Algorithms in Invariant Theory
Bernd Sturmfels
Second Edition 2008
Comments and corrections by Darij Grinberg

The following list contains my comments to Bernd Sturmfels's book *Algorithms in Invariant Theory*, specifically its 2nd edition (Springer, 2008). Some of these comments are corrections; others are more subjective improvements.

I have read the following parts of the book: Chapter 1, the elementary parts of Chapter 2 (§2.1, §2.2 until Example 2.2.4, §2.4), most of Chapter 3 (§3.1, §3.2, §3.3 until Algorithm 3.3.4, §3.6, §3.7 sans the second proof of Proposition 3.7.4 and until Lemma 3.7.6), and the early bits of Chapter 4 (§4.1, §4.3).

7. Errata and comments

1. **page 4, proof of Proposition 1.1.2:** It should be said that p_0 is to be understood as 1. (Otherwise, "$p_{i_1} p_{i_2} \ldots p_{i_n}$" might not make sense.)

2. **page 5, Proposition 1.1.3:** This proposition requires $n > 1$.

3. **page 5, proof of Proposition 1.1.3:** After "and therefore D divides $\widetilde{h}$.", add "Furthermore, the polynomial $\widetilde{h}/D$ is symmetric, since each permutation $\sigma \in S_n$ satisfies

$$\left(\widetilde{h}/D\right)\left(x_{\sigma_1}, \ldots, x_{\sigma_n}\right) = \frac{\widetilde{h}\left(x_{\sigma_1}, \ldots, x_{\sigma_n}\right)}{D\left(x_{\sigma_1}, \ldots, x_{\sigma_n}\right)} = \frac{\operatorname{sign}(\sigma) \cdot \widetilde{h}\left(x_1, \ldots, x_n\right)}{\operatorname{sign}(\sigma) \cdot D\left(x_1, \ldots, x_n\right)} = \frac{\widetilde{h}\left(x_1, \ldots, x_n\right)}{D\left(x_1, \ldots, x_n\right)} = \widetilde{h}/D.$$

".

4. **page 6:** After "The polynomials a_λ are precisely the nonzero images of monomials under antisymmetrization", add "(up to sign)".

5. **page 6:** The definition of "antisymmetrization" here uses the notion of "canonical projection", which is not explained here. Better to say that the *antisymmetrization* of a polynomial $f\left(x_1, \ldots, x_n\right)$ is defined to be $\sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot f\left(x_{\sigma_1}, \ldots, x_{\sigma_n}\right)$.

6. **page 6:** The word "discriminant", as it is used here, simply means the polynomial $D = D\left(x_1, \ldots, x_n\right)$. This should be clarified.

7. **page 6, Corollary 1.1.4:** It should be explained that "partitions of d into at most n parts" means "partitions of d with n entries (not necessarily all positive)".

8. **page 8:** "The largest monomial of a polynomial" → "The largest monomial of a nonzero polynomial".

9. **page 8:** "by the initial monomials of all polynomials" → "by the initial monomials of all nonzero polynomials".

10. **page 8, Example 1.2.1:** "versus a Gröbner bases" → "versus a Gröbner basis".

11. **page 11, proof of Theorem 1.2.4:** "By Corollary 1.2.3, the set of initial monomials $\{\text{init}(f) : f \in I \setminus \{\mathcal{G}\}\}$ has a minimal element $\text{init}(f_0)$ with respect to "$\prec$"": In my opinion, this is not obvious enough to be left unexplained, at least for a reader who has not already seen this kind of (highly non-constructive) argument before. The argument being made here is the following: The set of initial monomials $\{\text{init}(f) : f \in I \setminus \langle\mathcal{G}\rangle\}$ is nonempty (since we assumed that $I \setminus \langle\mathcal{G}\rangle$ is nonempty). If this set had no minimal element (with respect to "$\prec$"), then we could recursively construct an infinite chain $m_1 \succ m_2 \succ m_3 \succ \cdots$ of monomials in this set by picking m_1 arbitrary (this is possible, since the set $\{\text{init}(f) : f \in I \setminus \langle\mathcal{G}\rangle\}$ is nonempty) and then choosing each further element m_i to be smaller than m_{i-1} (this is possible, since m_{i-1} is not minimal in this set, because this set has no minimal element). But such a chain would contradict Corollary 1.2.3. Hence, the set $\{\text{init}(f) : f \in I \setminus \langle\mathcal{G}\rangle\}$ must have a minimal element after all.

12. **page 12, proof of Theorem 1.2.7:** Here is a proof of the identity

$$h_k(x_k, \ldots, x_n) + \sum_{i=1}^{k} (-1)^i h_{k-i}(x_k, \ldots, x_n)\, \sigma_i(x_1, \ldots, x_{k-1}, x_k, \ldots, x_n) = 0. \tag{1}$$

Proof of (1): We work in the polynomial ring $\mathbf{C}[x_1, x_2, \ldots, x_n]$ (actually, $\mathbf{C}$ could be replaced by any base ring here). Let $\mathcal{P}(n)$ denote the set of all subsets of $\{1, 2, \ldots, n\}$. We define the complete homogeneous symmetric polynomials

$$h_i(x_{k..n}) := (\text{sum of all monomials in } x_k, x_{k+1}, \ldots, x_n \text{ that have degree } i) \tag{2}$$

for all $i \in \mathbf{N}$ (so that $h_0(x_{k..n}) = 1$ because the only monomial of degree 0 is 1) and the elementary symmetric polynomials

$$\sigma_i(x_{1..n}) := \sum_{j_1 < j_2 < \cdots < j_i \text{ in } \{1,2,\ldots,n\}} x_{j_1} x_{j_2} \cdots x_{j_i} = \sum_{\substack{S \in \mathcal{P}(n);\\ |S| = i}} \prod_{s \in S} x_s \tag{3}$$

for all $i \in \mathbf{N}$ (so that $\sigma_0(x_{1..n}) = 1$ because the only 0-element subset of $\{1, 2, \ldots, n\}$ is $\varnothing$). Note that these polynomials $h_i(x_{k..n})$ and $\sigma_i(x_{1..n})$ are denoted by $h_i(x_k, \ldots, x_n)$ and $\sigma_i(x_1, \ldots, x_{k-1}, x_k, \ldots, x_n)$ in the book. Thus, the identity (1), which we must prove, can be rewritten as follows:

$$h_k(x_{k..n}) + \sum_{i=1}^{k} (-1)^i h_{k-i}(x_{k..n}) \sigma_i(x_{1..n}) = 0. \tag{4}$$

But we can simplify this identity even further. Namely, we observe that

$$\begin{aligned}
&\sum_{i=0}^{k} (-1)^i h_{k-i}(x_{k..n}) \sigma_i(x_{1..n}) \\
&= \underbrace{(-1)^0}_{=1} \underbrace{h_{k-0}(x_{k..n})}_{=h_k(x_{k..n})} \underbrace{\sigma_0(x_{1..n})}_{=1} + \sum_{i=1}^{k} (-1)^i h_{k-i}(x_{k..n}) \sigma_i(x_{1..n}) \\
&= h_k(x_{k..n}) + \sum_{i=1}^{k} (-1)^i h_{k-i}(x_{k..n}) \sigma_i(x_{1..n}).
\end{aligned}$$

Hence, the identity (4), which we must prove, can be rewritten as

$$\sum_{i=0}^{k} (-1)^i h_{k-i}(x_{k..n}) \sigma_i(x_{1..n}) = 0. \tag{5}$$

It is this latter identity (5) that we shall now prove; the original identity (1) will then follow.

Let $\mathfrak{M}$ be the set of all monomials in the variables $x_k, x_{k+1}, \ldots, x_n$. Thus, we can restate the definition (2) of $h_i(x_{k..n})$ as follows: For all $i \in \mathbf{N}$, we have

$$h_i(x_{k..n}) = \sum_{\substack{\mathfrak{m} \in \mathfrak{M}; \\ \deg \mathfrak{m} = i}} \mathfrak{m}. \tag{6}$$

For any subset S of $[n]$, let x_S denote the monomial $\prod_{s \in S} x_s$. Thus, we can restate the definition (3) of $\sigma_i(x_{1..n})$ as follows: For all $i \in \mathbf{N}$, we have

$$\sigma_i(x_{1..n}) = \sum_{\substack{S \in \mathcal{P}(n); \\ |S| = i}} x_S. \tag{7}$$

Thus,

$$\sum_{i=0}^{k} (-1)^i \underbrace{h_{k-i}(x_{k..n})}_{\substack{=\sum\limits_{\substack{\mathfrak{m}\in\mathfrak{M};\\ \deg\mathfrak{m}=k-i}} \mathfrak{m} \\ \text{(by (6), applied to } k-i \\ \text{instead of } i)}} \underbrace{\sigma_i(x_{1..n})}_{\substack{=\sum\limits_{\substack{S\in\mathcal{P}(n);\\ |S|=i}} x_S \\ \text{(by (7))}}}$$

$$= \sum_{i=0}^{k} (-1)^i \left(\sum_{\substack{\mathfrak{m}\in\mathfrak{M};\\ \deg\mathfrak{m}=k-i}} \mathfrak{m} \right) \sum_{\substack{S\in\mathcal{P}(n);\\ |S|=i}} x_S$$

$$= \sum_{i=0}^{k} \sum_{\substack{\mathfrak{m}\in\mathfrak{M};\\ \deg\mathfrak{m}=k-i}} \sum_{\substack{S\in\mathcal{P}(n);\\ |S|=i}} \underbrace{(-1)^i}_{\substack{=(-1)^{|S|} \\ \text{(since } i=|S|)}} \mathfrak{m} x_S$$

$$= \sum_{i=0}^{k} \sum_{\substack{\mathfrak{m}\in\mathfrak{M};\\ \deg\mathfrak{m}=k-i}} \sum_{\substack{S\in\mathcal{P}(n);\\ |S|=i}} (-1)^{|S|} \mathfrak{m} x_S. \tag{8}$$

However, the triple summation sign $\sum\limits_{i=0}^{k} \sum\limits_{\substack{\mathfrak{m}\in\mathfrak{M};\\ \deg\mathfrak{m}=k-i}} \sum\limits_{\substack{S\in\mathcal{P}(n);\\ |S|=i}}$ can be rewritten as a single summation sign $\sum\limits_{\substack{(\mathfrak{m},S)\in\mathfrak{M}\times\mathcal{P}(n);\\ \deg\mathfrak{m}+|S|=k}}$, because the conditions $\deg\mathfrak{m} = k-i$ and $|S| = i$ add up to $\deg\mathfrak{m} + |S| = (k-i) + i = k$ (and conversely, $\deg\mathfrak{m} + |S| = k$ implies that $\deg\mathfrak{m} = k-i$ and $|S| = i$ for a unique i). Hence, we can rewrite (8) as

$$\sum_{i=0}^{k} (-1)^i h_{k-i}(x_{k..n}) \sigma_i(x_{1..n}) = \sum_{\substack{(\mathfrak{m},S)\in\mathfrak{M}\times\mathcal{P}(n);\\ \deg\mathfrak{m}+|S|=k}} (-1)^{|S|} \mathfrak{m} x_S. \tag{9}$$

Now, for any monomial $\mathfrak{m}$ in the variables $x_1, x_2, \ldots, x_n$, let $\operatorname{Supp}\mathfrak{m}$ denote the set of all $i \in \{1, 2, \ldots, n\}$ such that the variable x_i appears in $\mathfrak{m}$ (that is, such that $\mathfrak{m}$ is divisible by x_i as a monomial). Clearly, if $\mathfrak{m} \in \mathfrak{M}$, then $\operatorname{Supp}\mathfrak{m} \subseteq \{k, k+1, \ldots, n\}$ (since $\mathfrak{m} \in \mathfrak{M}$ means that $\mathfrak{m}$ is a monomial in $x_k, x_{k+1}, \ldots, x_n$ only). Furthermore, for any $S \in \mathcal{P}(n)$, we have $\operatorname{Supp}(x_S) = S$. More generally, for any $\mathfrak{m} \in \mathfrak{M}$ and $S \in \mathcal{P}(n)$, we have

$$\operatorname{Supp}(\mathfrak{m} x_S) = S \cup \operatorname{Supp}\mathfrak{m} \tag{10}$$

(and more generally, we have $\operatorname{Supp}(\mathfrak{m}\mathfrak{n}) = \operatorname{Supp}\mathfrak{m} \cup \operatorname{Supp}\mathfrak{n}$ for any two monomials $\mathfrak{m}$ and $\mathfrak{n}$).

It is easy to see that for every pair $(\mathfrak{m}, S) \in \mathfrak{M} \times \mathcal{P}(n)$ satisfying $\deg\mathfrak{m} + |S| = k$, we have

$$\max(\operatorname{Supp}(\mathfrak{m}x_S)) \in \{k, k+1, \ldots, n\} \tag{11}$$

[1]. Hence, we can split up the sum on the right hand side of (9) according to the value of $\max(\operatorname{Supp}(\mathfrak{m}x_S))$ as follows:

$$\sum_{\substack{(\mathfrak{m},S)\in\mathfrak{M}\times\mathcal{P}(n);\\ \deg\mathfrak{m}+|S|=k}} (-1)^{|S|}\,\mathfrak{m}x_S = \sum_{j=k}^{n} \sum_{\substack{(\mathfrak{m},S)\in\mathfrak{M}\times\mathcal{P}(n);\\ \deg\mathfrak{m}+|S|=k;\\ \max(\operatorname{Supp}(\mathfrak{m}x_S))=j}} (-1)^{|S|}\,\mathfrak{m}x_S. \tag{12}$$

Now, fix $j \in \{k, k+1, \ldots, n\}$. We shall prove that the sum

$$\sum_{\substack{(\mathfrak{m},S)\in\mathfrak{M}\times\mathcal{P}(n);\\ \deg\mathfrak{m}+|S|=k;\\ \max(\operatorname{Supp}(\mathfrak{m}x_S))=j}} (-1)^{|S|}\,\mathfrak{m}x_S \tag{13}$$

is 0 by breaking it up into two mutually cancelling parts. Namely, we define the set

$$\mathfrak{A} := \{(\mathfrak{m}, S) \in \mathfrak{M} \times \mathcal{P}(n) \mid \deg\mathfrak{m} + |S| = k \text{ and } \max(\operatorname{Supp}(\mathfrak{m}x_S)) = j\};$$

[1] *Proof of (11):* Let $(\mathfrak{m}, S) \in \mathfrak{M} \times \mathcal{P}(n)$ be a pair satisfying $\deg\mathfrak{m} + |S| = k$. We must prove that $\max(\operatorname{Supp}(\mathfrak{m}x_S)) \in \{k, k+1, \ldots, n\}$. In other words, we must prove that $\max(S \cup \operatorname{Supp}\mathfrak{m}) \in \{k, k+1, \ldots, n\}$ (since (10) says that $\operatorname{Supp}(\mathfrak{m}x_S) = S \cup \operatorname{Supp}\mathfrak{m}$).

We are in one of the following two cases:

- *Case 1:* We have $\mathfrak{m} = 1$. Then, $\deg\mathfrak{m} = 0$ and thus $k = \underbrace{\deg\mathfrak{m}}_{=0} + |S| = |S|$. Hence, S is a k-element subset of $\{1, 2, \ldots, n\}$ and thus contains at least one of the numbers $k, k+1, \ldots, n$ (since otherwise, S would be a subset of $\{1, 2, \ldots, k-1\}$ and therefore would have at most $k-1$ many elements). Therefore, the union $S \cup \operatorname{Supp}\mathfrak{m}$ must also contain at least one of the numbers $k, k+1, \ldots, n$ (since it contains any element that S contains). Since $S \cup \operatorname{Supp}\mathfrak{m}$ is a subset of $\{1, 2, \ldots, n\}$, we conclude that $\max(S \cup \operatorname{Supp}\mathfrak{m}) \in \{k, k+1, \ldots, n\}$.
- *Case 2:* We have $\mathfrak{m} \neq 1$. Hence, the monomial $\mathfrak{m}$ is not constant, and thus must contain at least one of the indeterminates $x_k, x_{k+1}, \ldots, x_n$ (since it is a monomial in these indeterminates). In other words, the set $\operatorname{Supp}\mathfrak{m}$ contains at least one of the numbers $k, k+1, \ldots, n$. Therefore, the union $S \cup \operatorname{Supp}\mathfrak{m}$ must also contain at least one of the numbers $k, k+1, \ldots, n$ (since it contains any element that $\operatorname{Supp}\mathfrak{m}$ contains). Since $S \cup \operatorname{Supp}\mathfrak{m}$ is a subset of $\{1, 2, \ldots, n\}$, we conclude that $\max(S \cup \operatorname{Supp}\mathfrak{m}) \in \{k, k+1, \ldots, n\}$.

Hence, in both of these cases, we have shown that $\max(S \cup \operatorname{Supp}\mathfrak{m}) \in \{k, k+1, \ldots, n\}$. This completes the proof of (11).

this is the indexing set of our sum (13).

Now, if $(\mathfrak{m}, S) \in \mathfrak{A}$ is a pair satisfying $j \notin S$, then $x_j \mid \mathfrak{m}$ (because $(\mathfrak{m}, S) \in \mathfrak{A}$ entails $\max(\operatorname{Supp}(\mathfrak{m}x_S)) = j$, so that $j = \max(\operatorname{Supp}(\mathfrak{m}x_S)) \in \operatorname{Supp}(\mathfrak{m}x_S) = S \cup \operatorname{Supp}\mathfrak{m}$ (by (10)), which yields $j \in \operatorname{Supp}\mathfrak{m}$ (because $j \notin S$) and therefore $x_j \mid \mathfrak{m}$), and therefore $\mathfrak{m}/x_j \in \mathfrak{M}$ and thus $(\mathfrak{m}/x_j, S \cup \{j\}) \in \mathfrak{A}$ [2]. Thus, we obtain a map

$$\{(\mathfrak{m}, S) \in \mathfrak{A} \mid j \notin S\} \to \{(\mathfrak{m}, S) \in \mathfrak{A} \mid j \in S\}, \quad (\mathfrak{m}, S) \mapsto (\mathfrak{m}/x_j, S \cup \{j\}). \tag{14}$$

Conversely, if $(\mathfrak{m}, S) \in \mathfrak{A}$ is a pair satisfying $j \in S$, then $(\mathfrak{m}x_j, S \setminus \{j\}) \in \mathfrak{A}$ [3]. Thus, we obtain a map

$$\{(\mathfrak{m}, S) \in \mathfrak{A} \mid j \in S\} \to \{(\mathfrak{m}, S) \in \mathfrak{A} \mid j \notin S\}, \quad (\mathfrak{m}, S) \mapsto (\mathfrak{m}x_j, S \setminus \{j\}). \tag{15}$$

[2]*Proof.* We must show that $(\mathfrak{m}/x_j, S \cup \{j\}) \in \mathfrak{M} \times \mathcal{P}(n)$ and $\deg(\mathfrak{m}/x_j) + |S \cup \{j\}| = k$ and $\max\left(\operatorname{Supp}\left((\mathfrak{m}/x_j)\, x_{S\cup\{j\}}\right)\right) = j$.

The first of these three statements is clear because $\mathfrak{m}/x_j \in \mathfrak{M}$ and $S \cup \{j\} \in \mathcal{P}(n)$. The second statement follows from

$$\underbrace{\deg(\mathfrak{m}/x_j)}_{=\deg\mathfrak{m}-1} + \underbrace{|S \cup \{j\}|}_{\substack{=|S|+1 \\ (\text{since } j \notin S)}} = (\deg\mathfrak{m} - 1) + (|S| + 1) = \deg\mathfrak{m} + |S| = k \qquad (\text{since } (\mathfrak{m}, S) \in \mathfrak{A}).$$

It remains to prove the third statement. But $x_{S\cup\{j\}} = x_S x_j$ (since $j \notin S$) and thus $(\mathfrak{m}/x_j)\, x_{S\cup\{j\}} = (\mathfrak{m}/x_j)(x_S x_j) = \mathfrak{m}x_S$. Hence, $\max\left(\operatorname{Supp}\left((\mathfrak{m}/x_j)\, x_{S\cup\{j\}}\right)\right) = \max(\operatorname{Supp}(\mathfrak{m}x_S)) = j$ because of $(\mathfrak{m}, S) \in \mathfrak{A}$. This proves the third of the three statements we needed.

Hence, $(\mathfrak{m}/x_j, S \cup \{j\}) \in \mathfrak{A}$ follows.

[3]*Proof.* We must show that $(\mathfrak{m}x_j, S \setminus \{j\}) \in \mathfrak{M} \times \mathcal{P}(n)$ and $\deg(\mathfrak{m}x_j) + |S \setminus \{j\}| = k$ and $\max\left(\operatorname{Supp}\left((\mathfrak{m}x_j)\, x_{S\setminus\{j\}}\right)\right) = j$.

The first of these three statements is clear because $\mathfrak{m}x_j \in \mathfrak{M}$ (since $j \in \{k, k+1, \ldots, n\}$) and $S \setminus \{j\} \in \mathcal{P}(n)$. The second statement follows from

$$\underbrace{\deg(\mathfrak{m}x_j)}_{=\deg\mathfrak{m}+1} + \underbrace{|S \setminus \{j\}|}_{\substack{=|S|-1 \\ (\text{since } j \in S)}} = (\deg\mathfrak{m} + 1) + (|S| - 1) = \deg\mathfrak{m} + |S| = k \qquad (\text{since } (\mathfrak{m}, S) \in \mathfrak{A}).$$

It remains to prove the third statement. But $x_S = x_{S\setminus\{j\}} x_j$ (since $j \in S$) and thus $x_{S\setminus\{j\}} = x_S/x_j$, so that $(\mathfrak{m}x_j)\, x_{S\setminus\{j\}} = (\mathfrak{m}x_j)(x_S/x_j) = \mathfrak{m}x_S$. Hence, $\max\left(\operatorname{Supp}\left((\mathfrak{m}x_j)\, x_{S\setminus\{j\}}\right)\right) = \max(\operatorname{Supp}(\mathfrak{m}x_S)) = j$ because of $(\mathfrak{m}, S) \in \mathfrak{A}$. This proves the third of the three statements we needed.

Hence, $(\mathfrak{m}x_j, S \setminus \{j\}) \in \mathfrak{A}$ follows.

These two maps (14) and (15) are mutually inverse. Hence, they are bijections.

Now,

$$\underbrace{\sum_{\substack{(\mathfrak{m},S)\in\mathfrak{M}\times\mathcal{P}(n);\\ \deg\mathfrak{m}+|S|=k;\\ \max(\operatorname{Supp}(\mathfrak{m}x_S))=j}}}_{\substack{=\sum_{(\mathfrak{m},S)\in\mathfrak{A}}\\ \text{(by the definition of }\mathfrak{A})}} (-1)^{|S|}\,\mathfrak{m}x_S$$

$$= \sum_{(\mathfrak{m},S)\in\mathfrak{A}} (-1)^{|S|}\,\mathfrak{m}x_S$$

$$= \sum_{\substack{(\mathfrak{m},S)\in\mathfrak{A};\\ j\in S}} (-1)^{|S|}\,\mathfrak{m}x_S + \sum_{\substack{(\mathfrak{m},S)\in\mathfrak{A};\\ j\notin S}} (-1)^{|S|}\,\mathfrak{m}x_S$$

$$\text{(since each } (\mathfrak{m},S)\in\mathfrak{A} \text{ satisfies } j\in S \text{ or } j\notin S \text{ but never both)}$$

$$= \sum_{\substack{(\mathfrak{m},S)\in\mathfrak{A};\\ j\notin S}} \underbrace{(-1)^{|S\cup\{j\}|}}_{\substack{=(-1)^{|S|+1}\\ \text{(since } j\notin S\\ \text{entails } |S\cup\{j\}|=|S|+1)}} \left(\mathfrak{m}/x_j\right) \underbrace{x_{S\cup\{j\}}}_{\substack{=x_Sx_j\\ \text{(since } j\notin S)}} + \sum_{\substack{(\mathfrak{m},S)\in\mathfrak{A};\\ j\notin S}} (-1)^{|S|}\,\mathfrak{m}x_S$$

$$\left(\begin{array}{c}\text{here, we have substituted } \left(\mathfrak{m}/x_j, S\cup\{j\}\right) \text{ for } (\mathfrak{m},S)\\ \text{in the first sum, since the map (14) is a bijection}\end{array}\right)$$

$$= \sum_{\substack{(\mathfrak{m},S)\in\mathfrak{A};\\ j\notin S}} \underbrace{(-1)^{|S|+1}}_{=-(-1)^{|S|}} \underbrace{\left(\mathfrak{m}/x_j\right)x_Sx_j}_{=\mathfrak{m}x_S} + \sum_{\substack{(\mathfrak{m},S)\in\mathfrak{A};\\ j\notin S}} (-1)^{|S|}\,\mathfrak{m}x_S$$

$$= -\sum_{\substack{(\mathfrak{m},S)\in\mathfrak{A};\\ j\notin S}} (-1)^{|S|}\,\mathfrak{m}x_S + \sum_{\substack{(\mathfrak{m},S)\in\mathfrak{A};\\ j\notin S}} (-1)^{|S|}\,\mathfrak{m}x_S = 0. \tag{16}$$

Forget that we fixed j. We thus have proved (16) for each $j\in\{k,k+1,\ldots,n\}$. Hence, (12) becomes

$$\sum_{\substack{(\mathfrak{m},S)\in\mathfrak{M}\times\mathcal{P}(n);\\ \deg\mathfrak{m}+|S|=k}} (-1)^{|S|}\,\mathfrak{m}x_S = \sum_{j=k}^{n} \underbrace{\sum_{\substack{(\mathfrak{m},S)\in\mathfrak{M}\times\mathcal{P}(n);\\ \deg\mathfrak{m}+|S|=k;\\ \max(\operatorname{Supp}(\mathfrak{m}x_S))=j}} (-1)^{|S|}\,\mathfrak{m}x_S}_{\substack{=0\\ \text{(by (16))}}} = \sum_{j=k}^{n} 0 = 0.$$

Thus, (9) rewrites as

$$\sum_{i=0}^{k} (-1)^i\, h_{k-i}\left(x_{k..n}\right)\sigma_i\left(x_{1..n}\right) = 0.$$

This proves (5), and thus proves (1) as well. ■

13. **pages 12–13, proof of Theorem 1.2.7:** The uniqueness of the reduced Gröbner basis of I is not proved here. But it follows from a general fact saying that any ideal of a polynomial ring over a field has a unique reduced Gröbner basis (at least unique up to scaling; or literally unique, if we require the elements of a Gröbner basis to have leading coefficients 1). For a proof of this fact, see, for example, Satz 2.6.4 in Birgit Reinert, *Gröbnerbasen*, Wintersemester 1996/1997, or Theorem 5 in §2.7 of David A. Cox, John Little, Donal O'Shea, *Ideals, Varieties, and Algorithms*, 5th edition, Springer 2025.

