Zelevinsky, *Representations of Finite Classical Groups: A Hopf Algebra Approach*, Lecture Notes in Mathematics **869**, Springer 1981.

[Zimmer14] Alexander Zimmermann, *Representation Theory: A Homological Algebra Point of View*, Springer 2014.