14. **page 13, before the Exercises:** "monomials $\sigma_1^{i_1}\sigma_2^{i_2}\cdots\sigma_n^{i_n}$ in the elementary symmetric polynomial" $\to$ "monomials $\sigma_1^{i_1}\sigma_2^{i_2}\cdots\sigma_n^{i_n}$ in the elementary symmetric polynomials".

15. **page 17, Example 1.3.4:** "has distance 7" should be "has distance $\sqrt{7}$".

16. **page 19, §1.4:** The "*monoid defined by* $\mathcal{A}$" is a semigroup, not a monoid (since it does not contain $(0,0,\ldots,0)$ and thus has no neutral element).

17. **page 20:** "The invariant monomials are in bijection with the elements of the monoid $\mathcal{M}_{\mathcal{A}}$" should be "The invariant monomials are in bijection with the elements of the monoid $\mathcal{M}_{\mathcal{A}} \cup \{0\}$" (since $\mathcal{M}_{\mathcal{A}}$ itself is not a monoid).

18. **page 20, Lemma 1.4.2:** Replace "$\in \mathcal{M}_{\mathcal{A}}$" by "$\in \mathcal{M}_{\mathcal{A}} \cup \{0\}$".

19. **page 20, proof of Lemma 1.4.2:** Replace "of the monoid $\mathcal{M}_{\mathcal{A}}$" by "of the monoid $\mathcal{M}_{\mathcal{A}} \cup \{0\}$".

20. **page 21, proof of correctness for Algorithms 1.4.3 and 1.4.4:** Here it is claimed that "In each step in the reduction of $\mathbf{x}^\beta$ a monomial reduces to another monomial". This tacitly uses the fact that the reduced Gröbner basis $\mathcal{G}$ of I consists of "monomial differences" (i.e., of polynomials of the form $\mathfrak{m} - \mathfrak{n}$ where $\mathfrak{m}$ and $\mathfrak{n}$ are two monomials). To prove this fact, it suffices to show the same about $\mathcal{G}'$ (since $\mathcal{G} \subseteq \mathcal{G}'$). But $\mathcal{G}'$ is the reduced Gröbner basis of an ideal of $\mathbf{C}\left[t_0, t_1, \ldots, t_d, x_1, \ldots, x_n\right]$, and the latter ideal is generated by "monomial differences" (namely, the differences $t_0 t_1 \cdots t_d - 1$ and $x_i - \prod_{j=1}^{d} t_j^{a_{ij}}$ for $i = 1, 2, \ldots, n$, where all negative powers $t_j^{a_{i,j}}$ in the latter monomials should be replaced by $\left(t_0 t_1 \cdots t_{j-1} t_{j+1} t_{j+2} \cdots t_d\right)^{-a_{i,j}}$). It is easy to see that if a ideal of a polynomial ring is generated by "monomial differences", then its reduced Gröbner basis (with respect to any monomial order) consists of "monomial differences". Thus, $\mathcal{G}'$ consists of "monomial differences", and therefore so does $\mathcal{G}$.

21. **page 25, proof of Proposition 2.1.1:** The polynomial P_i is not "monic", but very close (its leading coefficient is $(-1)^{|\Gamma|}$, which is invertible; thus, the polynomial $(-1)^{|\Gamma|} P_i$ is monic).

22. **page 25, proof of Proposition 2.1.1:** "Hence the invariant subring $\mathbf{C}[\mathbf{x}]^{\Gamma}$ and the full polynomial ring $\mathbf{C}[\mathbf{x}]$ have the same transcendence degree n over the ground field $\mathbf{C}$": This argument seems to tacitly use the facts that

    a) the polynomial ring $\mathbf{C}[\mathbf{x}]$ has transcendence degree n over the ground field $\mathbf{C}$, and

    b) if $\mathbf{C} \subseteq A \subseteq B$ is an inclusion of integral domains such that B is integral over A, then the transcendence degrees of A and B are equal.

    I know how to prove the first of these facts (indeed, it would follow from Lemma 4.7.2 if not for the homogeneity requirement in the latter lemma; but this homogeneity requirement can be easily removed by introducing an extra slack variable), but I'm less sure about the second.

23. **page 26, proof of Theorem 2.1.3:** This proof tacitly uses the fact that if $I \in \mathbf{C}[\mathbf{x}]$ is an invariant of Γ, then all homogeneous components of I are invariants of Γ as well. (This is easy, since the action of Γ does not change the degree of a homogeneous polynomial.) This fact allows us to restrict ourselves to homogeneous invariants.

24. **page 26, proof of Theorem 2.1.3:** "Hence there exist finitely many *homogeneous* invariants $I_1, I_2, \ldots, I_m$ such that $\mathcal{I}_\Gamma = \langle I_1, I_2, \ldots, I_m \rangle$". Let me explain in more detail why this holds: Indeed, this is a consequence of the following general fact:

    > *Fact.* If an ideal J of some commutative ring is finitely generated, and if G is any generating set of J (not necessarily finite), then there exists a **finite** subset of G that already generates J.

    (To prove this general fact, pick a finite generating set K of J, and expand each $k \in K$ in terms of the generators in G. Only finitely many elements of G altogether are used in these expansions, and so they form a finite subset of G that already generates J.)

25. **page 26, proof of Theorem 2.1.3:** "Since $I \in \mathcal{I}_\Gamma$, we have $I = \sum\limits_{j=1}^{s} f_j I_j$ for some homogeneous polynomials $f_j \in \mathbf{C}[\mathbf{x}]$ of degree less than $\deg(I)$": More precisely, this follows from the fact that $I \in \mathcal{I}_\Gamma$ is homogeneous and that the polynomials $I_1, I_2, \ldots, I_m$ are also homogeneous of positive degrees. Namely, the fact that $I \in \mathcal{I}_\Gamma = \langle I_1, I_2, \ldots, I_m \rangle$ shows that I can be written as $I = \sum\limits_{j=1}^{s} g_j I_j$ for some polynomials $g_j \in \mathbf{C}[\mathbf{x}]$ that are not necessarily homogeneous. But now, projecting this equality onto the $\deg I$-th homogeneous component of $\mathbf{C}[\mathbf{x}]$, we conclude that $I = \sum\limits_{j=1}^{s} f_j I_j$, where f_j is the $\left(\deg I - \deg I_j\right)$-th homogeneous component of g_j; and this is the expansion we are looking for.

26. **page 26:** In "the remarkable statement that every *ideal basis* $\{I_1, \ldots, I_m\}$ of $\mathcal{I}_\Gamma$ is automatically an *algebra basis* for $\mathbf{C}[\mathbf{x}]^\Gamma$", add the word "homogeneous" in front of "*ideal basis*" (though arguably, the whole statement is probably meant informally, since there is no definition of "ideal basis" anywhere in the book).

    It is worth remarking that in some places (e.g., in Proposition 2.1.5), "algebra basis" means "graded generating set of $\mathbf{C}[\mathbf{x}]^\Gamma$" (that is, "generating set consisting of homogeneous elements").

27. **page 26:** "the finiteness of the group Γ has not been used until the last paragraph" is not literally true: The Reynolds operator $*$ has already made an appearance in the first paragraph of the proof.

28. **page 28, Proposition 2.1.5:** Replace "$n, p \geq 2$" by "$n, p \geq 1$".

29. **page 29, Theorem 2.2.1:** Add "Let $\Gamma \subset GL(\mathbf{C}^n)$ be a finite matrix group." at the beginning of this theorem.

30. **page 29, proof of Theorem 2.2.1:** The word "d-form" means "homogeneous polynomial of degree d" here.

31. **page 30, proof of Theorem 2.2.1:** When applying Lemma 2.2.2 here, one should be careful: As stated, Lemma 2.2.2 would yield a sum over all the elements of the group $\left\{\pi^{(d)} \mid \pi \in \Gamma\right\}$, not over all the elements of Γ. This difference sometimes matters, since different π's in Γ might lead to the same $\pi^{(d)}$'s. The cleanest way to correct this little discrepancy is to generalize Lemma 2.2.2 by replacing the subgroup $\Gamma \subset GL(\mathbf{C}^n)$ by an arbitrary finite group Γ that acts on $\mathbf{C}^n$. (Of course, $\operatorname{trace}(\pi)$ must then be understood as the trace of the action of π on $\mathbf{C}^n$.) If we generalize Lemma 2.2.2 this way, then we can apply it to the group Γ acting on $\mathbf{C}[\mathbf{x}]_d^\Gamma$ via $\pi \mapsto \pi^{(d)}$, and this immediately yields

    $$\dim\left(\mathbf{C}[\mathbf{x}]_d^\Gamma\right) = \frac{1}{|\Gamma|}\sum_{\pi\in\Gamma} \operatorname{trace}\left(\pi^{(d)}\right) = \frac{1}{|\Gamma|}\sum_{\pi\in\Gamma}\ \sum_{d_1+\cdots+d_n=d} \rho_{\pi,1}^{d_1}\cdots\rho_{\pi,n}^{d_n},$$

    as desired.

32. **page 30, Lemma 2.2.3:** Replace "$\sum_{n=0}^{\infty}$" by "$\sum_{d=0}^{\infty}$".

33. **page 31, proof of Example 2.2.4:** "Molien series" just means "Hilbert series" here (computed using Molien's formula, i.e., Theorem 2.2.1).

34. **page 37, §2.3:** The claims about Krull dimensions and h.s.o.p.s made here need some further assumptions. Clearly, the algebra R must be commutative, but even this does not seem to be enough; e.g., the algebra

$\mathbf{C}[x_1, x_2, x_3, \ldots] / (x_1^2, x_2^2, x_3^2, \ldots)$ has no algebraically independent elements at all, but is not finitely generated as a module over its subring **C**, so it is **not** free over a subalgebra generated by an h.s.o.p., at least not if h.s.o.p.s are defined as they are here.

The claim that the maximal number of algebraically independent elements of **C** is the Krull dimension of R is true when R is finitely generated as a **C**-algebra, according to Theorem 5.9 in Kemper's book (Gregor Kemper, *A Course in Commutative Algebra*, Springer 2011).

35. **page 41, Proposition 2.3.6 (ii):** Replace "$\phi_\Gamma(z)$" by "$\Phi_\Gamma(z)$".

36. **pages 44–50, §2.4:** For most of the book (and, in particular, for all of the text before §2.4), the group Γ has always been acting from the left on $\mathbf{C}^n$, and thus from the right on the polynomial ring $\mathbf{C}[V]$ (since the polynomials in $\mathbf{C}[V]$ are viewed as polynomial functions from V to **C**). However, in §2.4, the group Γ suddenly acts from the left on $\mathbf{C}[V]$ instead (as witnessed, e.g., in the notation "σf" in Lemma 2.4.2). Here are two ways how this can be reconciled with the rest of the book:

    - One way is to replace all the "σf"s (for $\sigma \in \Gamma$ and $f \in \mathbf{C}[V]$) in §2.4 by "$f \circ \sigma$"s. (Thus, for example, in the proof of Proposition 2.4.3, each "σh_1" should become a "$h_1 \circ \sigma$".) Also, in the proof of Proposition 2.4.3, on page 46, the computation "

    $$\pi h_1 - h_1 = \sum_{i=1}^{l-1} (\sigma_1 \ldots \sigma_i \sigma_{i+1} h_1 - \sigma_1 \ldots \sigma_i h_1) = \sum_{i=1}^{l-1} (\sigma_1 \ldots \sigma_i (\sigma_{i+1} h_1 - h_1)) \in \mathcal{I}_\Gamma$$

    " should be replaced by "

    $$h_1 \circ \pi - h_1 = \sum_{i=1}^{l-1} (h_1 \circ \sigma_i \sigma_{i+1} \ldots \sigma_l - h_1 \circ \sigma_{i+1} \sigma_{i+2} \ldots \sigma_l) = \sum_{i=1}^{l-1} ((h_1 \circ \sigma_i - h_1) \circ \sigma_{i+1} \sigma_{i+2} \ldots \sigma_l) \in \mathcal{I}_\Gamma$$

    ". This is likely the intended way.

    - An alternative way is to define $\sigma f := f \circ \sigma^{-1}$ for each $\sigma \in \Gamma$ and $f \in \mathbf{C}[V]$. This way, the right Γ-action on $\mathbf{C}[V]$ is "translated" into a left Γ-action on $\mathbf{C}[V]$ that carries the same information and has the same invariants (since a polynomial $f \in \mathbf{C}[V]$ and a group element $\sigma \in \Gamma$ satisfy $\sigma f = f$ if and only if $f \circ \sigma^{-1} = f$, that is, if and only if $f \circ \sigma = f$).

37. **page 44, §2.4:** The definition of a reflection given here (as a linear transformation $\pi \in GL(\mathbf{C}^n)$ such that "precisely one eigenvalue of π is not equal to one") is somewhat eccentric. In the context of a finite reflection group, it does its job well; however, viewed in isolation, it does not agree with any of the commonly used definitions. Normally, one defines a reflection (or pseudo-reflection) to be a linear transformation $\pi \in GL(\mathbf{C}^n)$ such that $\dim(\operatorname{Ker}(\pi - id)) = n - 1$. Some authors also require π to be of finite order, but this is automatically satisfies when π is an element of a finite group. The fact that precisely one eigenvalue of π is not equal to one follows automatically from the condition $\dim(\operatorname{Ker}(\pi - id)) = n - 1$, but the converse implication again requires the assumption that π is an element of a finite group (this ensures that π is diagonalizable, so that all the $n-1$ eigenvalues equal to 1 cause $\operatorname{Ker}(\pi - id)$ to have dimension $n-1$). If not for the latter assumption, the matrix $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ would be a "reflection" in the sense of the book (a linear transformation $\pi \in GL(\mathbf{C}^n)$ such that precisely one eigenvalue of π is not equal to one), but would satisfy rather little of what a reflection is commonly expected to satisfy.

38. **page 45, proof of Lemma 2.4.2:** The use of Hilbert's Nullstellensatz here is overkill. The only thing needed is the following easy fact:

    > *Fact:* Let $L \in \mathbf{C}[\mathbf{x}]$ be a homogeneous linear polynomial (i.e., a homogeneous polynomial of degree 1). Let $g \in \mathbf{C}[\mathbf{x}]$ be a polynomial such that each $\mathbf{v} \in \mathbf{C}^n$ satisfying $L(\mathbf{v}) = 0$ satisfies $g(\mathbf{v}) = 0$. Then, L is a divisor of g in $\mathbf{C}[\mathbf{x}]$.

    (*Proof of the fact:* By an appropriate coordinate transformation (i.e., composition with some invertible matrix $\pi \in GL(\mathbf{C}^n)$), we can ensure that the linear polynomial L is simply x_1. Thus, WLOG assume that $L = x_1$. Then, the condition "each $\mathbf{v} \in \mathbf{C}^n$ satisfying $L(\mathbf{v}) = 0$ satisfies $g(\mathbf{v}) = 0$" becomes "g vanishes on each vector whose 1-st coordinate is 0", that is, "$g(0, x_2, x_3, \ldots, x_n) = 0$". But this entails that all monomials in g that have degree 0 with respect to x_1 have coefficient 0, and thus all monomials that actually appear in g must have x_1 in some nonzero power. Consequently, g is divisible by x_1, that is, by L. In other words, L is a divisor of g. This proves the fact.)

39. **page 45, proof of Proposition 2.4.3:** The claim that "$\widetilde{h}_1 \cdot L_\sigma \in \mathcal{I}_\Gamma$" at the end of page 45 is not entirely obvious: It requires showing that $\deg \widetilde{h}_1 > 0$ (since $\mathcal{I}_\Gamma$ is the ideal generated by all homogeneous invariants **of positive degree**). But this is easy: If $\widetilde{h}_1 = 0$, then $\widetilde{h}_1 \cdot L_\sigma = 0 \in \mathcal{I}_\Gamma$ is obvious; if $\widetilde{h}_1$ is a nonzero constant, then $g_1 \widetilde{h}_1 + g_2 \widetilde{h}_2 + \cdots + g_m \widetilde{h}_m = 0$ entails $g_1 = -\frac{1}{\widetilde{h}_1}\left(g_2 \widetilde{h}_2 + \cdots + g_m \widetilde{h}_m\right) \in \langle g_2, \ldots, g_m \rangle$, contradicting $g_1 \notin \langle g_2, \ldots, g_m \rangle$.

40. **page 46, proof of Theorem 2.4.1 (if-part):** The first sentence of this proof again uses the following fact (already mentioned above):

    > *Fact.* If an ideal J of some commutative ring is finitely generated, and if G is any generating set of J (not necessarily finite), then there exists a **finite** subset of G that already generates J.

    Here this fact is being applied to $J = \mathcal{I}_\Gamma$ and $G = \{\text{homogeneous invariants of } \Gamma\}$.

41. **page 46, proof of Theorem 2.4.1 (if-part):** "We need to prove that $m = n$, or, equivalently, that the invariants $f_1, f_2, \ldots, f_m$ are algebraically independent over $\mathbf{C}$": I don't see why these two claims are equivalent. The proof given here is showing the second claim (i.e., that the invariants $f_1, f_2, \ldots, f_m$ are algebraically independent over $\mathbf{C}$). The first claim $m = n$ can then be derived from this by asymptotically comparing the Hilbert series (essentially following the proof of Corollary 2.4.5, but without assuming that the number of generators is n [4]).

42. **page 47, proof of Theorem 2.4.1 (if-part):** "Euler's formula" is the (easily verified) fact that any homogeneous polynomial $f \in \mathbf{C}[\mathbf{x}]$ satisfies $\sum_{s=1}^{n} x_s \dfrac{\partial f}{\partial x_s} = (\deg f)\, f$.

43. **page 47, Lemma 2.4.4:** Talking about "the Laurent expansion of the Molien series about $z = 1$", it is important to keep in mind that the Laurent series is a rational function (not just a formal power series), by Theorem 2.2.1, and therefore can be expanded into a Laurent series about any complex number.

44. **page 49, proof of Theorem 2.4.2 (only-if-part):** The claim that "the Jacobian determinant $\det\left(\partial\theta_i/\partial\psi_j\right)$ is nonzero" could use a bit of explanation. It relies on the so-called *Jacobian criterion*, which says that n polynomials $p_1, p_2, \ldots, p_n$ in the polynomial ring $\mathbf{C}[\mathbf{x}] = \mathbf{C}[x_1, x_2, \ldots, x_n]$ are algebraically independent over $\mathbf{C}$ if and only if their Jacobian $\det\left(\partial p_i/\partial x_j\right)$ is nonzero. Here, this criterion is being applied not to the polynomial ring $\mathbf{C}[\mathbf{x}] = \mathbf{C}[x_1, x_2, \ldots, x_n]$ but rather to its subring $\mathbf{C}[\psi_1, \psi_2, \ldots, \psi_n]$ (which, too, is a polynomial ring, since the polynomials $\psi_1, \psi_2, \ldots, \psi_n$ are algebraically independent) and to the n polynomials $\theta_1, \theta_2, \ldots, \theta_n$ therein.

---

[4] That is, we need the following generalization of Corollary 2.4.5:

> *Corollary 2.4.5′.* Let $\Gamma \subset GL(\mathbf{C}^n)$ be a finite matrix group whose invariant ring $\mathbf{C}[\mathbf{x}]^\Gamma$ is generated by m algebraically independent homogeneous invariants $\theta_1, \ldots, \theta_m$ where $d_i := \deg \theta_i$. Let r be the number of reflections in Γ. Then,
>
> $$m = n \qquad \text{and} \qquad |\Gamma| = d_1 d_2 \cdots d_n \qquad \text{and} \qquad r = d_1 + d_2 + \cdots + d_n - n.$$

The proof of this generalization is just a slight modification of the proof of the original Corollary 2.4.5.

For a proof of the Jacobian criterion, see, e.g., Theorem 2.2 in the paper Richard Ehrenborg, Gian-Carlo Rota, *Apolarity and Canonical Forms for Homogeneous Polynomials*, European Journal of Combinatorics **14**, Issue 3, May 1993, pp. 157–181. (This proof relies on the fact that if $p_1, p_2, \ldots, p_n$ are n algebraically independent polynomials in $\mathbf{C}[\mathbf{x}]$, then each of the $(n+1)$-tuples $(x_i, p_1, p_2, \ldots, p_n)$ is algebraically dependent. This follows from Lemma 4.7.2 in Sturmfels's book, after some tweaking to make the polynomials homogeneous.)

45. **page 52:** The formula "$f^* := \frac{1}{|\Gamma|} \sum_{\sigma \in \Gamma} \sigma(f)$" should be "$f^* := \frac{1}{|\Gamma|} \sum_{\sigma \in \Gamma} f \circ \sigma$" (since the group Γ acts on $\mathbf{C}[\mathbf{x}]$ from the right, not from the left).

46. **page 53, proof of Lemma 2.5.7:** This is unnecessary. After all, Lemma 2.5.7 simply follows from Proposition 2.6.4 (applied to M instead of I).

47. **page 55, proof of Lemma 2.5.11:** In the first displayed equation, replace "$\prod_{i=1}^{n} \left(1 - z^{d_j}\right)$" by either "$\prod_{i=1}^{n} \left(1 - z^{d_i}\right)$" or "$\prod_{j=1}^{n} \left(1 - z^{d_j}\right)$".

48. **page 56:** "where $\alpha \in$ ranges" should be "where α ranges".

49. **page 57, Algorithm 2.5.14:** In step 0, replace "$\Phi(z)$" by "$\Phi_\Gamma(z)$".

50. **page 57, Algorithm 2.5.14:** In step 2, replace "$\prod_{i=1}^{n} \left(1 - z^{d_j}\right)$" by either "$\prod_{i=1}^{n} \left(1 - z^{d_i}\right)$" or "$\prod_{j=1}^{n} \left(1 - z^{d_j}\right)$".

51. **page 59, Algorithm 2.6.2:** "Gröbner basis $\mathcal{G}_1$ for $\mathcal{F} \cup \mathcal{G}_0$" should be "Gröbner basis $\mathcal{G}_1$ for $\langle \mathcal{F} \cup \mathcal{G}_0 \rangle$".

52. **page 61, proof of Proposition 2.6.4:** This proof can be significantly simplified, removing the use of the Nullstellensatz:

    Simpler proof of Proposition 2.6.4. Since $I' \subseteq I$, we clearly have $\operatorname{Rad}(I') \subseteq \operatorname{Rad}(I)$. It remains to show that $\operatorname{Rad}(I) \subseteq \operatorname{Rad}(I')$. For this purpose, it suffices to show that $I \subseteq \operatorname{Rad}(I')$ (since this would entail $\operatorname{Rad}(I) \subseteq \operatorname{Rad}(\operatorname{Rad}(I')) = \operatorname{Rad}(I')$).

    So let $f \in I$. We must show that $f \in \operatorname{Rad}(I')$. We shall show that $f^{|\Gamma|} \in I'$; this will clearly do the trick.

    Consider the polynomial $\prod_{\sigma \in \Gamma} (z - f(\sigma \mathbf{x})) \in (\mathbf{C}[\mathbf{x}])[z]$ in the new indeterminate z over $\mathbf{C}[\mathbf{x}]$. This is clearly a monic polynomial of degree $|\Gamma|$ in z, hence can be written as

    $$\prod_{\sigma \in \Gamma} (z - f(\sigma \mathbf{x})) = z^{|\Gamma|} + \sum_{j=0}^{|\Gamma|-1} p_j(\mathbf{x}) z^j, \qquad (17)$$

where the $p_j(\mathbf{x}) \in \mathbf{C}[\mathbf{x}]$ are its coefficients (and are themselves polynomials in $\mathbf{x}$). Consider these $p_j(\mathbf{x})$. Furthermore, the polynomial $\prod_{\sigma\in\Gamma}(z - f(\sigma\mathbf{x}))$ is invariant under the action of Γ (since the action of Γ merely permutes the factors of the product $\prod_{\sigma\in\Gamma}(z - f(\sigma\mathbf{x}))$). Hence, all its coefficients $p_j(\mathbf{x})$ are invariant under Γ as well.

Now, let $j \in \{0, 1, \ldots, |\Gamma| - 1\}$ be arbitrary. Then, as we just showed, $p_j(\mathbf{x})$ is invariant under Γ. Furthermore, (17) shows that $p_j(\mathbf{x})$ equals (up to sign) the $(|\Gamma| - j)$-th elementary symmetric polynomial evaluated at the $|\Gamma|$ many inputs $f(\sigma\mathbf{x})$ for $\sigma \in \Gamma$; thus, in particular, it is a polynomial in these $|\Gamma|$ many inputs $f(\sigma\mathbf{x})$ with no constant term (since $j < |\Gamma|$ entails $|\Gamma| - j > 0$). Hence, $p_j(\mathbf{x})$ belongs to the ideal I (since all the $f(\sigma\mathbf{x})$ belong to I (because f belongs to I, and because I is Γ-invariant)). Since $p_j(\mathbf{x})$ is furthermore invariant under Γ, we thus conclude that $p_j(\mathbf{x})$ is an invariant in I. Hence, $p_j(\mathbf{x}) \in I'$ (by the definition of I').

Forget that we fixed j. We thus have shown that

$$p_j(\mathbf{x}) \in I' \qquad \text{for each } j \in \{0, 1, \ldots, |\Gamma| - 1\}. \tag{18}$$

Now, substituting $f = f(\mathbf{x})$ for z on both sides of (17), we obtain

$$\prod_{\sigma\in\Gamma}(f - f(\sigma\mathbf{x})) = f^{|\Gamma|} + \sum_{j=0}^{|\Gamma|-1} p_j(\mathbf{x}) f^j.$$

Since the product $\prod_{\sigma\in\Gamma}(f - f(\sigma\mathbf{x}))$ is 0 (because one of the factors of this product is $f - f(\operatorname{id}\mathbf{x}) = f - f(\mathbf{x}) = f - f = 0$), this can be rewritten as

$$0 = f^{|\Gamma|} + \sum_{j=0}^{|\Gamma|-1} p_j(\mathbf{x}) f^j.$$

Hence,

$$f^{|\Gamma|} = -\sum_{j=0}^{|\Gamma|-1} \underbrace{p_j(\mathbf{x})}_{\substack{\in I' \\ (\text{by } (18))}} f^j \in I' \qquad \left(\text{since } I' \text{ is an ideal}\right),$$

and therefore $f \in \operatorname{Rad}(I')$, qed. ■

53. **page 63, proof of Proposition 2.6.6:** "By Hilbert's Nullstellensatz" → "By Hilbert's Nullstellensatz, this shows that $\mathbf{C}[\mathbf{x}] = I + \mathcal{I}(\Gamma\mathbf{a})$, where $\Gamma\mathbf{a}$ is the ideal of polynomials vanishing on the finite set $\Gamma\mathbf{a}$. Hence, by the Chinese Remainder Theorem, $\mathbf{C}[\mathbf{x}] / (I)$".

(Alternatively, we can avoid the use of Hilbert's Nullstellensatz by defining a maximal ideal

$$\mathcal{I}_{\mathbf{b}} := \langle x_1 - b_1,\ x_2 - b_2,\ \ldots,\ x_n - b_n \rangle$$

of $\mathbf{C}[\mathbf{x}]$ for each point $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \Gamma\mathbf{a}$. Then, any two distinct points $\mathbf{b}, \mathbf{c} \in \Gamma\mathbf{a}$ satisfy $\mathbf{C}[\mathbf{x}] = \mathcal{I}_{\mathbf{b}} + \mathcal{I}_{\mathbf{c}}$, and furthermore $\Gamma\mathbf{a} \cap \mathcal{V}(I) = \varnothing$ shows that all $\mathbf{b} \in \Gamma\mathbf{a}$ satisfy $\mathbf{C}[\mathbf{x}] = I + \mathcal{I}_{\mathbf{b}}$ (since $\mathbf{b} \notin \mathcal{V}(I)$ and thus $I \not\subseteq \mathcal{I}_{\mathbf{b}}$). Hence, the Chinese Remainder Theorem can be applied to the ideals $I, \mathcal{I}_{\mathbf{b}_1}, \mathcal{I}_{\mathbf{b}_2}, \ldots, \mathcal{I}_{\mathbf{b}_m}$ where $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m\} := \Gamma\mathbf{a}$.)

54. **page 69, Algorithm 2.7.3, step 3:** "solution monoid" → "solution monoid $\mathcal{F} \subset \mathbf{N}^n$".

55. **page 71:** "This means that each factor group Γ_i/Γ_{i+1} is cyclic of prime order p_i": This is not what a composition series means. Rather, it is an extra requirement that you want to impose here.

56. **page 71, algorithm:** "It follows from Theorem 2.3.5 that $\mathbf{C}[\mathbf{x}]^{\Gamma_{i+1}}$ is a free module of rank p_i over $\mathbf{C}[\mathbf{x}]^{\Gamma_i}$" is not true in general. For example, in Example 2.7.6, $\mathbf{C}[\mathbf{x}]^{\{\mathrm{id}\}}$ is not a free module over $\mathbf{C}[\mathbf{x}]^{\Gamma = A_3}$.

57. **page 72, Remark 2.7.7:** The meaning of "cycle type $\ell(\sigma) = (\ell_1, \ell_2, \ldots, \ell_n)$" should be explained (since the notation is quite nonstandard). It means that σ has exactly ℓ_i cycles of length i for each $i \in \{1, 2, \ldots, n\}$.

58. **page 81, second paragraph:** What is called a "standard tableau" here is better known as a "transpose semistandard tableau of rectangular shape" to combinatorialists (although invariant theorists often do call it "standard tableau").

59. **page 81, third paragraph:** In the definition of "straightening syzygy", the condition "$\beta_s < \gamma_1$" is understood to be tautologically true if γ_1 does not exist (i.e., if $s = d$).

60. **page 82, proof of Theorem 3.1.7:** Lemma 3.1.8 is not entirely obvious. Its proof requires checking three facts:

Fact 1: Sorting each column in an arbitrary tableau T yields a standard tableau $\widetilde{T}$.

Fact 2: This resulting tableau $\widetilde{T}$ satisfies $\operatorname{init} \phi_{n,d}\left(\widetilde{T}\right) = \prod_{i=1}^{k} x_{\lambda_1^i 1} x_{\lambda_2^i 2} \cdots x_{\lambda_d^i d}$.

Fact 3: This tableau $\widetilde{T}$ is the only standard tableau S such that $\operatorname{init} \phi_{n,d}(S) = \prod_{i=1}^{k} x_{\lambda_1^i 1} x_{\lambda_2^i 2} \cdots x_{\lambda_d^i d}$.

Fact 1 is a version of the *non-messing-up lemma*, saying that if the entries of an integer matrix are strictly increasing along the rows, then sorting each column of the matrix will result in a matrix whose entries are again strictly increasing along the rows (as well as weakly increasing down the columns, which is clear because we just sorted them). Proving this is a nice and easy exercise[5]. Fact 2 is clear, because if we write $\widetilde{T}$ as $\widetilde{T} = \left[\widetilde{\lambda}^1\right]\left[\widetilde{\lambda}^2\right]\cdots\left[\widetilde{\lambda}^k\right]$, then the standardness of $\widetilde{T}$ (see Fact 1) yields

$$\begin{aligned}
\operatorname{init}\phi_{n,d}\left(\widetilde{T}\right) &= \prod_{i=1}^{k} x_{\widetilde{\lambda}^i_1 1} x_{\widetilde{\lambda}^i_2 2} \cdots x_{\widetilde{\lambda}^i_d d} = \prod_{i=1}^{k} \prod_{j=1}^{d} x_{\widetilde{\lambda}^i_j j} \\
&= \prod_{j=1}^{d} \underbrace{\prod_{i=1}^{k} x_{\widetilde{\lambda}^i_j j}}_{\substack{=\prod_{i=1}^{k} x_{\lambda^i_j j} \\ \text{(since the entries in the } j\text{-th column of } \widetilde{T} \\ \text{are the same as those in the } j\text{-th column of } T)}} = \prod_{j=1}^{d} \prod_{i=1}^{k} x_{\lambda^i_j j} \\
&= \prod_{i=1}^{k} \prod_{j=1}^{d} x_{\lambda^i_j j} = \prod_{i=1}^{k} x_{\lambda^i_1 1} x_{\lambda^i_2 2} \cdots x_{\lambda^i_d d}.
\end{aligned}$$

Finally, it remains to prove Fact 3. It suffices to show that different standard tableaux S induce different monomials $\operatorname{init}\phi_{n,d}(S)$. But this is not hard: If we write a standard tableau S as $S = \left[\mu^1\right]\left[\mu^2\right]\cdots\left[\mu^k\right]$, then the monomial $\operatorname{init}\phi_{n,d}(S) = \prod_{i=1}^{k} x_{\mu^i_1 1} x_{\mu^i_2 2} \cdots x_{\mu^i_d d}$ uniquely determines which entries lie in which column of S (namely: for each $j \leq d$, the entries in the j-th column of S are exactly the integers k such that x_{kd} appears in this monomial, and their multiplicities in the j-th column are precisely their multiplicities in this monomial), and thus uniquely determines S itself (since S is standard, so the entries in each column appear in increasing order). In other words, different standard tableaux S induce different monomials $\operatorname{init}\phi_{n,d}(S)$. This yields Fact 3.

This all said, Lemma 3.1.8 is not actually used in the proof of Theorem 3.1.7 (though it is used later on). Indeed, in the last sentence of the proof of Theorem 3.1.7 ("This is a contradiction to Lemma 3.1.8"), the only thing that is actually being used is that different standard tableaux S induce

[5]The analogous claim with "strictly increasing" replaced by "weakly increasing" is Remark 6.42 in Darij Grinberg, *Notes on the combinatorial fundamentals of algebra*, arXiv:2008.09862v3. The version with "strictly increasing" can be proved in the exact same way, or reduced to the "weakly increasing" version by an appropriate tweak of the matrix (if we subtract j from each entry in the j-th column of a matrix, then its strictly increasing rows become weakly increasing, while the relative order of entries in a given column does not change).

different monomials init $\phi_{n,d}(S)$. This is a much simpler claim than Lemma 3.1.8 (and we have proved it above, during our proof of Fact 3).

61. **page 88, proof of Lemma 3.2.5:** Replace "$\phi_{n+2d,n}$" by "$\phi_{n+2d,d}$" twice in this proof.

62. **page 88, proof of Lemma 3.2.5:** The matrix denoted by $\operatorname{Adj}(A)$ and called the "adjoint matrix of A" in this proof is not actually the adjoint matrix of A in the standard meaning of this word, but rather the transpose of the adjoint matrix of A. However, this does not affect the argument, since its determinant is the same as that of the actual adjoint matrix.

63. **page 88, proof of Lemma 3.2.5:** The computation (specifically, the equality "$\det(A)^{p(d-1)} \cdot I(x_{ij}) = \det(\operatorname{Adj}(A))^p \cdot I(x_{ij})$") relies on the classical result that $\det(\operatorname{Adj}(A)) = (\det(A))^{d-1}$ (see, for instance, Theorem 5.12 **(a)** in Darij Grinberg, *The trace Cayley-Hamilton theorem*, arXiv:2510.20689v1).

64. **page 89, proof of Theorem 3.2.1:** In the first paragraph of this proof, the straightening algorithm is applied not to the polynomials

$$[b_1 b_2 \cdots b_d]^{p(d-1)} \cdot I([a_1 \ldots a_{j-1} x_i a_{j+1} \ldots a_d]) \qquad \text{and}$$
$$[a_1 a_2 \cdots a_d]^{p(d-1)} \cdot I([b_1 \ldots b_{j-1} x b_{j+1} \ldots b_d]),$$

but rather to the polynomials

$$I([a_1 \ldots a_{j-1} x_i a_{j+1} \ldots a_d]) \qquad \text{and}$$
$$I([b_1 \ldots b_{j-1} x b_{j+1} \ldots b_d]).$$

The factors $[b_1 b_2 \cdots b_d]^{p(d-1)}$ and $[a_1 a_2 \cdots a_d]^{p(d-1)}$ are then multiplied onto the resulting expansions. This does not destroy the standardness of the tableaux in these expansions, because

- the factor $[b_1 b_2 \cdots b_d]^{p(d-1)}$ merely inserts $p(d-1)$ many rows of the form $[b_1 b_2 \cdots b_d]$ at the bottom of the standard tableaux appearing in the expansion of $I([a_1 \ldots a_{j-1} x_i a_{j+1} \ldots a_d])$ (and these new rows do not destroy the standardness of the tableaux, since each of the b_j's is larger than any of the letters $a_1, a_2, \ldots, a_d, x_1, x_2, \ldots, x_n$ that can appear in the expansion of $I([a_1 \ldots a_{j-1} x_i a_{j+1} \ldots a_d])$);
- the factor $[a_1 a_2 \cdots a_d]^{p(d-1)}$ merely inserts $p(d-1)$ many rows of the form $[a_1 a_2 \cdots a_d]$ at the top of the standard tableaux appearing in the expansion of $I([b_1 \ldots b_{j-1} x b_{j+1} \ldots b_d])$ (and these new rows do not destroy the standardness of the tableaux, since each of the a_j's is smaller than any of the letters $b_1, b_2, \ldots, b_d, x_1, x_2, \ldots, x_n$ that can appear in the expansion of $I([b_1 \ldots b_{j-1} x b_{j+1} \ldots b_d])$).

65. **page 89, proof of Theorem 3.2.1:** "both polynomials in question are equal to" should be "both polynomials in question evaluate in $\mathbf{C}\left[x_{ij}\right]$ to".

66. **page 90, third paragraph:** I think that the order "$\prec$" defined here is **not** the "diagonal order" from Sect. 3.1, but a close relative of it (the variable order is exactly the opposite one); fortunately all its important properties are analogous.

67. **page 91:** In the description of the "subduction algorithm", replace "$f - f_1^{\nu_1} \cdots f_m^{\nu_m}$" by "$f - c f_1^{\nu_1} \cdots f_m^{\nu_m}$, where c is the coefficient of $\operatorname{init}_\prec f$ in f" (don't forget that the leading coefficient is not necessarily 1).

68. **page 95, definition of an extensor:** An "*extensor (of step k)*" should be defined not as an element of the form $A = a_1 \vee a_2 \vee \cdots \vee a_k$ for some $a_1, \ldots, a_k \in V$, but rather as an element of the form $A = \lambda \cdot a_1 \vee a_2 \vee \cdots \vee a_k$ for some $a_1, \ldots, a_k \in V$ and $\lambda \in \mathbf{C}$. The scalar factor λ does not make a real difference when $k \geq 1$ (since it can be simply incorporated into the a_1 factor), but it becomes important for $k = 0$, where it allows any scalar (rather than just 1) to be considered as an extensor. And this is necessary for Theorem 3.3.2 **(b)** to be true (since the meet of two extensors of steps j and k with $j + k = d$ can be any scalar).

69. **page 96, definition of the meet $A \wedge B$:** It is not obvious that the meet operation $\wedge$ on $\Lambda(V)$ is well-defined. To prove this, we need to show that the right hand side of (3.3.6) is multilinear in the inputs $a_1, a_2, \ldots, a_j, b_1, b_2, \ldots, b_k$ as well as alternating in $a_1, a_2, \ldots, a_j$ and alternating in $b_1, b_2, \ldots, b_k$ (because then, by the universal property of the exterior powers $\Lambda^j(V)$ and $\Lambda^k(V)$, it will follow that this right hand side is a function of $a_1 a_2 \cdots a_j$ and $b_1 b_2 \cdots b_k$).

    The multilinearity is obvious (since the determinant form $[v_1, v_2, \ldots, v_d]$ is multilinear in its inputs $v_1, v_2, \ldots, v_d$, and since the wedge product $w_1 w_2 \cdots w_p$ is multilinear in its factors $w_1, w_2, \ldots, w_p$). The alternatingness in $b_1, b_2, \ldots, b_k$ is also clear (since the determinant form $[v_1, v_2, \ldots, v_d]$ is alternating in its inputs $v_1, v_2, \ldots, v_d$). It remains to show that the right hand side of (3.3.6) is alternating in $a_1, a_2, \ldots, a_j$. In other words, we must prove that if two of $a_1, a_2, \ldots, a_j$ are equal, then the right hand side of (3.3.6) is 0. Actually, it suffices to show that if some $p \in \{1, 2, \ldots, j-1\}$ satisfies $a_p = a_{p+1}$, then the right hand side of (3.3.6) is 0 (because it is well-known that this condition, combined with the multilinearity of the right hand side of (3.3.6), entails that this right hand side is alternating[6]).

[6] Let me explicitly state the fact that I am using here:

> *Fact:* Let V be a vector space over a field $\mathbf{k}$. Let $f : V^j \to \mathbf{k}$ be any multilinear form on V. If we have
>
> $$f(v_1, v_2, \ldots, v_j) = 0 \qquad \text{whenever some } p \in \{1, 2, \ldots, j-1\} \text{ satisfies } v_p = v_{p+1},$$

Let me prove this now. Assume that some $p \in \{1, 2, \ldots, j-1\}$ satisfies $a_p = a_{p+1}$. Consider this p. Now, the right hand side of (3.3.6) is

$$\sum_{\sigma} \operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)},$$

where the sum is over all permutations σ of $\{1, 2, \ldots, j\}$ such that $\sigma(1) < \sigma(2) < \cdots < \sigma(d-k)$ and $\sigma(d-k+1) < \sigma(d-k+2) < \cdots < \sigma(j)$. We shall refer to these permutations σ as the *shuffles*. For any permutation $\sigma \in S_j$ (thus, in particular, for any shuffle σ), we shall denote the sets $\{\sigma(1), \sigma(2), \ldots, \sigma(d-k)\}$ and $\{\sigma(d-k+1), \sigma(d-k+2), \ldots, \sigma(j)\}$ as L_σ and R_σ, respectively; we call them the *left half* and the *right half* of σ (although they are not really "halves" as they usually have different sizes). Clearly, any $\sigma \in S_j$ satisfies $L_\sigma \cap R_\sigma = \varnothing$ and $L_\sigma \cup R_\sigma = \{1, 2, \ldots, d\}$, so that each $i \in \{1, 2, \ldots, d\}$ lies in either L_σ or R_σ (but not in both). Hence, we can classify the permutations $\sigma \in S_j$ into four classes:

- *Class LL:* those that satisfy $p, p+1 \in L_\sigma$.
- *Class RR:* those that satisfy $p, p+1 \in R_\sigma$.
- *Class LR:* those that satisfy $p \in L_\sigma$ and $p+1 \in R_\sigma$.
- *Class RL:* those that satisfy $p \in R_\sigma$ and $p+1 \in L_\sigma$.

Each permutation $\sigma \in S_j$ is of exactly one of the four Classes LL, RR, LR and RL.

If $\sigma \in S_j$ is a permutation of Class LL, then

$$\begin{aligned} &\operatorname{sign}(\sigma) \underbrace{\left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right]}_{\substack{=0 \\ \text{(since this bracket contains} \\ \text{the equal vectors } a_p \text{ and } a_{p+1} \\ \text{(because } p, p+1 \in L_\sigma = \{\sigma(1), \sigma(2), \ldots, \sigma(d-k)\}))}} \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\ &= 0. \end{aligned} \tag{19}$$

then f is alternating (i.e., satisfies $f(v_1, v_2, \ldots, v_j) = 0$ whenever two of the j vectors $v_1, v_2, \ldots, v_j$ are equal).

For a proof of this fact, see Lemma 1.6 in Bernhard Leeb, *Some multilinear algebra*, January 25, 2020, or Corollary 2.9 in Keith Conrad, *Exterior powers*, 2026, or Proposition 6.4 in Jean Gallier and Jocelyn Quaintance, *Linear Algebra for Computer Vision, Robotics, and Machine Learning*, October 24, 2025, or Theorem 6 in Jordan Bell, *Alternating multilinear forms*, August 21, 2018 (but beware that the last two sources define "alternating" differently from us).

If $\sigma \in S_j$ is a permutation of Class RR, then

$$\operatorname{sign}(\sigma)\left[a_{\sigma(1)},\ldots,a_{\sigma(d-k)},b_1,\ldots,b_k\right]\cdot\underbrace{a_{\sigma(d-k+1)}\cdots a_{\sigma(j)}}_{\substack{=0\\ \text{(since this wedge product contains}\\ \text{the equal vectors } a_p \text{ and } a_{p+1}\\ \text{(because } p,p+1\in R_\sigma=\{\sigma(d-k+1),\sigma(d-k+2),\ldots,\sigma(j)\}))}}$$
$$=0. \tag{20}$$

Hence, in the sum on the right hand side of (3.3.6), all the addends corresponding to shuffles σ of Class LL and of Class RR are 0 and can thus be discarded. The remaining addends correspond to shuffles σ of Class LR and of Class RL. These addends are not (usually) 0, but rather can be paired up in such a way that any two addends in a pair cancel out; here is how this pairing works: Let $s_p \in S_j$ be the transposition that swaps p and $p+1$. Clearly, $s_p^2 = \mathrm{id}$ and $\operatorname{sign}(s_p) = -1$. If σ is a shuffle of Class LR, then $s_p\sigma \in S_j$ is a shuffle of Class RL[7]. Hence, we can define a map

$$\{\text{shuffles of Class LR}\} \to \{\text{shuffles of Class RL}\},\qquad \sigma \mapsto s_p\sigma. \tag{21}$$

Likewise, we can define a map

$$\{\text{shuffles of Class RL}\} \to \{\text{shuffles of Class LR}\},\qquad \sigma \mapsto s_p\sigma.$$

[7] *Proof.* Let σ be a shuffle of Class LR. Then, $s_p\sigma$ is obtained from σ by swapping the entries p and $p+1$ in the list $(\sigma(1),\sigma(2),\ldots,\sigma(j))$. Since σ is of Class LR, we know that

$$p \in L_\sigma = \{\sigma(1),\sigma(2),\ldots,\sigma(d-k)\} \qquad \text{and}$$
$$p+1 \in R_\sigma = \{\sigma(d-k+1),\sigma(d-k+2),\ldots,\sigma(j)\}.$$

In other words, p is one of the first $d-k$ entries of the list $(\sigma(1),\sigma(2),\ldots,\sigma(j))$, while $p+1$ is one of the last $j-(d-k)$ entries of this list. When we swap the entries p and $p+1$ in the list $(\sigma(1),\sigma(2),\ldots,\sigma(j))$, this positionality clearly gets reversed; thus, we have $p \in R_{s_p\sigma}$ and $p+1 \in L_{s_p\sigma}$. Hence, the permutation $s_p\sigma$ is of Class RL.

Moreover, σ is a shuffle, so we have $\sigma(1) < \sigma(2) < \cdots < \sigma(d-k)$ and $\sigma(d-k+1) < \sigma(d-k+2) < \cdots < \sigma(j)$. The number p appears in the chain of inequalities $\sigma(1) < \sigma(2) < \cdots < \sigma(d-k)$ (since $p \in L_\sigma = \{\sigma(1),\sigma(2),\ldots,\sigma(d-k)\}$), but the number $p+1$ does not (since $p+1 \in R_\sigma$ and thus $p+1 \notin L_\sigma = \{\sigma(1),\sigma(2),\ldots,\sigma(d-k)\}$). When we swap the entries p and $p+1$ in the list $(\sigma(1),\sigma(2),\ldots,\sigma(j))$, the number p is replaced by $p+1$, but the chain of inequalities $\sigma(1) < \sigma(2) < \cdots < \sigma(d-k)$ remains valid (indeed, the only inequality that would be invalidated when we swap the entries p and $p+1$ is the inequality $p < p+1$; but this inequality is not part of the chain, because $p+1$ does not appear in this chain), and so does the chain of inequalities $\sigma(d-k+1) < \sigma(d-k+2) < \cdots < \sigma(j)$ (for a similar reason: the number $p+1$ appears in this chain, but p does not). Thus, when we swap the entries p and $p+1$ in the list $(\sigma(1),\sigma(2),\ldots,\sigma(j))$, the permutation σ remains a shuffle. In other words, $s_p\sigma$ is a shuffle.

Hence, $s_p\sigma$ is a shuffle of Class RL (since we have shown that $s_p\sigma$ is of Class RL).

These two maps are mutually inverse (since $s_p\left(s_p\sigma\right) = \underbrace{s_p^2}_{=\operatorname{id}}\sigma = \sigma$ for any $\sigma \in S_j$), and thus are bijections.

However, if we replace a given shuffle σ by $s_p\sigma$, then the product

$$\left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \tag{22}$$

does not change (indeed, $s_p\sigma$ is obtained from σ by swapping the values p and $p+1$; but our assumption $a_p = a_{p+1}$ ensures that this swap does not change the vectors $a_1, a_2, \ldots, a_j$ and therefore the expression (22)). In other words, for any shuffle σ, we have

$$\begin{aligned}&\left[a_{\left(s_p\sigma\right)(1)}, \ldots, a_{\left(s_p\sigma\right)(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\left(s_p\sigma\right)(d-k+1)} \cdots a_{\left(s_p\sigma\right)(j)} \\ &= \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)}.\end{aligned} \tag{23}$$

Hence, for any shuffle σ, we have

$$\begin{aligned}&\underbrace{\operatorname{sign}\left(s_p\sigma\right)}_{\substack{=\operatorname{sign}\left(s_p\right)\cdot\operatorname{sign}(\sigma) \\ =-\operatorname{sign}(\sigma) \\ (\text{since } \operatorname{sign}\left(s_p\right)=-1)}} \underbrace{\left[a_{\left(s_p\sigma\right)(1)}, \ldots, a_{\left(s_p\sigma\right)(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\left(s_p\sigma\right)(d-k+1)} \cdots a_{\left(s_p\sigma\right)(j)}}_{\substack{=\left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\ (\text{by (23)})}} \\ &= -\operatorname{sign}(\sigma)\left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)}.\end{aligned} \tag{24}$$

Now let us summarize: The right hand side of (3.3.6) is

$$\begin{aligned}
&\sum_{\sigma \text{ is a shuffle}} \operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\
&= \sum_{\substack{\sigma \text{ is a shuffle} \\ \text{of Class LL}}} \underbrace{\operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)}}_{\substack{=0 \\ \text{(by (19))}}} \\
&\qquad + \sum_{\substack{\sigma \text{ is a shuffle} \\ \text{of Class RR}}} \underbrace{\operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)}}_{\substack{=0 \\ \text{(by (20))}}} \\
&\qquad + \sum_{\substack{\sigma \text{ is a shuffle} \\ \text{of Class RL}}} \operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\
&\qquad + \sum_{\substack{\sigma \text{ is a shuffle} \\ \text{of Class LR}}} \operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\
&\qquad \left(\begin{array}{c}\text{since each shuffle belongs to exactly one of the} \\ \text{four Classes LL, RR, RL and LR}\end{array}\right) \\
&= \sum_{\substack{\sigma \text{ is a shuffle} \\ \text{of Class RL}}} \operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\
&\qquad + \sum_{\substack{\sigma \text{ is a shuffle} \\ \text{of Class LR}}} \operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\
&= \sum_{\substack{\sigma \text{ is a shuffle} \\ \text{of Class LR}}} \underbrace{\operatorname{sign}(s_p\sigma) \left[a_{(s_p\sigma)(1)}, \ldots, a_{(s_p\sigma)(d-k)}, b_1, \ldots, b_k\right] \cdot a_{(s_p\sigma)(d-k+1)} \cdots a_{(s_p\sigma)(j)}}_{\substack{= -\operatorname{sign}(\sigma)\left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\ \text{(by (24))}}} \\
&\qquad + \sum_{\substack{\sigma \text{ is a shuffle} \\ \text{of Class LR}}} \operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } s_p\sigma \text{ for } \sigma \text{ in the first sum,} \\ \text{since the map (21) is a bijection}\end{array}\right) \\
&= -\sum_{\substack{\sigma \text{ is a shuffle} \\ \text{of Class LR}}} \operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\
&\qquad + \sum_{\substack{\sigma \text{ is a shuffle} \\ \text{of Class LR}}} \operatorname{sign}(\sigma) \left[a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k\right] \cdot a_{\sigma(d-k+1)} \cdots a_{\sigma(j)} \\
&= 0.
\end{aligned}$$

This completes the proof of the claim that the right hand side of (3.3.6) is alternating in $a_1, a_2, \ldots, a_j$, and thus (ultimately) the proof that the meet operation on $\Lambda(V)$ is well-defined. ■

70. **page 97, proof of Theorem 3.3.2:** I am not sure how part **(a)** of the theorem is being proved here (e.g., where is the associativity of $\wedge$ proved?). Let me instead prove Theorem 3.3.2 using the Hodge star operation. Here is an outline:

   - Let $[d] := \{1, 2, \ldots, d\}$. For any subset $J = \{j_1 < j_2 < \cdots < j_p\}$ of $[d]$, let e_J be the extensor $e_{j_1} \vee e_{j_2} \vee \cdots \vee e_{j_p} \in \Lambda^p(V)$. It is known that the family $(e_J)_{J \subseteq [d]}$ is a basis of the $\mathbf{C}$-vector space $\Lambda(V)$.

     In particular, $e_\varnothing = 1 \in \Lambda^0(V)$ and $e_{[d]} = e_1 \vee e_2 \vee \cdots \vee e_d \in \Lambda^d(V)$.

     It is well-known (and easy to see) that any two subsets U and V of $[d]$ satisfy
     $$e_U \vee e_V = (-1)^{|U| \cdot |V|} e_V \vee e_U. \qquad (25)$$
     More generally, any $a \in \Lambda^k(V)$ and any $b \in \Lambda^l(V)$ satisfy
     $$a \vee b = (-1)^{kl} b \vee a. \qquad (26)$$
     Moreover, if U and V are two subsets of $[d]$ that are not disjoint, then
     $$e_U \vee e_V = 0 \qquad (27)$$
     (since there is a common element $r \in U \cap V$, and the corresponding basis vector e_r appears as a factor in both e_U and e_V). On the other hand, if U and V are two disjoint subsets of $[d]$, then
     $$e_U \vee e_V = \operatorname{sign}(\eta) \cdot e_{U \cup V}, \qquad (28)$$
     where $\eta \in S_{|U|+|V|}$ is the permutation that transforms the increasing list[8] of $U \cup V$ into the concatenation of the increasing lists of U and of V.

   - For each subset J of $[d]$, we let $\sigma_J \in S_d$ be the permutation that sends the numbers $1, 2, \ldots, |J|$ to the elements of J listed in increasing order and sends the numbers $|J|+1, |J|+2, \ldots, d$ to the elements of $[d] \setminus J$ listed in increasing order. For instance:
     - If $n = 5$ and $J = \{2, 5\}$, then $\sigma_J \in S_5$ is the permutation sending $1, 2, 3, 4, 5$ to $2, 5, 1, 3, 4$.
     - If $n = 4$ and $J = \{2\}$, then $\sigma_J \in S_4$ is the permutation sending $1, 2, 3, 4$ to $2, 1, 3, 4$.

     It is easy to see that any subset J of $[d]$ satisfies
     $$e_J \vee e_{[d] \setminus J} = \operatorname{sign}(\sigma_J) \cdot e_{[d]} \qquad (29)$$

[8] The *increasing list* of a finite set K of integers means the list of all elements of K in increasing order.

(since $e_J \vee e_{[d]\setminus J}$ is the extensor obtained by multiplying the e_j for $j \in J$ in increasing order and then the e_j for $j \in [d] \setminus J$ in increasing order; but this is just the wedge product $e_1 \vee e_2 \vee \cdots \vee e_d = e_{[d]}$ with its factors permuted by σ_J).

- Define the *Hodge star* operation to be the $\mathbf{C}$-linear map $\star : \Lambda(V) \to \Lambda(V)$ that sends each element e_J (with $J \subseteq [d]$) to $\operatorname{sign}(\sigma_J) \cdot e_{[d]\setminus J}$. For instance:
  - If $n = 5$ and $J = \{2,5\}$, then $\sigma_J \in S_5$ is the permutation sending $1,2,3,4,5$ to $2,5,1,3,4$, and thus we have $\star e_{\{2,5\}} = \underbrace{\operatorname{sign}\left(\sigma_{\{2,5\}}\right)}_{=1} \cdot \underbrace{e_{[5]\setminus\{2,5\}}}_{=e_{\{1,3,4\}}} = e_{\{1,3,4\}}$.
  - If $n = 4$ and $J = \{2\}$, then $\sigma_J \in S_4$ is the permutation sending $1,2,3,4$ to $2,1,3,4$, and thus we have $\star e_{\{2\}} = \underbrace{\operatorname{sign}\left(\sigma_{\{2\}}\right)}_{=-1} \cdot \underbrace{e_{[4]\setminus\{2\}}}_{=e_{\{1,3,4\}}} = -e_{\{1,3,4\}}$.

  It is easy to see that every $k \in \mathbf{N}$ and $A \in \Lambda^k(V)$ satisfy

$$\star A \in \Lambda^{d-k}(V) \qquad (30)$$

  and

$$\star(\star A) = (-1)^{k(d-k)} A. \qquad (31)$$

  (*Proof sketch:* By linearity, it suffices to prove both (30) and (31) in the case when $A = e_J$ for some k-element subset J of $[d]$. So let J be a k-element subset of $[d]$, and let $A = e_J$. Then, $|J| = k$, so that $|[d] \setminus J| = d - k$. But $A = e_J$, and thus the definition of the Hodge star $\star$ yields $\star A = \operatorname{sign}(\sigma_J) \cdot e_{[d]\setminus J} \in \Lambda^{d-k}(V)$ (since $|[d] \setminus J| = d-k$); this immediately proves (30). It remains to prove (31).

  From $\star A = \operatorname{sign}(\sigma_J) \cdot e_{[d]\setminus J}$, we obtain

$$\begin{aligned}
\star(\star A) &= \star\left(\operatorname{sign}(\sigma_J) \cdot e_{[d]\setminus J}\right) = \operatorname{sign}(\sigma_J) \cdot \underbrace{\star e_{[d]\setminus J}}_{\substack{=\operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot e_{[d]\setminus([d]\setminus J)} \\ \text{(by the definition of } \star)}} \\
&= \underbrace{\operatorname{sign}(\sigma_J) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right)}_{=\operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}\left(\sigma_J\right)} \cdot \underbrace{e_{[d]\setminus([d]\setminus J)}}_{\substack{=e_J \\ \text{(since } [d]\setminus([d]\setminus J)=J)}} \\
&= \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}(\sigma_J) \cdot \underbrace{e_J}_{=A} \\
&= \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}(\sigma_J) \cdot A.
\end{aligned}$$

Hence, in order to complete the proof of (31), we only need to show that

$$\operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}\left(\sigma_J\right) = (-1)^{k(d-k)} . \tag{32}$$

For this purpose, we let $\eta_k \in S_d$ be the permutation that sends the numbers $1, 2, \ldots, k$ to $d-k+1, d-k+2, \ldots, d$ and sends the numbers $k+1, k+2, \ldots, d$ to the numbers $1, 2, \ldots, d-k$. Then, it is easy to see that $\operatorname{sign}(\eta_k) = (-1)^{k(d-k)}$ and $\sigma_J = \sigma_{[d]\setminus J}\eta_k$ [9]. The latter equality yields

$$\begin{aligned} \operatorname{sign}(\sigma_J) &= \operatorname{sign}\left(\sigma_{[d]\setminus J}\eta_k\right) = \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \underbrace{\operatorname{sign}(\eta_k)}_{=(-1)^{k(d-k)}} \\ &= \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot (-1)^{k(d-k)} , \end{aligned}$$

and therefore we have

$$\begin{aligned} \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}(\sigma_J) &= \underbrace{\operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right)}_{\substack{=\left(\operatorname{sign}\left(\sigma_{[d]\setminus J}\right)\right)^2=1 \\ \text{(since } \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \text{ is 1 or } -1)}} \cdot (-1)^{k(d-k)} \\ &= (-1)^{k(d-k)} . \end{aligned} \tag{33}$$

This proves (32), and so the proof of (31) is complete.)

- The Hodge star $\star : \Lambda(V) \to \Lambda(V)$ is a vector space isomorphism. (Indeed, (31) shows that the composition $\star \circ \star$ is the linear map that sends each $A \in \Lambda^k(V)$ to $(-1)^{k(d-k)} A$. But the latter map is clearly a vector space isomorphism. Thus, $\star$ is an isomorphism as well.)
- Now we claim that any $a, b \in \Lambda(V)$ satisfy

$$\star(a \wedge b) = (\star a) \vee (\star b) . \tag{34}$$

(*Proof sketch:* By linearity, we WLOG assume that $a = e_I$ and $b = e_J$, where I and J are two subsets of $[d]$. Then, the definition of the Hodge

[9] *Proof.* The permutation $\sigma_{[d]\setminus J}\eta_k$ sends the numbers $1, 2, \ldots, k$ to the elements of J in increasing order (because the permutation η_k sends the numbers $1, 2, \ldots, k$ to the numbers $d-k+1, d-k+2, \ldots, d$, and then the permutation $\sigma_{[d]\setminus J}$ sends the latter numbers to the elements of $[d] \setminus ([d] \setminus J) = J$ in increasing order), and furthermore sends the numbers $k+1, k+2, \ldots, d$ to the elements of $[d] \setminus J$ in increasing order (because the permutation η_k sends the numbers $k+1, k+2, \ldots, d$ to the numbers $1, 2, \ldots, d-k$, and then the permutation $\sigma_{[d]\setminus J}$ sends the latter numbers to the elements of $[d] \setminus J$ in increasing order). But the permutation σ_J does the exact same things (by its definition). Hence, the two permutations $\sigma_{[d]\setminus J}\eta_k$ and σ_J agree on all the numbers $1, 2, \ldots, k$ and on all the numbers $k+1, k+2, \ldots, d$. In other words, these two permutations are equal. That is, $\sigma_J = \sigma_{[d]\setminus J}\eta_k$.

star entails that $\star a = \operatorname{sign}(\sigma_I) \cdot e_{[d]\setminus I}$ and $\star b = \operatorname{sign}(\sigma_J) \cdot e_{[d]\setminus J}$. Thus,

$$(\star a) \vee (\star b) = \left(\operatorname{sign}(\sigma_I) \cdot e_{[d]\setminus I}\right) \vee \left(\operatorname{sign}(\sigma_J) \cdot e_{[d]\setminus J}\right)$$
$$= \operatorname{sign}(\sigma_I) \cdot \operatorname{sign}(\sigma_J) \cdot e_{[d]\setminus I} \vee e_{[d]\setminus J}. \tag{35}$$

Meanwhile, let us write the subsets I and J as $I = \{i_1 < i_2 < \cdots < i_p\}$ and $J = \{j_1 < j_2 < \cdots < j_q\}$. Thus, $a = e_I = e_{i_1} e_{i_2} \cdots e_{i_p}$ and $b = e_J = e_{j_1} e_{j_2} \cdots e_{j_q}$, so that

$$a \wedge b = \left(e_{i_1} e_{i_2} \cdots e_{i_p}\right) \wedge \left(e_{j_1} e_{j_2} \cdots e_{j_q}\right)$$
$$= \sum_{\sigma} \operatorname{sign}(\sigma) \left[e_{i_{\sigma(1)}}, \ldots, e_{i_{\sigma(d-q)}}, e_{j_1}, \ldots, e_{j_q}\right] \cdot e_{i_{\sigma(d-q+1)}} \cdots e_{i_{\sigma(p)}} \tag{36}$$

(by the definition of the meet operation), where the sum ranges over certain permutations $\sigma \in S_p$ (namely, those that satisfy $\sigma(1) < \sigma(2) < \cdots < \sigma(d-q)$ and $\sigma(d-q+1) < \sigma(d-q+2) < \cdots < \sigma(p)$). We shall denote the latter permutations σ as the *shuffles*. Thus, the sum in (36) ranges over the shuffles σ.

Now we are in one of the following two cases:

Case 1: We have $I \cup J \neq [d]$.

Case 2: We have $I \cup J = [d]$.

Let us first consider Case 1. In this case, we have $I \cup J \neq [d]$. Hence, there exists some $r \in [d]$ such that $r \notin I$ and $r \notin J$. Consider this r. Now, **each** permutation $\sigma \in S_p$ satisfies $\left[e_{i_{\sigma(1)}}, \ldots, e_{i_{\sigma(d-q)}}, e_{j_1}, \ldots, e_{j_q}\right] = 0$ (since $r \notin I$ and $r \notin J$ shows that the basis vector e_r does not appear among $e_{i_{\sigma(1)}}, \ldots, e_{i_{\sigma(d-q)}}, e_{j_1}, \ldots, e_{j_q}$, which entails by the pigeonhole principle that two entries of the list $e_{i_{\sigma(1)}}, \ldots, e_{i_{\sigma(d-q)}}, e_{j_1}, \ldots, e_{j_q}$ are equal[10]). Hence, all the addends in the sum on the right hand side of (36) equal 0. Thus, (36) simplifies to $a \wedge b = 0$. Hence $\star(a \wedge b) = \star 0 = 0$. On the other hand, we have $r \in [d] \setminus I$ (since $r \notin I$) and $r \in [d] \setminus J$ (since $r \notin J$); thus, the extensor $e_{[d]\setminus I} \vee e_{[d]\setminus J}$ has two equal factors (namely, e_r, which appears as a factor in both $e_{[d]\setminus I}$ and $e_{[d]\setminus J}$). Hence, this extensor is 0. Thus, $e_{[d]\setminus I} \vee e_{[d]\setminus J} = 0$. Therefore, (35) simplifies to $(\star a) \vee (\star b) = 0$. Comparing this with $\star(a \wedge b) = 0$, we obtain $\star(a \wedge b) = (\star a) \vee (\star b)$. Thus, (34) is proved in Case 1.

Now, let us consider Case 2. In this case, we have $I \cup J = [d]$. Hence, $[d] \setminus J \subseteq I$. Thus, $[d] \setminus J$ is a $(d-q)$-element subset of I (since $|J| = q$

[10] Indeed, there are only d basis vectors $e_1, e_2, \ldots, e_d$ available. Thus, any list that contains d of these vectors but does not contain e_r must contain two equal entries.

and thus $|[d] \setminus J| = d - q$). In other words,

$$[d] \setminus J = \left\{ i_{r_1} < i_{r_2} < \cdots < i_{r_{d-q}} \right\}$$

for some $d - q$ elements $r_1 < r_2 < \cdots < r_{d-q}$ of $\{1, 2, \ldots, p\}$ (since $I = \{i_1 < i_2 < \cdots < i_p\}$). Consider these elements.

It is easy to see that for any $d - q$ elements $h_1 < h_2 < \cdots < h_{d-q}$ of $\{1, 2, \ldots, p\}$, there is a unique shuffle σ_h whose first $d - q$ values $\sigma_h(1), \sigma_h(2), \ldots, \sigma_h(d-q)$ are the elements $h_1, h_2, \ldots, h_{d-q}$ (indeed, its remaining $p - (d-q)$ values $\sigma_h(d-q+1), \sigma_h(d-q+2), \ldots, \sigma_h(p)$ must then be the remaining $p - (d-q)$ elements of $\{1, 2, \ldots, p\}$ listed in increasing order). Applying this to $h_k = r_k$, we see that there is a unique shuffle σ_0 whose first $d - q$ values $\sigma_0(1), \sigma_0(2), \ldots, \sigma_0(d-q)$ are the elements $r_1, r_2, \ldots, r_{d-q}$ (because $r_1 < r_2 < \cdots < r_{d-q}$). Consider this shuffle σ_0. It satisfies

$$\begin{aligned}\left[e_{i_{\sigma_0(1)}}, \ldots, e_{i_{\sigma_0(d-q)}}, e_{j_1}, \ldots, e_{j_q}\right] &= \left[e_{i_{r_1}}, \ldots, e_{i_{r_{d-q}}}, e_{j_1}, \ldots, e_{j_q}\right] \\ &= \operatorname{sign}\left(\sigma_{[d]\setminus J}\right), \end{aligned} \tag{37}$$

since the vectors $e_{i_{r_1}}, \ldots, e_{i_{r_{d-q}}}, e_{j_1}, \ldots, e_{j_q}$ inside the bracket are precisely the d basis vectors $e_1, e_2, \ldots, e_d$ of V permuted using the permutation $\sigma_{[d]\setminus J}$ (because we have $\left\{ i_{r_1} < i_{r_2} < \cdots < i_{r_{d-q}} \right\} = [d] \setminus J$ and $\{j_1 < j_2 < \cdots < j_q\} = J$).

Furthermore, since σ_0 is a shuffle, we have $\sigma_0(d-q+1) < \sigma_0(d-q+2) < \cdots < \sigma_0(p)$, so that $i_{\sigma_0(d-q+1)} < i_{\sigma_0(d-q+2)} < \cdots < i_{\sigma_0(p)}$ (since $i_1 < i_2 < \cdots < i_p$). Since we also have

$$\begin{aligned}&\left\{ i_{\sigma_0(d-q+1)}, i_{\sigma_0(d-q+2)}, \ldots, i_{\sigma_0(p)} \right\} \\ &= \underbrace{\{i_1, i_2, \ldots, i_p\}}_{\substack{=\{i_1 < i_2 < \cdots < i_p\} \\ = I}} \setminus \underbrace{\left\{ i_{\sigma_0(1)}, i_{\sigma_0(2)}, \ldots, i_{\sigma_0(d-q)} \right\}}_{\substack{= \left\{ i_{r_1}, i_{r_2}, \ldots, i_{r_{d-q}} \right\} \\ \text{(since the numbers } \sigma_0(1), \sigma_0(2), \ldots, \sigma_0(d-q) \\ \text{are the elements } r_1, r_2, \ldots, r_{d-q})}} \\ &\qquad \text{(because } \sigma_0 \text{ is a permutation of } \{1, 2, \ldots, p\}) \\ &= I \setminus \underbrace{\left\{ i_{r_1}, i_{r_2}, \ldots, i_{r_{d-q}} \right\}}_{\substack{= \left\{ i_{r_1} < i_{r_2} < \cdots < i_{r_{d-q}} \right\} \\ = [d] \setminus J}} = I \setminus ([d] \setminus J) = I \cap J,\end{aligned}$$

we thus conclude that the numbers $i_{\sigma_0(d-q+1)}, i_{\sigma_0(d-q+2)}, \ldots, i_{\sigma_0(p)}$ are the elements of $I \cap J$ listed in increasing order. Hence,

$$e_{i_{\sigma_0(d-q+1)}} e_{i_{\sigma_0(d-q+2)}} \cdots e_{i_{\sigma_0(p)}} = e_{I \cap J}. \tag{38}$$

Recall that the sum on the right hand side of (36) ranges over all shuffles σ; one of these shuffles σ is σ_0. All other shuffles σ are distinct from σ_0, and thus satisfy

$$\left[e_{i_{\sigma(1)}}, \ldots, e_{i_{\sigma(d-q)}}, e_{j_1}, \ldots, e_{j_q}\right] = 0,$$

since the bracket $\left[e_{i_{\sigma(1)}}, \ldots, e_{i_{\sigma(d-q)}}, e_{j_1}, \ldots, e_{j_q}\right]$ contains two equal vectors[11]. Thus, all the addends in the sum on the right hand side of (36) equal 0, except for the addend for $\sigma = \sigma_0$, which is

$$\begin{aligned} & \operatorname{sign}(\sigma_0) \underbrace{\left[e_{i_{\sigma_0(1)}}, \ldots, e_{i_{\sigma_0(d-q)}}, e_{j_1}, \ldots, e_{j_q}\right]}_{\substack{=\operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \\ \text{(by (37))}}} \cdot \underbrace{e_{i_{\sigma_0(d-q+1)}} \cdots e_{i_{\sigma_0(p)}}}_{\substack{=e_{I\cap J} \\ \text{(by (38))}}} \\ & = \operatorname{sign}(\sigma_0) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot e_{I\cap J}. \end{aligned}$$

Thus, (36) rewrites as

$$a \wedge b = \operatorname{sign}(\sigma_0) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot e_{I\cap J}.$$

Thus,

$$\begin{aligned} \star(a \wedge b) &= \star\left(\operatorname{sign}(\sigma_0) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot e_{I\cap J}\right) \\ &= \operatorname{sign}(\sigma_0) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \underbrace{\star e_{I\cap J}}_{\substack{=\operatorname{sign}\left(\sigma_{I\cap J}\right)\cdot e_{[d]\setminus(I\cap J)} \\ \text{(by the definition of } \star)}} \\ &= \operatorname{sign}(\sigma_0) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}(\sigma_{I\cap J}) \cdot e_{[d]\setminus(I\cap J)}. \end{aligned} \tag{39}$$

On the other hand, $I \cup J = [d]$ entails that the sets $[d] \setminus I$ and $[d] \setminus J$ are disjoint. Hence, (28) shows that

$$e_{[d]\setminus I} \vee e_{[d]\setminus J} = \operatorname{sign}(\eta) \cdot e_{([d]\setminus I)\cup([d]\setminus J)},$$

[11] Why? Because if it didn't, then the indices $i_{\sigma(1)}, \ldots, i_{\sigma(d-q)}, j_1, \ldots, j_q$ would be distinct, so that we would have

$$\left\{i_{\sigma(1)}, \ldots, i_{\sigma(d-q)}\right\} = [d] \setminus \underbrace{\{j_1, \ldots, j_q\}}_{=J} = [d] \setminus J = \left\{i_{r_1} < i_{r_2} < \cdots < i_{r_{d-q}}\right\};$$

but this would entail $\{\sigma(1), \ldots, \sigma(d-q)\} = \left\{r_1, r_2, \ldots, r_{d-q}\right\}$ (since the $i_1, i_2, \ldots, i_p$ are distinct), and therefore $(\sigma(1), \ldots, \sigma(d-q)) = \left(r_1, r_2, \ldots, r_{d-q}\right)$ (since $\sigma(1) < \cdots < \sigma(d-q)$ (because σ is a shuffle) and $r_1 < r_2 < \cdots < r_{d-q}$), and this would entail $\sigma = \sigma_0$ (since σ is a shuffle, but σ_0 is the only shuffle such that the numbers $\sigma_0(1), \sigma_0(2), \ldots, \sigma_0(d-q)$ are the elements $r_1, r_2, \ldots, r_{d-q}$).

where $\eta \in S_{|[d]\setminus I|+|[d]\setminus J|}$ is the permutation that transforms the increasing list of $([d] \setminus I) \cup ([d] \setminus J)$ into the concatenation of the increasing lists of $[d] \setminus I$ and of $[d] \setminus J$. In view of $([d] \setminus I) \cup ([d] \setminus J) = [d] \setminus (I \cap J)$, we can rewrite this as

$$e_{[d]\setminus I} \vee e_{[d]\setminus J} = \operatorname{sign}(\eta) \cdot e_{[d]\setminus(I\cap J)}. \tag{40}$$

Hence, (35) rewrites as

$$(\star a) \vee (\star b) = \operatorname{sign}(\sigma_I) \cdot \operatorname{sign}(\sigma_J) \cdot \operatorname{sign}(\eta) \cdot e_{[d]\setminus(I\cap J)}.$$

Comparing this with (39), we see that the vectors $\star(a \wedge b)$ and $(\star a) \vee (\star b)$ are equal up to sign. All that remains to be proved now is that their signs agree as well. In other words, we must prove that

$$\begin{aligned} &\operatorname{sign}(\sigma_0) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}(\sigma_{I\cap J}) \\ &= \operatorname{sign}(\sigma_I) \cdot \operatorname{sign}(\sigma_J) \cdot \operatorname{sign}(\eta). \end{aligned} \tag{41}$$

The easiest way to prove this equality is as follows: First, we observe that $[d] \setminus I = J \setminus I$ (since $I \cup J = [d]$) and thus $|I \cap J| + |[d] \setminus I| = |I \cap J| + |J \setminus I| = |J| = q$. Thus, the extensor $e_{I\cap J} \vee e_{[d]\setminus I}$ has step $|I \cap J| + |[d] \setminus I| = q$. Hence, $e_{I\cap J} \vee e_{[d]\setminus I} \in \Lambda^q(V)$. Moreover, $|[d] \setminus J| = d - q$ (since $|J| = q$) and thus $e_{[d]\setminus J} \in \Lambda^{d-q}(V)$. Hence, (26) (applied to $k = q$ and $l = d - q$ and $a = e_{I\cap J} \vee e_{[d]\setminus I}$ and $b = e_{[d]\setminus J}$) yields

$$e_{I\cap J} \vee e_{[d]\setminus I} \vee e_{[d]\setminus J} = (-1)^{q(d-q)} e_{[d]\setminus J} \vee e_{I\cap J} \vee e_{[d]\setminus I}. \tag{42}$$

But (29) (applied to $I \cap J$ instead of J) yields

$$e_{I\cap J} \vee e_{[d]\setminus(I\cap J)} = \operatorname{sign}(\sigma_{I\cap J}) \cdot e_{[d]}.$$

Hence,

$$\begin{aligned} &\operatorname{sign}(\sigma_{I\cap J}) \cdot e_{[d]} \\ &= e_{I\cap J} \vee \underbrace{e_{[d]\setminus(I\cap J)}}_{\substack{=\frac{1}{\operatorname{sign}(\eta)} e_{[d]\setminus I} \vee e_{[d]\setminus J} \\ \text{(by (40))}}} \\ &= \underbrace{\frac{1}{\operatorname{sign}(\eta)}}_{\substack{=\operatorname{sign}(\eta) \\ \text{(since } \operatorname{sign}(\eta) \\ \text{is 1 or } -1)}} \underbrace{e_{I\cap J} \vee e_{[d]\setminus I} \vee e_{[d]\setminus J}}_{\substack{=(-1)^{q(d-q)} e_{[d]\setminus J} \vee e_{I\cap J} \vee e_{[d]\setminus I} \\ \text{(by (42))}}} \\ &= \operatorname{sign}(\eta) \cdot (-1)^{q(d-q)} e_{[d]\setminus J} \vee e_{I\cap J} \vee e_{[d]\setminus I}. \end{aligned} \tag{43}$$

Now, recall that the values $\sigma_0(1), \sigma_0(2), \ldots, \sigma_0(d-q)$ are the elements $r_1, r_2, \ldots, r_{d-q}$ (by the definition of σ_0). Hence,

$$e_{i_{\sigma_0(1)}} e_{i_{\sigma_0(2)}} \cdots e_{i_{\sigma_0(d-q)}} = e_{i_{r_1}} e_{i_{r_2}} \cdots e_{i_{r_{d-q}}} = e_{[d]\setminus J}$$

(since $[d] \setminus J = \left\{ i_{r_1} < i_{r_2} < \cdots < i_{r_{d-q}} \right\}$). Multiplying this equality with (38) (using the join operation $\vee$), we obtain

$$\begin{aligned} &\left(e_{i_{\sigma_0(1)}} e_{i_{\sigma_0(2)}} \cdots e_{i_{\sigma_0(d-q)}} \right) \vee \left(e_{i_{\sigma_0(d-q+1)}} e_{i_{\sigma_0(d-q+2)}} \cdots e_{i_{\sigma_0(p)}} \right) \\ &= e_{[d]\setminus J} \vee e_{I \cap J}. \end{aligned}$$

Hence,

$$\begin{aligned} &e_{[d]\setminus J} \vee e_{I\cap J} \\ &= \left(e_{i_{\sigma_0(1)}} e_{i_{\sigma_0(2)}} \cdots e_{i_{\sigma_0(d-q)}} \right) \vee \left(e_{i_{\sigma_0(d-q+1)}} e_{i_{\sigma_0(d-q+2)}} \cdots e_{i_{\sigma_0(p)}} \right) \\ &= e_{i_{\sigma_0(1)}} e_{i_{\sigma_0(2)}} \cdots e_{i_{\sigma_0(p)}} = \operatorname{sign}(\sigma_0) \cdot \underbrace{e_{i_1} e_{i_2} \cdots e_{i_p}}_{= e_I} = \operatorname{sign}(\sigma_0) \cdot e_I. \end{aligned}$$

Hence, (43) becomes

$$\begin{aligned} &\operatorname{sign}(\sigma_{I\cap J}) \cdot e_{[d]} \\ &= \operatorname{sign}(\eta) \cdot (-1)^{q(d-q)} \underbrace{e_{[d]\setminus J} \vee e_{I\cap J}}_{= \operatorname{sign}(\sigma_0) \cdot e_I} \vee e_{[d]\setminus I} \\ &= \operatorname{sign}(\eta) \cdot (-1)^{q(d-q)} \cdot \operatorname{sign}(\sigma_0) \cdot \underbrace{e_I \vee e_{[d]\setminus I}}_{\substack{= \operatorname{sign}(\sigma_I) \cdot e_{[d]} \\ \text{(by (29))}}} \\ &= \operatorname{sign}(\eta) \cdot (-1)^{q(d-q)} \cdot \operatorname{sign}(\sigma_0) \cdot \operatorname{sign}(\sigma_I) \cdot e_{[d]}. \end{aligned}$$

Since $e_{[d]}$ is nonzero (and in fact an element of a basis of $\Lambda(V)$), we can cancel $e_{[d]}$ from this equality, and obtain

$$\begin{aligned} &\operatorname{sign}(\sigma_{I\cap J}) \\ &= \operatorname{sign}(\eta) \cdot \underbrace{(-1)^{q(d-q)}}_{\substack{= \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}\left(\sigma_J\right) \\ \text{(since } |J| = q \text{, and thus (33) (applied to } k = q \text{)} \\ \text{yields } \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}\left(\sigma_J\right) = (-1)^{q(d-q)} \text{)}}} \cdot \operatorname{sign}(\sigma_0) \cdot \operatorname{sign}(\sigma_I) \\ &= \operatorname{sign}(\eta) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}(\sigma_J) \cdot \operatorname{sign}(\sigma_0) \cdot \operatorname{sign}(\sigma_I). \end{aligned}$$

Multiplying this equality by $\operatorname{sign}(\sigma_0) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right)$, we obtain

$$\begin{aligned}
&\operatorname{sign}(\sigma_0) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}(\sigma_{I\cap J}) \\
&= \operatorname{sign}(\sigma_0) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}(\eta) \cdot \operatorname{sign}\left(\sigma_{[d]\setminus J}\right) \cdot \operatorname{sign}(\sigma_J) \cdot \operatorname{sign}(\sigma_0) \cdot \operatorname{sign}(\sigma_I) \\
&= \operatorname{sign}(\sigma_I) \cdot \operatorname{sign}(\sigma_J) \cdot \operatorname{sign}(\eta) \cdot \underbrace{\left(\operatorname{sign}\left(\sigma_{[d]\setminus J}\right)\right)^2}_{=1} \cdot \underbrace{(\operatorname{sign}(\sigma_0))^2}_{=1} \\
&= \operatorname{sign}(\sigma_I) \cdot \operatorname{sign}(\sigma_J) \cdot \operatorname{sign}(\eta).
\end{aligned}$$

This proves (41), and thus completes the proof of (34) in Case 2.

Thus, (34) is proved in both Cases 1 and 2; this completes the proof of (34).)

- Now, it is easy to see that the meet is anticommutative: i.e., any $a \in \Lambda^j(V)$ and $b \in \Lambda^k(V)$ satisfy

$$a \wedge b = (-1)^{(d-k)(d-j)} b \wedge a. \tag{44}$$

(*Proof sketch:* Let $a \in \Lambda^j(V)$ and $b \in \Lambda^k(V)$. Then, (30) yields $\star a \in \Lambda^{d-j}(V)$ and $\star b \in \Lambda^{d-k}(V)$. Hence, (26) (applied to $d-j$, $d-k$, $\star a$ and $\star b$ instead of k, j, a and b) yields

$$(\star a) \vee (\star b) = (-1)^{(d-j)(d-k)} (\star b) \vee (\star a). \tag{45}$$

However, (34) yields $\star(a \wedge b) = (\star a) \vee (\star b)$ and $\star(b \wedge a) = (\star b) \vee (\star a)$. In light of these two equalities, we can rewrite (45) as

$$\star(a \wedge b) = (-1)^{(d-j)(d-k)} \star(b \wedge a) = \star\left((-1)^{(d-j)(d-k)} (b \wedge a)\right).$$

Since the map $\star$ is injective (because $\star$ is a vector space isomorphism), this entails $a \wedge b = (-1)^{(d-j)(d-k)} (b \wedge a) = (-1)^{(d-k)(d-j)} b \wedge a$. This proves (44).)

- Now, it is easy to see that the meet is associative: i.e., any $a, b, c \in \Lambda(V)$ satisfy

$$a \wedge (b \wedge c) = (a \wedge b) \wedge c. \tag{46}$$

(*Proof sketch:* This is similar to the proof of (44): Again, we use (34) to reduce the claim $a \wedge (b \wedge c) = (a \wedge b) \wedge c$ to the equality $(\star a) \vee ((\star b) \vee (\star c)) = ((\star a) \vee (\star b)) \vee (\star c)$, which is true because the join is associative.)

Thus, Theorem 3.3.2 **(a)** is proved. ■

71. **page 97, proof of Theorem 3.3.2:** "In view of the assumption $\overline{A}+\overline{B}=V$": This is somewhat inappropriate, since no such assumption has been made. Instead, the case $\overline{A}+\overline{B}\neq V$ must be considered separately. Fortunately, this case is very easy: In this case, we can easily see that $A\wedge B=0$, because all the brackets $\left[a_{\sigma(1)},\ldots,a_{\sigma(d-k)},b_1,\ldots,b_k\right]$ on the right hand side of (3.3.6) are 0 (indeed, $\overline{A}+\overline{B}\neq V$ shows that $\overline{A}+\overline{B}$ is a proper subspace of V, and obviously all the d vectors $a_{\sigma(1)},\ldots,a_{\sigma(d-k)},b_1,\ldots,b_k$ belong to this proper subspace; but this entails that these d vectors cannot be linearly independent, and thus must satisfy $\left[a_{\sigma(1)},\ldots,a_{\sigma(d-k)},b_1,\ldots,b_k\right]=0$).

72. **page 97, proof of Theorem 3.3.2:** Let me add a few remarks about Theorem 3.3.2 **(b)**.

    a) Somewhat surprisingly, Theorem 3.3.2 **(b)** is not true if we replace the base field $\mathbf{C}$ by a commutative ring. For a specific counterexample, let $\mathbf{k}$ be the polynomial ring $\mathbf{R}[x,y,z]$, and let $d=4$ and $k=3$ and $j=3$. In the exterior algebra $\Lambda_{\mathbf{k}}\left(\mathbf{k}^4\right)$, let $A=e_1e_2e_3$ and $B=b_1b_2b_3$, where (e_1,e_2,e_3,e_4) is the standard basis of $\mathbf{k}^4$, and where b_1,b_2,b_3 are the vectors $(1,0,0,x)^T,(0,1,0,y)^T,(0,0,1,z)^T$. Then, A and B are extensors, but

    $$\begin{aligned} A\wedge B &= [e_1,b_1,b_2,b_3]\,e_2e_3-[e_2,b_1,b_2,b_3]\,e_1e_3+[e_3,b_1,b_2,b_3]\,e_1e_2\\ &= xe_2e_3-ye_1e_3+ze_1e_2\end{aligned}$$

    is not. (Indeed, if $xe_2e_3-ye_1e_3+ze_1e_2$ was an extensor $v\vee w$, then we could write v as $v=(\alpha,\beta,\gamma,\delta)^T$, and obtain

    $$\begin{aligned} 0 &= \underbrace{v}_{\substack{=(\alpha,\beta,\gamma,\delta)^T\\ =\alpha e_1+\beta e_2+\gamma e_3+\delta e_4}} \vee \underbrace{v\vee w}_{=xe_2e_3-ye_1e_3+ze_1e_2}\\ &= (\alpha e_1+\beta e_2+\gamma e_3+\delta e_4)\vee(xe_2e_3-ye_1e_3+ze_1e_2)\\ &= (\alpha x+\beta y+\gamma z)\,e_1e_2e_3+\delta e_4\,(xe_2e_3-ye_1e_3+ze_1e_2)\,,\end{aligned}$$

    which would entail $\alpha x+\beta y+\gamma z=0$ and $\delta=0$, and thus the vector $(\alpha,\beta,\gamma)^T$ would be a nonvanishing tangent vector field on the 2-sphere $\mathbf{R}^3$ (nonvanishing because $v\vee w=xe_2e_3-ye_1e_3+ze_1e_2$ is nonvanishing and thus v is nonvanishing); but this would contradict the hairy ball theorem (which says that no such tangent vector fields exist, even if we replace polynomials by continuous functions). Thus, $xe_2e_3-ye_1e_3+ze_1e_2$ is not an extensor.)

    This explains why the proof of Theorem 3.3.2 **(b)** must use linear algebra that is specific to vector spaces (over fields).

b) The Hodge star $\star$ has a similar property to Theorem 3.3.2 **(b)**: If $A \in \Lambda(V)$ is any extensor, then $\star A$ is an extensor again. This, too, would fail if we replaced $\mathbf{C}$ by a commutative ring such as $\mathbf{R}[x, y, z]$ (and again, a counterexample can be built based on the hairy ball theorem).

73. **page 103, Theorem:** This theorem is very easy to prove without any algebra.

    Proof sketch: WLOG assume that (ℓ_3, ℓ_4) is the pair of lines whose intersection we must prove (while all the other five pairs are known to intersect). The two lines ℓ_1 and ℓ_2 are distinct (otherwise, ℓ_1, ℓ_2, ℓ_3 would lie in a plane), and thus span a plane H and intersect at a point P. The line ℓ_3 intersects both ℓ_1 and ℓ_2 but cannot lie on the plane H (otherwise, ℓ_1, ℓ_2, ℓ_3 would lie in a plane); thus, ℓ_3 must pass through P. Similarly, ℓ_4 must pass through P. Hence, all four lines $\ell_1, \ell_2, \ell_3, \ell_4$ pass through P; in particular, the pair (ℓ_3, ℓ_4) intersects. ■

74. **page 111:** "the *weight* of a tableaux" should be "the *weight* of a tableau".

75. **page 111:** It is not true that "the property of being homogeneous depends only on the image in $\mathcal{B}_{n,d}$". Indeed, an inhomogeneous bracket polyomial can lie in $I_{n,d}$ (an example is a linear combination of syzygies of different weights) and thus turn into the homogeneous polynomial 0 in $\mathcal{B}_{n,d}$.

76. **page 118, Proposition 3.6.1:** It is worth saying that the proof of Proposition 3.6.1 is entirely straightforward:

Proof of Proposition 3.6.1. The equality (3.6.3) becomes[12]

$$\begin{aligned}
&\overline{f}\left(\overline{x},\overline{y}\right) \\
&= \sum_{k=0}^{n} \binom{n}{k} a_k \underbrace{\left(c_{11}\overline{x}+c_{12}\overline{y}\right)^k}_{\substack{=\sum_j \binom{k}{j} (c_{11}\overline{x})^j (c_{12}\overline{y})^{k-j} \\ \text{(by the binomial theorem)}}} \underbrace{\left(c_{21}\overline{x}+c_{22}\overline{y}\right)^{n-k}}_{\substack{=\sum_l \binom{n-k}{l} (c_{21}\overline{x})^l (c_{22}\overline{y})^{n-k-l} \\ \text{(by the binomial theorem)}}} \\
&= \sum_{k=0}^{n} \binom{n}{k} a_k \sum_j \binom{k}{j} \left(c_{11}\overline{x}\right)^j \left(c_{12}\overline{y}\right)^{k-j} \sum_l \binom{n-k}{l} \left(c_{21}\overline{x}\right)^l \left(c_{22}\overline{y}\right)^{n-k-l} \\
&= \sum_{k=0}^{n} \sum_j \sum_l \underbrace{\binom{n}{k} a_k \binom{k}{j} \left(c_{11}\overline{x}\right)^j \left(c_{12}\overline{y}\right)^{k-j} \binom{n-k}{l} \left(c_{21}\overline{x}\right)^l \left(c_{22}\overline{y}\right)^{n-k-l}}_{=\binom{n}{k}\binom{k}{j}\binom{n-k}{l} a_k c_{11}^j c_{12}^{k-j} c_{21}^l c_{22}^{n-k-l} \overline{x}^{j+l} \overline{y}^{n-j-l}} \\
&= \sum_{k=0}^{n} \sum_j \sum_l \binom{n}{k}\binom{k}{j}\binom{n-k}{l} a_k c_{11}^j c_{12}^{k-j} c_{21}^l c_{22}^{n-k-l} \overline{x}^{j+l} \overline{y}^{n-j-l} \\
&= \underbrace{\sum_{k=0}^{n} \sum_j \sum_i}_{=\sum_i \sum_{k=0}^{n} \sum_j} \underbrace{\binom{n}{k}\binom{k}{j}\binom{n-k}{i-j}}_{\substack{=\binom{n}{i}\binom{i}{j}\binom{n-i}{k-j} \\ \text{(this is easy to check using the} \\ \text{definition of binomial coefficients)}}} a_k c_{11}^j c_{12}^{k-j} c_{21}^{i-j} \underbrace{c_{22}^{n-k-(i-j)}}_{=c_{22}^{n-i-k+j}} \underbrace{\overline{x}^{j+(i-j)}}_{=\overline{x}^i} \underbrace{\overline{y}^{n-j-(i-j)}}_{=\overline{y}^{n-i}} \\
&\qquad \text{(here, we substituted } i-j \text{ for } l \text{ in the third sum)} \\
&= \sum_i \sum_{k=0}^{n} \sum_j \binom{n}{i}\binom{i}{j}\binom{n-i}{k-j} a_k c_{11}^j c_{12}^{k-j} c_{21}^{i-j} c_{22}^{n-i-k+j} \overline{x}^i \overline{y}^{n-i} \\
&= \sum_k \sum_{i=0}^{n} \sum_j \binom{n}{k}\binom{k}{j}\binom{n-k}{i-j} a_i c_{11}^j c_{12}^{i-j} c_{21}^{k-j} c_{22}^{n-k-i+j} \overline{x}^k \overline{y}^{n-k} \\
&\qquad \text{(here, we renamed the summation indices } i \text{ and } k \text{ as } k \text{ and } i\text{)} \\
&= \sum_k \binom{n}{k} \left(\sum_{i=0}^{n} \left(\sum_j \binom{k}{j}\binom{n-k}{i-j} c_{11}^j c_{12}^{i-j} c_{21}^{k-j} c_{22}^{n-k-i+j} \right) a_i \right) \overline{x}^k \overline{y}^{n-k}.
\end{aligned}$$

In other words,

$$\overline{f}\left(\overline{x},\overline{y}\right) = \sum_k \binom{n}{k} \overline{a}_k \overline{x}^k \overline{y}^{n-k},$$

where

$$\overline{a}_k = \sum_{i=0}^{n} \left(\sum_j \binom{k}{j}\binom{n-k}{i-j} c_{11}^j c_{12}^{i-j} c_{21}^{k-j} c_{22}^{n-k-i+j} \right) a_i. \tag{47}$$

[12]Any sum with no upper or lower limits is understood to range over all integers.

The latter equality (47) is precisely the equality (3.6.5), except that the second sum ranges over all $j \in \mathbf{Z}$ instead of only ranging from $\max(0, i-n+k)$ to $\min(i,k)$. But this little difference in summation ranges does not affect the value of the sum, since all addends with $j < \max(0, i-n+k)$ or $j > \min(i,k)$ are 0 anyway (indeed, $\binom{k}{j} = 0$ if $j < 0$ or $j > k$, whereas $\binom{n-k}{i-j} = 0$ if $j < i-n+k$ or $j > i$). Thus, by proving (47), we have proved (3.6.5) and therefore Proposition 3.6.1. ■

77. **page 122:** "defines a $\mathbf{C}$-algebra homomorphism" should be "defines a $\mathbf{C}[x,y]$-algebra homomorphism" (or you should say separately that Ψ is supposed to send $x \mapsto x$ and $y \mapsto y$).

78. **page 122, (3.6.7):** In "$\frac{(-1)^k}{n!}\mu_1 \cdots \mu_n \cdot \sigma_k\left(\frac{\nu_1}{\mu_1}, \ldots, \frac{\nu_n}{\mu_n}\right)$", the "$n!$" should be "$\binom{n}{k}$". (But the "$n!$" one line above is correct.)

79. **page 123, proof of Lemma 3.6.3:** Here is an alternative proof:

    Second proof of Lemma 3.6.3. Equip the polynomial ring $\mathbf{C}[\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y]$ with a monomial order that is lexicographic with $\mu_1 > \mu_2 > \cdots > \mu_n > \nu_1 > \nu_2 > \cdots > \nu_n > x > y$. Then, for each $k \in \{0, 1, \ldots, n\}$, we have

    $$\Psi(a_k) = \frac{(-1)^{n-k}}{n!} \sum_{\pi \in S_n} \nu_{\pi(1)}\nu_{\pi(2)} \cdots \nu_{\pi(n-k)}\mu_{\pi(n-k+1)}\mu_{\pi(n-k+2)} \cdots \mu_{\pi(n)}$$

    and therefore

    $$\operatorname{init}(\Psi(a_k)) = \mu_1\mu_2 \cdots \mu_k\nu_{k+1}\nu_{k+2} \cdots \nu_n \tag{48}$$

    (with coefficient $\frac{(-1)^{n-k}}{n!}$, which we ignore). Also, of course, $\operatorname{init}(\Psi(x)) = x$ (since $\Psi(x) = x$) and $\operatorname{init}(\Psi(y)) = y$ (likewise). Hence, for any product

$a_0^{u_0} a_1^{u_1} \cdots a_n^{u_n} x^v y^w$ (with $u_0, u_1, \ldots, u_n, v, w \in \mathbf{N}$), we have

$$\begin{aligned}
&\operatorname{init}\left(\Psi\left(a_0^{u_0} a_1^{u_1} \cdots a_n^{u_n} x^v y^w\right)\right) \\
&= \operatorname{init}\left(\Psi\left(x^v y^w a_0^{u_0} a_1^{u_1} \cdots a_n^{u_n}\right)\right) \\
&= \operatorname{init}\left(\Psi\left(x\right)^v \Psi\left(y\right)^w \Psi\left(a_0\right)^{u_0} \Psi\left(a_1\right)^{u_1} \cdots \Psi\left(a_n\right)^{u_n}\right) \\
&\qquad \text{(since } \Psi \text{ is an algebra morphism)} \\
&= \Bigg(\underbrace{\operatorname{init}\left(\Psi\left(x\right)\right)}_{=x}\Bigg)^v \Bigg(\underbrace{\operatorname{init}\left(\Psi\left(y\right)\right)}_{=y}\Bigg)^w \prod_{k=0}^{n} \Bigg(\underbrace{\operatorname{init}\left(\Psi\left(a_k\right)\right)}_{\substack{=\mu_1\mu_2\cdots\mu_k\nu_{k+1}\nu_{k+2}\cdots\nu_n \\ \text{(by (48))}}}\Bigg)^{u_k} \\
&\qquad \left(\begin{array}{c}\text{since the initial monomial of a product of several polynomials} \\ \text{is the product of the initial monomials of the factors}\end{array}\right) \\
&= \underbrace{x^v y^w \prod_{k=0}^{n} \left(\mu_1\mu_2\cdots\mu_k\nu_{k+1}\nu_{k+2}\cdots\nu_n\right)^{u_k}}_{=\prod_{j=1}^{n} \mu_j^{u_j+u_{j+1}+\cdots+u_n} \nu_j^{u_0+u_1+\cdots+u_{j-1}}} \\
&= x^v y^w \prod_{j=1}^{n} \mu_j^{u_j+u_{j+1}+\cdots+u_n} \nu_j^{u_0+u_1+\cdots+u_{j-1}}.
\end{aligned}$$

We can easily recover the original exponents $u_0, u_1, \ldots, u_n, v, w$ from this monomial (indeed, v and w are simply the exponents of x and y, whereas the exponents $u_j + u_{j+1} + \cdots + u_n$ on the μ_j's allow us to find $u_1, u_2, \ldots, u_n$ by taking differences, and the exponents $u_0 + u_1 + \cdots + u_{j-1}$ on the ν_j's allow us to recover $u_0, u_1, \ldots, u_{n-1}$ by taking differences). Thus, the initial monomials $\operatorname{init}\left(\Psi\left(a_0^{u_0} a_1^{u_1} \cdots a_n^{u_n} x^v y^w\right)\right)$ for all $u_0, u_1, \ldots, u_n, v, w \in \mathbf{N}$ are pairwise distinct. Hence, the polynomials $\Psi\left(a_0^{u_0} a_1^{u_1} \cdots a_n^{u_n} x^v y^w\right)$ are $\mathbf{C}$-linearly independent (since a family of polynomials that have pairwise distinct initial monomials must always be $\mathbf{C}$-linearly independent).

But the family $\left(a_0^{u_0} a_1^{u_1} \cdots a_n^{u_n} x^v y^w\right)_{u_0, u_1, \ldots, u_n, v, w \in \mathbf{N}}$ of monomials is a basis of the $\mathbf{C}$-vector space $\mathbf{C}\left[a_0, a_1, \ldots, a_n, x, y\right]$. We have just proved that the map Ψ sends this basis to a linearly independent family (since the polynomials $\Psi\left(a_0^{u_0} a_1^{u_1} \cdots a_n^{u_n} x^v y^w\right)$ are $\mathbf{C}$-linearly independent). Thus, the map Ψ is injective (since a linear map that sends a basis of its domain to a linearly independent family in its target must always be injective). This proves Lemma 3.6.3 again. ■

80. **page 124, proof of Proposition 3.6.4:** "Then R can be rewritten as

$$R\left(\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y\right) = \left(\mu_1 \ldots \mu_n\right)^d \cdot \widehat{R}\left(\frac{\nu_1}{\mu_1}, \ldots, \frac{\nu_n}{\mu_n}, x, y\right),$$

where $\widehat{R}$ is a symmetric function (in the usual sense) in the ratios $\frac{\nu_i}{\mu_i}$":

Let me explain why this is the case. We define the polynomial $\widehat{R} = \widehat{R}(\alpha_1, \ldots, \alpha_n, x, y) \in \mathbf{C}[\alpha_1, \ldots, \alpha_n, x, y]$ by

$$\widehat{R}(\alpha_1, \ldots, \alpha_n, x, y) := R(1, \alpha_1, 1, \alpha_2, \ldots, 1, \alpha_n, x, y)$$

(that is, $\widehat{R}$ is the evaluation of R at $\mu_i := 1$ and $\nu_i := \alpha_i$). This polynomial $\widehat{R}$ is symmetric in $\alpha_1, \ldots, \alpha_n$, since R is symmetric. Furthermore, since R is regular of degree d, we know that if we multiply the inputs $\mu_1, \nu_1, \mu_2, \nu_2, \ldots, \mu_n, \nu_n$ by some scalars (or indeterminates) $\omega_1, \omega_1, \omega_2, \omega_2, \ldots, \omega_n, \omega_n$, respectively (that is, if we multiply both μ_i and ν_i by ω_i for each $i \in \{1, 2, \ldots, n\}$), then the value $R(\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y)$ gets multiplied by $\omega_1^d \omega_2^d \cdots \omega_n^d$. In other words, for any $\omega_1, \omega_2, \ldots, \omega_n$, we have

$$\begin{aligned} & R(\omega_1\mu_1, \omega_1\nu_1, \ldots, \omega_n\mu_n, \omega_n\nu_n, x, y) \\ &= \left(\omega_1^d \omega_2^d \cdots \omega_n^d\right) R(\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y). \end{aligned}$$

Applying this to $\omega_i = \frac{1}{\mu_i}$, we obtain

$$\begin{aligned} & R\left(\frac{1}{\mu_1}\mu_1, \frac{1}{\mu_1}\nu_1, \ldots, \frac{1}{\mu_n}\mu_n, \frac{1}{\mu_n}\nu_n, x, y\right) \\ &= \underbrace{\left(\left(\frac{1}{\mu_1}\right)^d \left(\frac{1}{\mu_2}\right)^d \cdots \left(\frac{1}{\mu_n}\right)^d\right)}_{=\frac{1}{(\mu_1 \ldots \mu_n)^d}} R(\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y) \\ &= \frac{1}{(\mu_1 \ldots \mu_n)^d} R(\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y). \end{aligned}$$

Solving this for $R(\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y)$, we obtain

$$\begin{aligned} & R(\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y) \\ &= (\mu_1 \ldots \mu_n)^d \cdot R\left(\underbrace{\frac{1}{\mu_1}\mu_1}_{=1}, \underbrace{\frac{1}{\mu_1}\nu_1}_{=\frac{\nu_1}{\mu_1}}, \ldots, \underbrace{\frac{1}{\mu_n}\mu_n}_{=1}, \underbrace{\frac{1}{\mu_n}\nu_n}_{=\frac{\nu_n}{\mu_n}}, x, y\right) \\ &= (\mu_1 \ldots \mu_n)^d \cdot \underbrace{R\left(1, \frac{\nu_1}{\mu_1}, \ldots, 1, \frac{\nu_n}{\mu_n}, x, y\right)}_{\substack{=\widehat{R}\left(\frac{\nu_1}{\mu_1}, \ldots, \frac{\nu_n}{\mu_n}, x, y\right) \\ \text{(by the definition of } \widehat{R})}} \\ &= (\mu_1 \ldots \mu_n)^d \cdot \widehat{R}\left(\frac{\nu_1}{\mu_1}, \ldots, \frac{\nu_n}{\mu_n}, x, y\right). \end{aligned}$$

Thus, we have rewritten R as desired.

81. **page 124, proof of Proposition 3.6.4:** "Multiplying Q by $(\mu_1 \ldots \mu_n)^d$ and distributing factors of $\mu_1 \ldots \mu_n$, we obtain a representation of R as a polynomial function in the magnitudes $\mu_1 \cdots \mu_n \sigma_k \left(\frac{\nu_1}{\mu_1}, \ldots, \frac{\nu_n}{\mu_n}\right)$": This argument is slightly incomplete. If we do what is suggested here, then we obtain a representation of R as a polynomial function in the magnitudes $\mu_1 \cdots \mu_n$ and $\sigma_k \left(\frac{\nu_1}{\mu_1}, \ldots, \frac{\nu_n}{\mu_n}\right)$. In order to rewrite this as a polynomial function in the products $\mu_1 \cdots \mu_n \sigma_k \left(\frac{\nu_1}{\mu_1}, \ldots, \frac{\nu_n}{\mu_n}\right)$, we must make sure that each of the monomials contains at least as many $\mu_1 \cdots \mu_n$'s as it contains $\sigma_k \left(\frac{\nu_1}{\mu_1}, \ldots, \frac{\nu_n}{\mu_n}\right)$'s; otherwise, we get only a Laurent polynomial.

    There are several ways to fix this. In my opinion, the easiest way to prove the "if" direction of Proposition 3.6.4 correctly is to proceed as in the proof of Theorem 1.1.1 (that is, in essence, by showing that the $\Psi(a_0), \Psi(a_1), \ldots, \Psi(a_n)$ form a Sagbi basis of the ring of regular symmetric polynomials $R \in \mathbf{C}[\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y]$ as a $\mathbf{C}[x, y]$-algebra, even though we don't ever have to use this language). Here are the details:

    Proof of the "if" direction of Proposition 3.6.4: We give an algorithm to express each regular symmetric polynomial $R \in \mathbf{C}[\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y]$ as a polynomial function of the images $\Psi(a_0), \Psi(a_1), \ldots, \Psi(a_n)$ over $\mathbf{C}[x, y]$ (which will, of course, show that R lies in the image of Ψ). To do so, we view x and

y as constants, and we equip the polynomial ring $\mathbf{C}[\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y] = (\mathbf{C}[x, y])[\mu_1, \nu_1, \ldots, \mu_n, \nu_n]$ with a monomial order that is lexicographic with $\mu_1 > \mu_2 > \cdots > \mu_n > \nu_1 > \nu_2 > \cdots > \nu_n$. Then, for a nonzero regular symmetric polynomial $R \in \mathbf{C}[\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y]$, the initial monomial $\operatorname{init} R$ must have the form $\mu_1^{u_1}\mu_2^{u_2}\cdots\mu_n^{u_n}\nu_1^{d-u_1}\nu_2^{d-u_2}\cdots\nu_n^{d-u_n}$ (since R is regular) with $u_1 \geq u_2 \geq \cdots \geq u_n$ (since otherwise, by the symmetry of R, we could obtain a larger monomial of R by swapping two of the μ_i's along with the corresponding ν_i's). But this is precisely the initial monomial $\operatorname{init}(R')$ of the polynomial $R' := \prod_{k=0}^{n} (\Psi(a_k))^{u_k - u_{k+1}}$, where we set $u_0 := d$ and $u_{n+1} := 0$ (this follows easily from (48)). Thus, subtracting an appropriate scalar multiple of R' from R, we cancel the initial term of R and obtain a smaller polynomial in $\mathbf{C}[\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y]$ that is still regular and symmetric (since R and R' are regular and symmetric). Repeating this procedure again and again, we eventually end up with the polynomial 0 (since Corollary 1.2.3 shows that this process cannot go on forever). Thus, R equals the sum of all the polynomials we have subtracted, hence a polynomial function of the $\Psi(a_0), \Psi(a_1), \ldots, \Psi(a_n)$ over $\mathbf{C}[x, y]$. Therefore, R lies in the image of Ψ. This finishes the proof. ■

82. **page 125, definition of the bracket $[i\ \mathbf{u}]$:** Replace "$[i\ \mathbf{u}] := \mu_i y - \nu_i x$" by "$[i\ \mathbf{u}] := \mu_i x - \nu_i y$".

83. **page 125, proof of Lemma 3.6.5:** Add commas before "$[1\ n]$", before "$[n-1\ n]$", and before "$[n\ \mathbf{u}]$".

84. **page 125, Theorem 3.6.6:** It is worth pointing out that "bracket polynomial" here means a polynomial in $\mathbf{C}[\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y]$ that lies in the bracket ring. The notions of "symmetric" and "regular" are inherited from $\mathbf{C}[\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y]$. (Hence, in particular, a "symmetric bracket polynomial" means a symmetric polynomial that lies in the bracket ring. Such a polynomial is symmetric under all permutations of the subscripts $1, 2, \ldots, n$ when written as a polynomial in $\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y$, and furthermore can be expressed as a polynomial in the brackets $[i\ j]$ and $[i\ \mathbf{u}]$; but the latter expression might not be invariant under all permutations of the symbols $1, 2, \ldots, n$.)

85. **page 126, proof of Theorem 3.6.6:** "The expansion map Ψ commutes with" should better be "The expansion map Ψ is injective (by Lemma 3.6.3) and commutes with".

86. **page 127, definition of the bracket $\left[i^{(k)}\ \mathbf{u}\right]$:** Replace "$\left[i^{(k)}\ \mathbf{u}\right] := \mu_i^{(k)} x - \mu_j^{(l)} y$" by "$\left[i^{(k)}\ \mathbf{u}\right] := \mu_i^{(k)} x - \nu_i^{(k)} y$".

87. **page 127, middle of the page:** In "$f(x, y) = a_2 x^2 + 2a_1 xy + a_0^2 y$", replace "$a_0^2 y$" by "$a_0 y^2$". Likewise, on the next line, replace "$b_0^2 y$" by "$b_0 y^2$".

88. **page 129, proof of Lemma 3.7.2:** I find this proof confusing (what does "canonical preimage" mean), so let me give my own:

    Proof of Lemma 3.7.2. **(a)** For any polynomial $f \in \mathbf{C}[x_1, \ldots, x_r]$, we let $\overline{f}$ denote its residue class $f + I$ (that is, the projection of f onto the quotient ring $\mathbf{C}[x_1, \ldots, x_r] / I$).

    Clearly, the residue classes $\overline{p_1}, \ldots, \overline{p_m}$ are invariant under Γ (since the polynomials $p_1, \ldots, p_m$ are invariant under Γ).

    Let $\overline{p} \in (\mathbf{C}[x_1, \ldots, x_r] / I)^\Gamma$ be any invariant residue class in $\mathbf{C}[x_1, \ldots, x_r] / I$. We must show that $\overline{p}$ can be written as a polynomial function of $\overline{p_1}, \ldots, \overline{p_m}$.

    The polynomial $p \in \mathbf{C}[x_1, \ldots, x_r]$ itself might not be Γ-invariant, but its image p^* under the Reynolds operator is. Furthermore, $\overline{p^*} = \overline{p}^*$, because the Reynolds operator (being defined as the average of the actions of all elements of Γ) commutes with the projection onto $\mathbf{C}[x_1, \ldots, x_r] / I$. But $\overline{p}^* = \overline{p}$ because $\overline{p}$ is Γ-invariant. Hence, $\overline{p} = \overline{p}^* = \overline{p^*}$.

    But p^* belongs to $\mathbf{C}[x_1, \ldots, x_r]^\Gamma$ and thus can be written as a polynomial function $u(p_1, \ldots, p_m)$ of $p_1, \ldots, p_m$. Hence, $\overline{p}$ can be written as the same polynomial function $u(\overline{p_1}, \ldots, \overline{p_m})$ of $\overline{p_1}, \ldots, \overline{p_m}$ (since $p^* = u(p_1, \ldots, p_m)$ entails $\overline{p^*} = \overline{u(p_1, \ldots, p_m)} = u(\overline{p_1}, \ldots, \overline{p_m})$ and thus $\overline{p} = \overline{p^*} = u(\overline{p_1}, \ldots, \overline{p_m})$). This completes the proof of Lemma 3.7.2 **(a)**.

    (b) In the proof of Noether's degree bound (Theorem 2.1.4), it was shown that the invariant ring $\mathbf{C}[x_1, \ldots, x_r]^\Gamma$ is generated by the Reynolds images $\left(x_1^{i_1} x_2^{i_2} \cdots x_r^{i_r}\right)^*$ of all monomials $x_1^{i_1} x_2^{i_2} \cdots x_r^{i_r}$ with degree $i_1 + i_2 + \cdots + i_r \leq |\Gamma|$. Hence, applying part **(a)** (with the $p_1, \ldots, p_m$ being these Reynolds images), we conclude that the ring $(\mathbf{C}[x_1, \ldots, x_r] / I)^\Gamma$ is generated by the images of these Reynolds images under the canonical surjection $\mathbf{C}[x_1, \ldots, x_r] \to \mathbf{C}[x_1, \ldots, x_r] / I$. This proves Lemma 3.7.2 **(b)**. ■

89. **page 130, proof of Theorem 3.7.1:** In the first sentence of this proof, add commas before "[1 n]", before "[$n-1$ n]", and before "[n $\mathbf{u}$]".

90. **page 130, proof of Theorem 3.7.1:** "Since minimally regular monomials remain minimally regular after permuting letters": This is not literally true, since (e.g.) the transposition $t_{1,2}$ sends the bracket $[1\ 2]$ to $-[1\ 2]$. The action of a permutation $\sigma \in S_n$ sends a bracket monomial to $\pm$ a bracket monomial.

    The easiest way to fix this is by introducing additional brackets $[i\ j]$ for $i > j$. These new brackets are defined in the same way as the old brackets (that is, $[i\ j] := \mu_i \nu_j - \nu_i \mu_j$). Of course, they are redundant as generators, since $[i\ j] = -[j\ i]$ for all $i \neq j$; but they serve to make the set of generators more symmetric. After we introduce these new brackets, the set of bracket monomials (i.e., products of brackets) becomes fixed under the S_n-action that permutes the letters $1, 2, \ldots, n$ (fixing $\mathbf{u}$), and so is the subset of regular bracket monomials; hence, the minimially regular bracket monomials also form an S_n-set. With these changes made, the proof goes through as stated.

91. **page 131, discussion of the monomial order $\prec_{\text{circ}}$:** "Establishing the existence of such a monomial order is a nontrivial exercise": But not a hard exercise at any rate. For instance, we can introduce a (rather exotic) grading on the polynomial ring $\mathbf{C}[\Lambda(n,2)]$ by assigning to each indeterminate $[i\ j]$ the degree $\left(2 - \dfrac{1}{j-i}\right) n!$ (this is a positive integer, because $j - i \in \{1, 2, \ldots, n\}$ is a positive divisor of $n!$). Then, we totally order the bracket monomials as follows: If two monomials $\mathfrak{m}$ and $\mathfrak{n}$ have different degrees, then we set $\mathfrak{m} \succ \mathfrak{n}$ if and only if $\mathfrak{m}$ has larger degree than $\mathfrak{n}$; otherwise, we let $\mathfrak{m} \succ \mathfrak{n}$ if $\mathfrak{m}$ is lexicographically larger than $\mathfrak{n}$. This is easily seen to be a monomial order (indeed, this holds for any grading on a polynomial ring in which each indeterminate is homogeneous and has its degree equal to a positive integer). Moreover, for any $1 \leq i_1 < i_2 < i_3 < i_4 \leq n$, the initial monomial of the syzygy

$$P_{i_1 i_2 i_3 i_4} = [i_1\ i_3]\,[i_2\ i_4] - [i_1\ i_2]\,[i_3\ i_4] - [i_1\ i_4]\,[i_2\ i_3]$$

is $[i_1\ i_3]\,[i_2\ i_4]$. (Indeed, the three monomials $[i_1\ i_3]\,[i_2\ i_4]$, $[i_1\ i_2]\,[i_3\ i_4]$ and $[i_1\ i_4]\,[i_2\ i_3]$ have respective degrees

$$\begin{aligned} &\left(2 - \frac{1}{i_3 - i_1}\right) n! + \left(2 - \frac{1}{i_4 - i_2}\right) n!, \\ &\left(2 - \frac{1}{i_2 - i_1}\right) n! + \left(2 - \frac{1}{i_4 - i_3}\right) n! \qquad \text{and} \\ &\left(2 - \frac{1}{i_4 - i_1}\right) n! + \left(2 - \frac{1}{i_3 - i_2}\right) n!, \end{aligned}$$

but the first of these three degrees is larger than the other two[13].)

[13] Let us show this. We must prove that

$$\begin{aligned} \left(2 - \frac{1}{i_3 - i_1}\right) n! + \left(2 - \frac{1}{i_4 - i_2}\right) n! &> \left(2 - \frac{1}{i_2 - i_1}\right) n! + \left(2 - \frac{1}{i_4 - i_3}\right) n! \qquad \text{and} \\ \left(2 - \frac{1}{i_3 - i_1}\right) n! + \left(2 - \frac{1}{i_4 - i_2}\right) n! &> \left(2 - \frac{1}{i_4 - i_1}\right) n! + \left(2 - \frac{1}{i_3 - i_2}\right) n!. \end{aligned}$$

Upon cancelling the 2's and $n!$'s, these two inequalities rewrite as

$$\begin{aligned} \frac{1}{i_3 - i_1} + \frac{1}{i_4 - i_2} &< \frac{1}{i_2 - i_1} + \frac{1}{i_4 - i_3} \qquad \text{and} \\ \frac{1}{i_3 - i_1} + \frac{1}{i_4 - i_2} &< \frac{1}{i_4 - i_1} + \frac{1}{i_3 - i_2}. \end{aligned}$$

Upon setting $p := i_2 - i_1$ and $q := i_3 - i_2$ and $r := i_4 - i_2$ (so that $i_3 - i_1 = p + q$ and

Thus, the monomial order we have just introduced is (one option for) the monomial order $\prec_{\text{circ}}$ that we need.

92. **page 131, first proof of Proposition 3.7.4:** This is somewhat confused. The set $\mathcal{S}_{n,2}$ is a subset of the bracket polynomial ring $\mathbf{C}[\Lambda(n,2)]$, not of the polynomial ring $\mathbf{C}[\mu_1, \nu_1, \ldots, \mu_n, \nu_n, x, y]$; hence, the brackets $[i\ j]$ are the indeterminates here. Thus, if the initial monomials of $P_{i_1 i_2 i_3 i_4}$ and $P_{j_1 j_2 j_3 j_4}$ are not relatively prime, then the sets $\{[i_1\ i_3], [i_2\ i_4]\}$ and $\{[j_1\ j_3], [j_2\ j_4]\}$ have nonempty intersection, which is a stronger claim than "the set of indices $\{i_1, i_2, i_3, i_4, j_1, j_2, j_3, j_4\}$ has cardinality at most seven" (in fact, it shows that the set of indices $\{i_1, i_2, i_3, i_4, j_1, j_2, j_3, j_4\}$ has cardinality at most 6, but even this is unnecessarily weak a statement). As a consequence, we only need to verify the Gröbner basis property in the cases $n = 5, 6$, not in the cases $n = 5, 6, 7$.

    This is not as painful as it sounds, even without the use of a computer. All we have to do is show that for any two distinct 4-tuples $(i_1 < i_2 < i_3 < i_4)$ and $(j_1 < j_2 < j_3 < j_4)$ such that the initial monomials of $P_{i_1 i_2 i_3 i_4}$ and $P_{j_1 j_2 j_3 j_4}$ are not relatively prime (i.e., the the sets $\{[i_1\ i_3], [i_2\ i_4]\}$ and $\{[j_1\ j_3], [j_2\ j_4]\}$ have nonempty intersection), the $\mathcal{S}$-polynomial $\mathcal{S}\left(P_{i_1 i_2 i_3 i_4}, P_{j_1 j_2 j_3 j_4}\right)$ can be reduced to 0 modulo $\mathcal{S}_{n,2}$. For $n = 5$, the only $\mathcal{S}$-polynomials we need to check are

$$\begin{array}{ll} \mathcal{S}(P_{1,2,3,4}, P_{1,2,3,5}), & \mathcal{S}(P_{1,2,4,5}, P_{1,3,4,5}), \\ \mathcal{S}(P_{1,2,3,4}, P_{2,3,4,5}), & \mathcal{S}(P_{1,2,3,5}, P_{1,2,4,5}), \\ \mathcal{S}(P_{1,3,4,5}, P_{2,3,4,5}) & \end{array}$$

    [14]. For $n = 6$, the only additional $\mathcal{S}$-polynomials we need to check (aside

---

$i_4 - i_2 = q + r$ and $i_4 - i_1 = p + q + r$), these two inequalities rewrite as

$$\frac{1}{p+q} + \frac{1}{q+r} < \frac{1}{p} + \frac{1}{r} \qquad \text{and}$$
$$\frac{1}{p+q} + \frac{1}{q+r} < \frac{1}{p+q+r} + \frac{1}{q}.$$

Since p, q, r are positive (because $i_1 < i_2 < i_3 < i_4$), the first of these two inequalities is obvious (since $\frac{1}{p+q} < \frac{1}{p}$ and $\frac{1}{q+r} < \frac{1}{r}$), whereas the second boils down (upon cross-multiplying and cancelling the numerator $p + 2q + r$) to $(p+q+r)\,q < (p+q)\,(q+r)$, which is again clear (since $(p+q)\,(q+r) - (p+q+r)\,q = pr > 0$). Thus, both inequalities are true.

[14] I have omitted the "mirror versions" of these $\mathcal{S}$-polynomials (where the *mirror version* of an $\mathcal{S}$-polynomial $\mathcal{S}(f, g)$ means the $\mathcal{S}$-polynomial $\mathcal{S}(g, f)$), because these need not be checked separately (in fact, we always have $\mathcal{S}(g, f) = -\mathcal{S}(f, g)$).

There is an additional symmetry in the definitions of $\Lambda(n, 2)$ and $\mathcal{S}_{n,2}$ that we could use to simplify our life (namely, we can replace $[u\ v]$ by $[v'\ u']$, where $u' = n + 1 - u$ and $v' = n + 1 - v$). We leave it to the reader to check how it can be exploited.

from the above ones and their variants with relabelled variables) are

$$\begin{array}{ll} \mathcal{S}\left(P_{1,2,3,5}, P_{2,4,5,6}\right), & \mathcal{S}\left(P_{1,2,4,5}, P_{2,3,5,6}\right), \\ \mathcal{S}\left(P_{1,3,4,6}, \mathcal{S}_{2,3,5,6}\right), & \mathcal{S}\left(P_{1,3,5,6}, P_{2,3,4,6}\right), \\ \mathcal{S}\left(P_{1,2,4,5}, P_{1,3,4,6}\right), & \mathcal{S}\left(P_{1,2,4,6}, P_{1,3,4,5}\right) \end{array}$$

[15]. In each case, we must check that the respective $\mathcal{S}$-polynomial can be reduced to 0 modulo $\mathcal{S}_{n,2}$. According to a known criterion[16], it suffices to express these $\mathcal{S}$-polynomials in the form $\sum_{s\in S} a_s \mathfrak{s}_s g_s$, where S is a finite set, where $a_s \in \mathbf{C}$ are scalars, where $\mathfrak{s}_s$ are monomials and where g_s are (not necessarily distinct) elements of $\mathcal{S}_{n,2}$ such that all the monomials $\mathfrak{s}_s \cdot \operatorname{init}(g_s)$ are distinct (where init means the initial monomial with respect to the monomial order $\prec_{\text{circ}}$). But such expressions can be explicitly provided:

$$\begin{aligned} \mathcal{S}\left(P_{1,2,3,4}, P_{1,2,3,5}\right) &= [2\ 5]\ \ P_{1,2,3,4} - [2\ 4]\ \ P_{1,2,3,5} \\ &= [1\ 2]\ \ P_{2,3,4,5} - [2\ 3]\ \ P_{1,2,4,5} \end{aligned}$$

(with [1 2] and [2 3] playing the role of the monomials $\mathfrak{s}_s$ and with $P_{2,3,4,5}$ and $P_{1,2,4,5}$ acting as the g_s; it is easy to see that the monomials $[1\ 2] \cdot \operatorname{init} P_{2,3,4,5} = [1\ 2] \cdot [2\ 4] \cdot [3\ 5]$ and $[2\ 3] \cdot \operatorname{init} P_{1,2,4,5} = [2\ 3] \cdot [1\ 4] \cdot [2\ 5]$ are distinct) and similarly

$$\begin{aligned}
\mathcal{S}\left(P_{1,2,4,5}, P_{1,3,4,5}\right) &= [4\ 5]\ \ P_{1,2,3,5} - [1\ 5]\ \ P_{2,3,4,5}; \\
\mathcal{S}\left(P_{1,2,3,4}, P_{2,3,4,5}\right) &= [3\ 4]\ \ P_{1,2,3,5} - [2\ 3]\ \ P_{1,3,4,5}; \\
\mathcal{S}\left(P_{1,2,3,5}, P_{1,2,4,5}\right) &= [1\ 5]\ \ P_{1,2,3,4} - [1\ 2]\ \ P_{1,3,4,5}; \\
\mathcal{S}\left(P_{1,3,4,5}, P_{2,3,4,5}\right) &= [3\ 4]\ \ P_{1,2,4,5} - [4\ 5]\ \ P_{1,2,3,4}; \\
\mathcal{S}\left(P_{1,2,3,5}, P_{2,4,5,6}\right) &= [5\ 6]\ \ P_{1,2,3,4} + [4\ 5]\ \ P_{1,2,3,6} - [2\ 3]\ \ P_{1,4,5,6} - [1\ 2]\ \ P_{3,4,5,6}; \\
\mathcal{S}\left(P_{1,2,4,5}, P_{2,3,5,6}\right) &= [3\ 5]\ \ P_{1,2,4,6} - [2\ 4]\ \ P_{1,3,5,6} - [5\ 6]\ \ P_{1,2,3,4} + [1\ 2]\ \ P_{3,4,5,6}; \\
\mathcal{S}\left(P_{1,3,4,6}, P_{2,3,5,6}\right) &= [1\ 6]\ \ P_{2,3,4,5} - [2\ 3]\ \ P_{1,4,5,6} - [4\ 6]\ \ P_{1,2,3,5} + [3\ 5]\ \ P_{1,2,4,6}; \\
\mathcal{S}\left(P_{1,3,5,6}, P_{2,3,4,6}\right) &= -\ [1\ 6]\ \ P_{2,3,4,5} + [2\ 3]\ \ P_{1,4,5,6} - [5\ 6]\ \ P_{1,2,3,4} + [3\ 4]\ \ P_{1,2,5,6}; \\
\mathcal{S}\left(P_{1,2,4,5}, P_{1,3,4,6}\right) &= -\ [1\ 6]\ \ P_{2,3,4,5} + [4\ 5]\ \ P_{1,2,3,6} + [1\ 3]\ \ P_{2,4,5,6} - [2\ 4]\ \ P_{1,3,5,6}; \\
\mathcal{S}\left(P_{1,2,4,6}, P_{1,3,4,5}\right) &= -\ [1\ 6]\ \ P_{2,3,4,5} + [4\ 5]\ \ P_{1,2,3,6} - [1\ 2]\ \ P_{3,4,5,6} + [3\ 4]\ \ P_{1,2,5,6}.
\end{aligned}$$

93. **page 132, Lemma 3.7.5:** This lemma is false. A counterexample is the bracket monomial

$$[1\ 2]\ [1\ 3]\ [2\ 3]^2 \cdot [4\ 5]\ [4\ 6]\ [5\ 6]^2 \cdot [7\ 8]\ [7\ 9]\ [8\ 9]^2 \cdot [1\ \mathbf{x}]\ [4\ \mathbf{x}]\ [7\ \mathbf{x}]$$

in $\mathbf{C}\left[\Lambda\left(10,2\right)\right]$, which is regular of degree 3 and elemental.

[15] Again, I have omitted the "mirror versions" of these $\mathcal{S}$-polynomials.

[16] See, e.g., Lemma 3.6 in Darij Grinberg, *t-unique reductions for Mészáros's subdivision algebra [detailed version]*, ancillary file to arXiv:1704.00839v6. (This lemma appears as Lemma 4.6 in the published version, but without the proof.)

However, there is an alternative way of proving Theorem 3.7.3, which bypasses Lemma 3.7.5. It goes back to the paper B. Howard, J. Millson, A. Snowden, R. Vakil, *The equations for the moduli space of n points on the line*, Duke Math. J. **146** (2009), no. 2, pp. 175–226 (see also the proof of Theorem 31 in Giorgio Ottaviani, *Five Lectures on Projective Invariants*, arXiv:1305.2749v1). Here is an outline:

Correct proof of Theorem 3.7.3 (sketched). We rename the letter $\mathbf{u}$ as $n+1$ and rename the indeterminates y and x as μ_{n+1} and ν_{n+1} (so that our brackets have the form $[i\ j]$ for $1 \leq i < j \leq n+1$). Furthermore, we rename n as $n-1$ (so that our brackets now have the form $[i\ j]$ for $1 \leq i < j \leq n$). We also introduce the shorthand $[i\ j] := [j\ i]$ for all $i > j$, so that the bracket $[i\ j]$ is defined for all $i \neq j$ (not only for $i < j$). (Note that this differs from the convention I suggested on page 130.)

It is easy to see (indeed, it follows from Example 3.1.6 upon renaming the variables) that

$$[i_1\ i_3]\,[i_2\ i_4] = [i_1\ i_2]\,[i_3\ i_4] + [i_1\ i_4]\,[i_2\ i_3] \tag{49}$$

for any $1 \leq i_1 < i_2 < i_3 < i_4 \leq n$. Hence, for any four distinct elements $\alpha, \beta, \gamma, \delta$ of $\{1, 2, \ldots, n\}$, we have

$$[\alpha\ \gamma]\,[\beta\ \delta] = \pm\,[\alpha\ \beta]\,[\gamma\ \delta] \pm [\alpha\ \delta]\,[\beta\ \gamma] \tag{50}$$

for an appropriate choice of $\pm$ signs (the two $\pm$ signs need not be the same sign). (Indeed, (50) follows by applying (49) with i_1, i_2, i_3, i_4 being the numbers $\alpha, \beta, \gamma, \delta$ listed in increasing order, and then solving for $[\alpha\ \gamma]\,[\beta\ \delta]$.)

In the following, a *lowlie* will mean a bracket monomial that is regular of degree 1 or 2. For example, $[1\ 2]\,[2\ 3] \cdots [n-1\ n]\,[n\ 1]$ is a lowlie that is regular of degree 2. Lowlies that are regular of degree 1 exist only when n is even; an example of such a lowlie is $[1\ 2]\,[3\ 4] \cdots [n-1\ n]$.

We view any bracket monomial $[i_1\ j_1]\,[i_2\ j_2] \cdots [i_k\ j_k]$ as a multigraph with vertex set $\{1, 2, \ldots, n\}$ and edges $\{i_1, j_1\}$, $\{i_2, j_2\}$, $\ldots$, $\{i_k, j_k\}$. Such multigraphs can have parallel edges, but cannot have loops. A bracket monomial is regular of degree d if and only if it is a d-regular multigraph (i.e., each vertex has degree d).

Now we want to prove Theorem 3.7.3. This theorem claims that the ring $\mathcal{B}_{\text{reg}}$ of regular bracket polynomials is generated by the lowlies. In other words, it claims that every bracket polynomial that is regular of degree d (for some $d \in \mathbf{N}$) is a polynomial function of the lowlies. We shall prove this by distinguishing between two cases, depending on the parity of d:

- *Case 1:* The number d is even. Hence, Petersen's 2-factor theorem shows that every d-regular multigraph (that is, every bracket monomial that is regular of degree d) can be partitioned into $d/2$ edge-disjoint 2-factors, i.e., (regarded as a bracket monomial), can be factored as a product of $d/2$ bracket monomials that are regular of degree

2. Of course, the latter monomials are lowlies. Hence, we have shown that every bracket monomial that is regular of degree d is a product of lowlies. Thus, every bracket polynomial that is regular of degree d is a polynomial function of the lowlies. So our proof is complete in Case 1.

- *Case 2:* The number d is odd. We must prove that every bracket polynomial that is regular of degree d is a polynomial function of the lowlies. Clearly, it is enough to prove this for bracket monomials. Thus, we let $\mathfrak{m}$ be a bracket monomial that is regular of degree d. As a multigraph, $\mathfrak{m}$ is thus a d-regular multigraph with n vertices. Since the sum of the degrees of all vertices in a multigraph is even, we thus conclude that nd is even, so that n must be even (since d is odd).

  If the multigraph $\mathfrak{m}$ is bipartite, then Frobenius's matching theorem shows that $\mathfrak{m}$ has a perfect matching (since $\mathfrak{m}$ is d-regular), i.e., a spanning 1-regular subgraph. Rewritten in terms of brackets, this is saying that if $\mathfrak{m}$ is bipartite, then $\mathfrak{m}$ is divisible by a bracket monomial $\mathfrak{n}$ that is regular of degree 1. The quotient $\mathfrak{m}/\mathfrak{n}$ of this division must then be a bracket monomial that is regular of degree $d-1$, and hence is a polynomial function of the lowlies (by Case 1, since $d-1$ is even). Multiplying this polynomial function by $\mathfrak{n}$ (which itself is a lowlie), we obtain an expression for $\mathfrak{m}$ as a polynomial function of the lowlies. Thus, our proof is finished if the multigraph $\mathfrak{m}$ is bipartite.

  What can we do if $\mathfrak{m}$ is not bipartite? In this case, $\mathfrak{m}$ contains at least one edge $\{e_1, e_2\}$ whose both endpoints are even[17], as well as at least one edge $\{o_1, o_2\}$ whose both endpoints are odd[18]. Consider these two edges. Thus, the bracket monomial $\mathfrak{m}$ is divisible by $[e_1\ e_2]\ [o_1\ o_2]$. In other words,

  $$\mathfrak{m} = [e_1\ e_2]\ [o_1\ o_2]\ \mathfrak{p} \tag{51}$$

  for some bracket monomial $\mathfrak{p}$. Consider this $\mathfrak{p}$. But (50) (applied to

[17] Recall that n is even, so the even vertices are $2, 4, 6, \ldots, n$, while the odd vertices are $1, 3, 5, \ldots, n-1$.

[18] Indeed:

- If neither $\{e_1, e_2\}$ nor $\{o_1, o_2\}$ existed, then $\mathfrak{m}$ would be bipartite (with all the even vertices being left vertices, and all the odd vertices being right vertices); but we have assumed that $\mathfrak{m}$ is not.
- If $\{e_1, e_2\}$ existed but $\{o_1, o_2\}$ did not, then the sum of the degrees of all odd vertices would equal the number of even-odd edges (i.e., edges that have an even and an odd endpoint), whereas the sum of the degrees of all even vertices would be larger than the number of even-odd edges (since it would count the edge $\{e_1, e_2\}$ as well); but this is impossible, since both of these sums are $\underbrace{d + d + \cdots + d}_{n/2 \text{ times}} = dn/2$.
- If $\{o_1, o_2\}$ existed but $\{e_1, e_2\}$ did not, then we would obtain a similar contradiction.

So the only possibility is that both $\{e_1, e_2\}$ and $\{o_1, o_2\}$ exist.

$\alpha = e_1$, $\beta = o_1$, $\gamma = e_2$ and $\delta = o_2$) shows that

$$[e_1\ e_2]\,[o_1\ o_2] = \pm\,[e_1\ o_1]\,[e_2\ o_2] \pm [e_1\ o_2]\,[o_1\ e_2]\,.$$

Hence, we can rewrite (51) as

$$\begin{aligned}\mathfrak{m} &= (\pm\,[e_1\ o_1]\,[e_2\ o_2] \pm [e_1\ o_2]\,[o_1\ e_2])\ \mathfrak{p} \\ &= \pm\,[e_1\ o_1]\,[e_2\ o_2]\ \mathfrak{p} \pm [e_1\ o_2]\,[o_1\ e_2]\ \mathfrak{p}. \end{aligned} \tag{52}$$

However, both bracket monomials $[e_1\ o_1]\,[e_2\ o_2]\ \mathfrak{p}$ and $[e_1\ o_2]\,[o_1\ e_2]\ \mathfrak{p}$ are regular of degree d again (since they are obtained from $[e_1\ e_2]\,[o_1\ o_2]\ \mathfrak{p} = \mathfrak{m}$ by replacing the $[e_1\ e_2]\,[o_1\ o_2]$ factor with $[e_1\ o_1]\,[e_2\ o_2]$ or $[e_1\ o_2]\,[o_1\ e_2]$, respectively; but this replacement takes away equally many edges from each vertex as it adds to that vertex), but have a higher number of even-odd edges[19] than $\mathfrak{m}$ does (since all the four new edges $\{e_1, o_1\}$, $\{e_2, o_2\}$, $\{e_1, o_2\}$, $\{o_1, e_2\}$ are even-odd edges, while the original two edges $\{e_1, e_2\}$ and $\{o_1, o_2\}$ are not). By descending induction on the number of even-odd edges, we can thus assume that $[e_1\ o_1]\,[e_2\ o_2]\ \mathfrak{p}$ and $[e_1\ o_2]\,[o_1\ e_2]\ \mathfrak{p}$ are polynomial functions of the lowlies (the base case is the case when $\mathfrak{m}$ is bipartite; this case has already been handled). Then, the equality (52) shows that $\mathfrak{m}$ is a polynomial function of the lowlies as well. This completes our proof in Case 2.

Hence, the proof of Theorem 3.7.3 is complete in both cases. ■

(Note that Case 1 and Case 2 in the above proof are not as different as they look, since the proof of Petersen's 2-factor theorem also relies on Frobenius's matching theorem.)

(Note also that the proof of Theorem 3.7.3 given in §6.3 of Joseph P. S. Kung, Gian-Carlo Rota, *The invariant theory of binary forms*, Bull. Amer. Math. Soc. (N.S.) **10**(1) (1984), pp. 27–85. appears fairly similar to the above.)

94. **page 135, exercise (4):** "set for" should be "sets for".

95. **page 137:** "A comprehensive introduction with many geometric applications can be found in Fulton and Harris (1991)": Here are a few other references (which I found more readable than Fulton and Harris):

    - William Fulton, *Young Tableaux, With Applications to Representation Theory and Geometry*, Cambridge University Press 1999, Chapter 8. (Errata.)

[19] An *even-odd edge* means an edge that has an even and an odd endpoint.

- Pavel Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, Elena Yudovina, *Introduction to Representation Theory*, Student Mathematical Library **59**, AMS 2011, Sections §5.19–5.23.
- William Crawley-Boevey, *Lectures on representation theory and invariant theory*, 1999. (Errata.)
- Hanspeter Kraft, Claudio Procesi, *Classical Invariant Theory: A Primer*, July 1996. (Errata.)

96. **page 138:** "Note that $\det(A)^{n-1}$ is a common denominator" should be "Note that $\det(A)$ is a common denominator".

97. **page 139, Examples 4.1.5 (e):** Replace "$V = S_2\mathbf{C}^3$" by "$V = S_3\mathbf{C}^2$".

98. **page 140, definition of irreducibility:** "We say that (V, ρ) is *irreducible* if" should be "We say that a nonzero Γ-representation (V, ρ) is *irreducible* if".

99. **page 140:** After "For instance, the representations $S_d\mathbf{C}^n$ and $\wedge_d\mathbf{C}^n$ are irreducible", add "(except that $\wedge_d\mathbf{C}^n = 0$ for $d > n$)".

100. **page 140, Theorem 4.1.7:** "of irreducible representation" should be "of irreducible representations".

101. **page 141, definition of standard Young tableaux:** It is worth pointing out that the notion of a "standard Young tableaux" as defined here is different from than the notion of a "standard tableau" defined in §3.1. (The latter tableaux can have equal entries, while the former can have non-rectangular shapes.)

102. **page 141, definition of the Young symmetrizer:** The Young symmetrizer c_T is not "an idempotent linear map", but only a quasi-idempotent linear map – i.e., we have $c_T^2 = kc_T$ for some nonzero rational number k. Specifically, this factor k is the positive integer $n!/f^\lambda$, where f^λ is the number of SYTλ. Thus, the scalar multiple $\dfrac{f^\lambda}{n!}c_T$ of c_T is an actual idempotent. Maybe it is this multiple that you want to call the "Young symmetrizer".

    (It is worth noting that many authors define the Young symmetrizer to be not the linear operator c_T, but rather the element

    $$\sum_{\sigma \in \operatorname{colstb}(T)} \sum_{T \in \operatorname{rowstb}(T)} (\operatorname{sign} \sigma) \cdot \sigma\tau \qquad \text{in the group algebra } \mathbf{C}[S_d],$$

    whose action on $\otimes_d\mathbf{C}^n$ (from the right, by permuting tensorands) is your linear operator c_T. Some also define it as the element

    $$\sum_{\sigma \in \operatorname{colstb}(T)} \sum_{T \in \operatorname{rowstb}(T)} (\operatorname{sign} \sigma) \cdot \tau\sigma \qquad \text{in the group algebra } \mathbf{C}[S_d],$$

whose action on $\otimes_d \mathbf{C}^n$ (now from the left, by permuting tensorands) is your linear operator c_T.)

103. **page 141:** "Thus $W_T \mathbf{C}^n$ is the subspace of all tensors in $\otimes_d \mathbf{C}^n$ which are symmetric with respect to the rows of T and antisymmetric with respect to the columns of T": This is false. For instance, for $n = 2$ and $d = 3$ and $\lambda = (2,1)$ and $T = \begin{matrix} 1 & 2 \\ 3 & \end{matrix}$, the only tensor in $\otimes_d \mathbf{C}^n$ that is symmetric in its first two tensor factors and antisymmetric in its first and third tensor factors is 0, but $W_T \mathbf{C}^n$ is not 0.

104. **page 141, Theorem 4.1.11:** Let me outline a proof of this theorem, as it is not easily found in the literature.

   Proof of Theorem 4.1.11 (sketched). We shall show that

$$\otimes_d \mathbf{C}^n = \bigoplus_{\lambda \vdash d} \bigoplus_{T \text{ SYT} \lambda} W_T \mathbf{C}^n \qquad (53)$$

   (an internal direct sum, not just an isomorphism). For this purpose, we define some notations. If T is any standard Young tableau with d entries, then $r(T)$ shall denote the sequence $(r_d, r_{d-1}, \ldots, r_1) \in \mathbf{N}^d$, where r_i is the number of the row of T that contains the entry i (that is, entry i appears in the r_i-th row of T). For example, $r\begin{pmatrix} 1 & 3 & 4 \\ 2 & 5 & \\ 6 & 7 & \end{pmatrix} = (3,3,2,1,1,2,1)$. Clearly, a standard Young tableau T is uniquely determined by this sequence $r(T)$ (since $r(T)$ tells us which row contains which entries, and the order of the entries in a row must be increasing in order for T to be standard).

   Now, consider the set

$$\text{SYT}(d) := \bigcup_{\lambda \vdash d} \{T \text{ SYT} \lambda\}$$

   of all standard Young tableaux of shape λ for all partitions λ of d. We equip this set $\text{SYT}(d)$ with the following total order: For two standard Young tableaux S and T of shapes λ and μ, we say that $S > T$ if and only if

   - either $\lambda > \mu$ in the lexicographic order on $\mathbf{N}^n$,
   - or $\lambda = \mu$ and $r(S) > r(T)$ in the lexicographic order on $\mathbf{N}^d$.

   It is clear that this gives a well-defined total order on $\text{SYT}(d)$ (in fact, it is simply the lexicographic order for the concatenations $\lambda \oplus r(T)$).

Next, for each $T \in \operatorname{SYT}(d)$, we let $\mathbf{E}_T$ denote the element

$$\left(\sum_{\sigma \in \operatorname{colstb}(T)} (\operatorname{sign} \sigma) \cdot \sigma \right) \left(\sum_{T \in \operatorname{rowstb}(T)} \tau \right) = \sum_{\sigma \in \operatorname{colstb}(T)} \sum_{T \in \operatorname{rowstb}(T)} (\operatorname{sign} \sigma) \cdot \sigma \tau \tag{54}$$

in the group algebra $\mathbf{C}[S_d]$. This element $\mathbf{E}_T$ is called the *Young symmetrizer* of T (and is denoted by $\mathbf{E}_T$) in my notes Darij Grinberg, *An introduction to the symmetric group algebra [Math 701, Spring 2024 lecture notes]*, arXiv:2507.20706v1 (which I shall henceforth cite as [sga]). Note that this element $\mathbf{E}_T$ is quasi-idempotent, and specifically, its square is

$$\mathbf{E}_T^2 = \frac{n!}{f^\lambda} \mathbf{E}_T, \tag{55}$$

where f^λ is the number of $\text{SYT}\lambda$ (by Theorem 5.11.3 in [sga]). The Young symmetrizer c_T as you define it is the action of this element $\mathbf{E}_T$ on $\otimes_d \mathbf{C}^n$ from the right (by permuting tensorands). Thus,

$$W_T \mathbf{C}^n = c_T \cdot \otimes_d \mathbf{C}^n = (\otimes_d \mathbf{C}^n) \, \mathbf{E}_T \tag{56}$$

for each $T \in \operatorname{SYT}(d)$.

Now, it can be shown that any two standard Young tableaux $S, T \in \operatorname{SYT}(d)$ satisfy

$$\mathbf{E}_S \mathbf{E}_T = 0 \qquad \text{if} \qquad S > T \tag{57}$$

(with respect to the above-defined total order on $\operatorname{SYT}(d)$). Indeed, letting S and T have shapes λ and μ, respectively, we can argue as follows:

- If $\lambda > \mu$ in the lexicographic order on $\mathbf{N}^n$, then $\lambda \neq \mu$, and thus $\mathbf{E}_S \mathbf{E}_T = 0$ follows from Proposition 5.11.15 in [sga] (which makes the even stronger claim that $\mathbf{E}_S \mathbf{a} \mathbf{E}_T = 0$ for each $\mathbf{a} \in \mathbf{C}[S_d]$).
- If $\lambda = \mu$ and $r(S) > r(T)$ in the lexicographic order on $\mathbf{N}^d$, then $\mathbf{E}_S \mathbf{E}_T = 0$ follows from Lemma 5.15.21 in [sga], since $r(S) > r(T)$ is saying that $\overline{S} > \overline{T}$ with respect to the Young last letter order on n-tabloids (see [sga] for details).

In either case, we get $\mathbf{E}_S \mathbf{E}_T = 0$, so that (57) is proved.

Now, we claim the following:

Claim 1: The sum

$$\sum_{T \in \operatorname{SYT}(d)} W_T \mathbf{C}^n \tag{58}$$

of subspaces of $\otimes_d \mathbf{C}^n$ is a direct sum.

Claim 2: We have

$$\otimes_d \mathbf{C}^n = \sum_{T \in \operatorname{SYT}(d)} W_T \mathbf{C}^n.$$

Proof of Claim 1. Assume the contrary. Thus, there exists a family $(w_T)_{T \in \operatorname{SYT}(d)}$ of vectors $w_T \in W_T \mathbf{C}^n$ such that not all these vectors w_T are zero, but the sum $\sum\limits_{T \in \operatorname{SYT}(d)} w_T$ is 0. Consider such a family. Thus, there exists some $Q \in \operatorname{SYT}(d)$ such that $w_Q \neq 0$. Pick the smallest such Q (with respect to the above-defined total order). Thus,

$$w_P = 0 \qquad \text{for all } P < Q. \tag{59}$$

We have $w_Q \in W_Q \mathbf{C}^n = (\otimes_d \mathbf{C}^n) \mathbf{E}_Q$ (by (56)). But (55) shows that $\mathbf{E}_Q^2 = \dfrac{n!}{f^\lambda} \mathbf{E}_Q$. Hence, each $a \in W_Q \mathbf{C}^n$ satisfies $a \mathbf{E}_Q = \dfrac{n!}{f^\lambda} a$ (because $a \in W_Q \mathbf{C}^n = (\otimes_d \mathbf{C}^n) \mathbf{E}_Q$ allows us to write a as $a = b \mathbf{E}_Q$ for some $b \in \otimes_d \mathbf{C}^n$, and therefore we have $a \mathbf{E}_Q = b \underbrace{\mathbf{E}_Q \mathbf{E}_Q}_{= \mathbf{E}_Q^2 = \frac{n!}{f^\lambda} \mathbf{E}_Q} = \dfrac{n!}{f^\lambda} \underbrace{b \mathbf{E}_Q}_{= a} = \dfrac{n!}{f^\lambda} a$). Applying this to $a = w_Q$, we obtain $w_Q \mathbf{E}_Q = \dfrac{n!}{f^\lambda} w_Q \neq 0$ (since $\dfrac{n!}{f^\lambda} \neq 0$ and $w_Q \neq 0$).

On the other hand, if $P \in \operatorname{SYT}(d)$ satisfies $P > Q$, then $\mathbf{E}_P \mathbf{E}_Q = 0$ (by (57)) and thus

$$\underbrace{w_P}_{\substack{\in W_P \mathbf{C}^n \\ = (\otimes_d \mathbf{C}^n) \mathbf{E}_P}} \mathbf{E}_Q \in (\otimes_d \mathbf{C}^n) \underbrace{\mathbf{E}_P \mathbf{E}_Q}_{= 0} = 0,$$

so that

$$w_P \mathbf{E}_Q = 0. \tag{60}$$

Now, we assumed that the sum $\sum\limits_{T \in \operatorname{SYT}(d)} w_T$ is 0. Thus, $0 = \sum\limits_{T \in \operatorname{SYT}(d)} w_T = \sum\limits_{P \in \operatorname{SYT}(d)} w_P$. Multiplying this equality by $\mathbf{E}_Q$ from the right, we obtain

$$\begin{aligned}
0 &= \left(\sum_{P \in \operatorname{SYT}(d)} w_P \right) \mathbf{E}_Q = \sum_{P \in \operatorname{SYT}(d)} w_P \mathbf{E}_Q \\
&= \sum_{\substack{P \in \operatorname{SYT}(d); \\ P < Q}} \underbrace{w_P}_{\substack{= 0 \\ \text{(by (59))}}} \mathbf{E}_Q + w_Q \mathbf{E}_Q + \sum_{\substack{P \in \operatorname{SYT}(d); \\ P > Q}} \underbrace{w_P \mathbf{E}_Q}_{\substack{= 0 \\ \text{(by (60))}}} \\
&\qquad \left(\begin{array}{c} \text{since each } P \in \operatorname{SYT}(d) \text{ satisfies} \\ \text{either } P < Q \text{ or } P = Q \text{ or } P > Q \end{array} \right) \\
&= w_Q \mathbf{E}_Q \neq 0.
\end{aligned}$$

This obvious contradiction shows that our assumption was false; hence, Claim 1 is proved. □

Proof of Claim 2. Let $\mathcal{A} = \mathbf{C}[S_d]$. From Corollary 5.15.25 in [sga], we know that

$$\mathcal{A} = \underbrace{\bigoplus_{\lambda \vdash d} \bigoplus_{T \text{ SYT} \lambda}}_{= \bigoplus_{T \in \text{SYT}(d)}} \mathcal{A}\mathbf{E}_T = \bigoplus_{T \in \text{SYT}(d)} \mathcal{A}\mathbf{E}_T = \sum_{T \in \text{SYT}(d)} \mathcal{A}\mathbf{E}_T.$$

Hence, in particular, $1_{\mathcal{A}} \in \sum_{T \in \text{SYT}(d)} \mathcal{A}\mathbf{E}_T$, so that we can write $1_{\mathcal{A}}$ in the form

$$1_{\mathcal{A}} = \sum_{T \in \text{SYT}(d)} a_T \mathbf{E}_T \qquad \text{for some } a_T \in \mathcal{A}. \tag{61}$$

Consider these a_T. Now, for each $w \in \otimes_d \mathbf{C}^n$, we have

$$\begin{aligned} w = w1_{\mathcal{A}} &= w \sum_{T \in \text{SYT}(d)} a_T \mathbf{E}_T \qquad \text{(by (61))} \\ &= \sum_{T \in \text{SYT}(d)} \underbrace{w a_T \mathbf{E}_T}_{\substack{\in (\otimes_d \mathbf{C}^n) \mathbf{E}_T \\ = W_T \mathbf{C}^n}} \in \sum_{T \in \text{SYT}(d)} W_T \mathbf{C}^n. \end{aligned}$$

In other words, $\otimes_d \mathbf{C}^n \subseteq \sum_{T \in \text{SYT}(d)} W_T \mathbf{C}^n$. Hence,

$$\otimes_d \mathbf{C}^n = \sum_{T \in \text{SYT}(d)} W_T \mathbf{C}^n.$$

This proves Claim 2. □

Combining Claim 1 with Claim 2, we obtain

$$\begin{aligned} \otimes_d \mathbf{C}^n &= \bigoplus_{T \in \text{SYT}(d)} W_T \mathbf{C}^n \qquad \text{(internal direct sum)} \\ &= \bigoplus_{\lambda \vdash d} \bigoplus_{T \text{ SYT} \lambda} W_T \mathbf{C}^n. \end{aligned}$$

This proves (53) and therefore also Theorem 4.1.11. ■

105. **page 142, (4.1.5):** Replace "$c_{\substack{1\\2}} \oplus c_{1\,2}$" by "$c_{1\,2} \oplus c_{\substack{1\\2}}$" in order to match the order of the addends to the right of the arrow.
Also, it would be better to multiply "$(v_1 \otimes v_2 + v_2 \otimes v_1) + (v_1 \otimes v_2 - v_2 \otimes v_1)$" by $\frac{1}{2}$, so that the map becomes the identity map.

106. **page 142, Theorem 4.1.12:** It is worth saying that this theorem is a direct consequence of Corollary 4.1.14 (1) (which I prove below) and of the direct sum decomposition (53).

107. **page 143, Corollary 4.1.14:** Let me give a reference for the proof of this corollary, since it is not easy to locate in the literature.

I will use the book William Fulton, *Young Tableaux, With Applications to Representation Theory and Geometry*, Cambridge University Press 1999, (see also the errata). In the following, this book will be cited as [Fulton].

Proof of Corollary 4.1.14. **(1)** Set $E := \mathbf{C}^n$. Let $\mathbf{E}_T \in \mathbf{C}[S_d]$ be the Young symmetrizer defined in (54). By the definition of $W_T\mathbf{C}^n$, we have

$$\begin{aligned} W_T\mathbf{C}^n &= c_T\left(\otimes_d \mathbf{C}^n\right) && \text{(by the definition of } W_T\mathbf{C}^n\text{)} \\ &= c_T\left(\otimes_d E\right) && \text{(since } \mathbf{C}^n = E\text{)} \\ &= c_T\left(E^{\otimes d}\right) && \left(\begin{array}{c}\text{here, we switched from the notation } \otimes_d E \text{ to the} \\ \text{more convenient notation } E^{\otimes d} \text{ for the same thing}\end{array}\right) \\ &= E^{\otimes d}\mathbf{E}_T && \text{(since } c_T \text{ is right multiplication by } \mathbf{E}_T\text{)}. \end{aligned}$$

Note that [Fulton] denotes $\mathbf{E}_T$ as c_T.

Now, [Fulton] considers the Γ-module E^λ, about which he claims (on page 119 of [Fulton], in the paragraph below Exercise 11 in §8.3) that (in his notations, which are different from ours!) "E^λ is isomorphic to the image of the map $E^{\otimes n} \to E^{\otimes n}$ that is right multiplication by c_U". Translated into our notations (noting that our $d, n, E, T, U, \mathbf{E}_T$ correspond to [Fulton]'s $n, d, \mathbf{C}^n, U, T, c_U$), this is saying that [Fulton]'s E^λ is isomorphic to the image of the map $E^{\otimes d} \to E^{\otimes d}$ that is right multiplication by $\mathbf{E}_T$. In other words, [Fulton]'s E^λ is isomorphic to $E^{\otimes d}\mathbf{E}_T = W_T\mathbf{C}^n$. This isomorphism is constructed as follows:

$$\begin{aligned} E^\lambda &\cong E\left(S^\lambda\right) && \text{(by Proposition 1 in [Fulton]'s §8.3)} \\ &= E^{\otimes d} \otimes_{\mathbf{C}[S_d]} \underbrace{S^\lambda}_{\cong \mathbf{C}[S_d]\cdot \mathbf{E}_T} && \text{(by the definition of the } E \text{ functor)} \\ &\cong E^{\otimes d} \otimes_{\mathbf{C}[S_d]} \left(\mathbf{C}[S_d] \cdot \mathbf{E}_T\right) \\ &\cong E^{\otimes d}\mathbf{E}_T \end{aligned} \tag{62}$$

(where the last $\cong$ sign is a particular case of the isomorphism $M \otimes_A (Ae) \cong Me$ for each ring A, each right A-module M and each idempotent $e \in A$).

Using the above isomorphism $E^\lambda \to W_T\mathbf{C}^n$, the basis of E^λ constructed in Theorem 1 of [Fulton]'s §8.1 can be translated into a basis of $W_T\mathbf{C}^n$. Let us see what comes out of this translation. Theorem 1 of [Fulton]'s §8.1 claims that the vectors e_U, where U ranges over all semistandard tableaux of shape λ, form a basis of the vector space E^λ (recall that our T is [Fulton]'s

U, and vice versa!). Consider a semistandard tableau U of shape λ, and the corresponding basis vector e_U. The isomorphism $E^\lambda \to E\left(S^\lambda\right)$ from Proposition 1 in [Fulton]'s §8.3 maps this basis vector $e_U \in E^\lambda$ to the tensor $\mathbf{v}(T) \otimes v_T$ (again, our T is [Fulton]'s U, so this looks like "$\mathbf{v}(U) \otimes v_U$" in [Fulton]), where $\mathbf{v}(T) \in E^{\otimes d}$ is the pure tensor that we call $e_{(T,U)}$ (defined in the exact same way), while v_T is the polytabloid in the Specht module S^λ corresponding to the standard tableau T. The isomorphism $S^\lambda \to \mathbf{C}[S_d] \cdot \mathbf{E}_T$ sends this latter polytabloid v_T to $\mathbf{E}_T$, and thus the induced isomorphism $E^{\otimes d} \otimes_{\mathbf{C}[S_d]} S^\lambda \to E^{\otimes d} \otimes_{\mathbf{C}[S_d]} (\mathbf{C}[S_d] \cdot \mathbf{E}_T)$ sends $\mathbf{v}(T) \otimes v_T$ to $\mathbf{v}(T) \otimes \mathbf{E}_T$. Finally, the isomorphism $E^{\otimes d} \otimes_{\mathbf{C}[S_d]} (\mathbf{C}[S_d] \cdot \mathbf{E}_T) \to E^{\otimes d}\mathbf{E}_T$ sends $\mathbf{v}(T) \otimes \mathbf{E}_T$ to $\mathbf{v}(T)\mathbf{E}_T = e_{(T,U)}\mathbf{E}_T$ (since $\mathbf{v}(T) = e_{(T,U)}$). Altogether, we thus see that our isomorphism $E^\lambda \to W_T\mathbf{C}^n$ constructed in (62) sends e_U to $e_{(T,U)}\mathbf{E}_T$ (with the intermediate steps being $e_U \mapsto \mathbf{v}(T) \otimes v_T \mapsto \mathbf{v}(T) \otimes \mathbf{E}_T \mapsto e_{(T,U)}\mathbf{E}_T$). Hence, it sends the whole basis $(e_U)_{U \text{ ssyt}\lambda}$ of E^λ to a basis $\left(e_{(T,U)}\mathbf{E}_T\right)_{U \text{ ssyt}\lambda}$ of $W_T\mathbf{C}^n$. We can rewrite this latter basis further as $\left(c_T\left(e_{(T,U)}\right)\right)_{U \text{ ssyt}\lambda}$ (since c_T is right multiplication by $\mathbf{E}_T$, so that we have $c_T\left(e_{(T,U)}\right) = e_{(T,U)}\mathbf{E}_T$ for each U ssytλ). Thus, we have shown that $\left(c_T\left(e_{(T,U)}\right)\right)_{U \text{ ssyt}\lambda}$ is a basis of $W_T\mathbf{C}^n$. This proves Corollary 4.1.14 **(1)**.

(2) Consider a diagonal matrix $\operatorname{diag}(t_1, t_2, \ldots, t_n) \in \Gamma$. We compute the trace $\operatorname{trace}(\rho(\operatorname{diag}(t_1, t_2, \ldots, t_n)))$ using the basis of $W_T\mathbf{C}^n$ given in Corollary 4.1.14 **(1)**. The map $\rho(\operatorname{diag}(t_1, t_2, \ldots, t_n))$ acts on a basis vector $c_T\left(e_{(T,U)}\right)$ by scaling it by the factor

$$\rho(\operatorname{diag}(t_1, t_2, \ldots, t_n)) \cdot c_T\left(e_{(T,U)}\right) = \prod_{i=1}^n t_i^{\#\ i\text{'s in } U}$$

(since $c_T\left(e_{(T,U)}\right)$ is a linear combination of permutations of the pure tensor $e_{(T,U)}$, and the latter pure tensor contains each e_i as many times as there are i's in U). Thus, $\operatorname{trace}(\rho(\operatorname{diag}(t_1, t_2, \ldots, t_n)))$ is the sum of these factors $\prod_{i=1}^n t_i^{\#\ i\text{'s in } U}$ over all U ssytλ. That is,

$$\operatorname{trace}(\rho(\operatorname{diag}(t_1, t_2, \ldots, t_n))) = \sum_{U \text{ ssyt}\lambda} \prod_{i=1}^n t_i^{\#\ i\text{'s in } U} = s_\lambda(t_1, t_2, \ldots, t_n).$$

This proves Corollary 4.1.14 **(2)**. ■

108. **page 144:** "Any partition $\lambda \vdash d$ can be encoded into a monomial $\omega(\lambda) := t_1^{\nu_1} t_2^{\nu_2} \cdots t_n^{\nu_n}$ as follows: the exponent ν_i is the cardinality of the i-th column in the Ferrers diagram of λ. Equivalently, $\nu_i = \#\left\{j : \lambda_j \geq i\right\}$. It is easy to

see that $\omega(\lambda)$ is the lexicographically leading monomial of the Schur polynomial $s_\lambda(t_1, t_2, \ldots, t_n)$" should be "Any partition $\lambda \vdash d$ can be encoded into a monomial $\omega(\lambda) := t_1^{\lambda_1} t_2^{\lambda_2} \cdots t_n^{\lambda_n}$ as follows. It is easy to see that $\omega(\lambda)$ is the lexicographically leading monomial of the Schur polynomial $s_\lambda(t_1, t_2, \ldots, t_n)$ (where the variables are ordered by $t_1 > t_2 > \cdots > t_n$)".

109. **page 144, (4.1.8):** The "$c_\lambda \ W_\lambda \mathbf{C}^n$" here means a direct sum of c_λ many copies of $W_\lambda \mathbf{C}^n$.

110. **page 144:** I am not sure I would call Algorithm 4.1.16 a "subduction algorithm (cf. Algorithm 3.2.8)"; the Schur polynomials form a $\mathbf{Z}$-module basis of $\mathbf{Z}[t_1, \ldots, t_n]^{S_n}$, not a sagbi basis.

111. **page 145, Algorithm 4.1.16:** In Step 2, replace "$t_1^{\nu_1} t_2^{\nu_2} \cdots t_n^{\nu_n}$" by "$t_1^{\lambda_1} t_2^{\lambda_2} \cdots t_n^{\lambda_n}$".

112. **page 145, Example 4.1.17:** "the space of polynomial functions" should be "the space of homogeneous polynomial functions".

113. **page 145, Example 4.1.17:** In (4.1.12), replace "$s_{(3,3,3)}$" by "$s_{(2,2,2)}$".

114. **page 146:** "the space of homogeneous polynomials of degree d in the coefficients of a generic homogeneous polynomials of degree m" should be "the space of homogeneous polynomials of degree m in the coefficients of a generic homogeneous polynomial of degree d".

115. **page 147, §4.2:** Replace "$A \circ f$" by "$f \circ A$" everywhere in the third paragraph of §4.2. (The group Γ acts on $\mathbf{C}[V]$ is a right action, not a left action.)

116. **page 147, §4.2:** "The symmetric power $S_k(V)$ is a vector space of dimension $\binom{m+k-1}{k}$. We identify it with the space of homogeneous polynomial functions of degree k on V": This identification is unnatural, and does not respect the Γ-action in general. The vector space $\mathbf{C}[V]_k^\Gamma$ of all homogeneous polynomial functions of degree k on V can be naturally identified with the symmetric power $S_k(V^*)$, where Γ acts on the dual space V^* of V by the transpose matrix (i.e., where the action of Γ on V^* is given by $(f \circ A)(v) = f(Av)$ for all $A \in \Gamma$, $f \in V^*$ and $v \in V$). On both of these spaces, Γ acts **from the right**. Meanwhile, on the symmetric power $S_k(V)$, the group Γ acts **from the left**. Even if we transform the left action into a right action by taking inverses (i.e., we define Af to be $f \circ A^{-1}$), and even if we assume that $V = \mathbf{C}^n$ (in which case there is a "natural" vector space isomorphism $\mathbf{C}[V]_k^\Gamma \cong S_k(V^*) \cong S_k(V)$ by way of the standard basis), the actions do not become the same (unless you restrict them to the orthogonal group $\mathrm{O}(\mathbf{C}^n)$).

This discrepancy prevents much of what is being done in Chapter 4 of the book from being true "on the nose". However, the approach to invariants by means of symmetric powers is nevertheless a valid one, and much of what is done in Chapter 4 can be adapted once proper fixes are made. In many cases, it suffices to read $S_k(V^*)$ for $S_k(V)$ and make sure to adapt the representation theory of Γ to right (as opposed to left) Γ-actions. Even without such an adaptation, the invariant-theoretical insights can often be salvaged, since it can be shown that for any $k \in \mathbf{N}$ and any representation (V, ρ) of Γ, we have

$$\dim\left((S_k(V^*))^\Gamma\right) = \dim\left((S_k(V))^\Gamma\right) \tag{63}$$

(despite the lack of a Γ-equivariant isomorphism $S_k(V^*) \cong S_k(V)$).

Let me outline a proof of (63). It relies on the following two facts:

> *Fact 1:* Let (W, τ) be any Γ-representation. Then, $\dim\left((W^*)^\Gamma\right) = \dim\left(W^\Gamma\right)$.
>
> *Fact 2:* Let (V, ρ) be any Γ-representation. Let $k \in \mathbf{N}$. Then, $S_k(V^*) \cong (S_k(V))^*$ as Γ-representations.

Proof of Fact 1 (sketched). Let's say that W is a left Γ-representation. Decompose W into a direct sum $I_1 \oplus I_2 \oplus \cdots \oplus I_k$ of irreducible left Γ-representations. Then, $\dim\left(W^\Gamma\right)$ is the number of $i \in \{1, 2, \ldots, k\}$ for which I_i is isomorphic to the trivial 1-dimensional representation $\mathbf{C}$. But $W \cong I_1 \oplus I_2 \oplus \cdots \oplus I_k$ as left Γ-representations yields

$$W^* \cong (I_1 \oplus I_2 \oplus \cdots \oplus I_k)^* \cong I_1^* \oplus I_2^* \oplus \cdots \oplus I_k^*$$

as right Γ-representations, and moreover, the addends I_j^* here are again irreducible (since the dual of an irreducible left Γ-representation is an irreducible right Γ-representation). Hence, $\dim\left((W^*)^\Gamma\right)$ is the number of $i \in \{1, 2, \ldots, k\}$ for which I_i^* is isomorphic to the trivial 1-dimensional representation $\mathbf{C}$.

But the $i \in \{1, 2, \ldots, k\}$ that satisfy $I_i \cong \mathbf{C}$ are precisely those $i \in \{1, 2, \ldots, k\}$ that satisfy $I_i^* \cong \mathbf{C}$ (since $\mathbf{C}^* \cong \mathbf{C}$). Thus, the above descriptions of $\dim\left((W^*)^\Gamma\right)$ and of $\dim\left(W^\Gamma\right)$ boil down to the same thing, and we conclude that $\dim\left((W^*)^\Gamma\right) = \dim\left(W^\Gamma\right)$. This proves Fact 1. □

Proof of Fact 2 (sketched). There is an obvious bilinear form

$$\begin{aligned}
&\beta : S_k(V^*) \times S_k(V) \to \mathbf{C}, \\
&(f_1 f_2 \cdots f_k,\ v_1 v_2 \cdots v_k) \mapsto \sum_{\sigma \in S_k} f_1\left(v_{\sigma(1)}\right) f_2\left(v_{\sigma(2)}\right) \cdots f_k\left(v_{\sigma(k)}\right) \\
&\qquad\qquad = \sum_{\sigma \in S_k} f_{\sigma(1)}(v_1) f_{\sigma(2)}(v_2) \cdots f_{\sigma(k)}(v_k),
\end{aligned}$$

which is easily seen to be Γ-tensorial (i.e., any $g \in S_k(V^*)$ and $w \in S_k(V)$ and $A \in \Gamma$ satisfy $\beta(g \circ A, w) = \beta(g, Aw)$) and nondegenerate (i.e., it produces an isomorphism $S_k(V^*) \to (S_k(V))^*$) [20]. Thus, it gives rise to an isomorphism $S_k(V^*) \cong (S_k(V))^*$ of right Γ-representations. This proves Fact 2. □

Now (63) follows from

$$\begin{aligned} \dim\left((S_k(V^*))^\Gamma\right) &= \dim\left(((S_k(V))^*)^\Gamma\right) && \text{(by Fact 2)} \\ &= \dim\left((S_k(V))^\Gamma\right) && \text{(by Fact 1, applied to } W = S_k(V)\text{)}. \end{aligned}$$

117. **page 147, §4.2:** After "This implies that f is a homogeneous polynomial of degree gn/d", I would add "(in fact, we can WLOG assume that f is homogeneous; then, the assumption $f \circ A = (\det A)^g \cdot f$ rewrites as $f(\rho(A) \cdot \mathbf{v}) = (\det A)^g \cdot f(\mathbf{v})$, but the left hand side of the latter equality is homogeneous of degree $d \cdot \deg f$ in the entries of A, while the right hand side is homogeneous of degree gn in the entries of A; thus, $d \cdot \deg f = gn$ and therefore $\deg f = gn/d$)".

118. **page 148, Proposition 4.2.1:** As remarked above, the polynomial ring $\mathbf{C}[V]$ is not isomorphic to the symmetric algebra $S(V)$ of V as a Γ-module. Consequently, "$\mathbf{C}[V]$" should be replaced by "$S(V)$" throughout this proposition and its proof.

119. **page 150, proof of Lemma 4.2.4:** The "typical such SSYTλ" looks not like

$$\begin{array}{ccccccccc} 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 1 & \cdots & 1 & 2 & \cdots & 2 & & & \end{array}$$

but like

$$\begin{array}{ccccccccc} 1 & \cdots & 1 & 1 & \cdots & 1 & 2 & \cdots & 2 \\ 2 & \cdots & 2 & & & & & & \end{array}.$$

120. **page 150, proof of Lemma 4.2.4:** "By Corollary 4.1.14 (b)" should be "By Corollary 4.1.14 (2)".

121. **page 157:** "which we denote with $\Omega_\mathbf{s}$, $\Omega_\mathbf{t}$ and $\Omega_\mathbf{u}$ respectively" should be "which we denote with $\Omega_\mathbf{t}$, $\Omega_\mathbf{s}$ and $\Omega_\mathbf{u}$ respectively".

122. **page 157:** "Given a matrix-valued polynomial function ϕ" should be "Given a polynomial function $\phi \in \mathbf{C}[\mathbf{t}]$".

123. **page 157, proof of Theorem 4.3.4:** Replace each "k" in this proof by an "n".

[20] For the nondegeneracy, we have to thank the facts that V is finite-dimensional and that our base field $\mathbf{C}$ has characteristic 0.

124. **page 157, proof of Theorem 4.3.4:** In the equality (4.3.5) and further on, the notation
$$\frac{\partial^n \phi\left(\mathbf{u}=\mathbf{st}\right)}{\partial u_{\sigma_1,\pi_1}\,\partial u_{\sigma_2,\pi_2}\,\cdots\,\partial u_{\sigma_n,\pi_n}}$$
should be understood as
$$\frac{\partial^n \phi\left(\mathbf{u}\right)}{\partial u_{\sigma_1,\pi_1}\,\partial u_{\sigma_2,\pi_2}\,\cdots\,\partial u_{\sigma_n,\pi_n}}\,|_{\mathbf{u}=\mathbf{st}}$$
(that is, we first compute the n-th partial derivative $\dfrac{\partial^n \phi\left(\mathbf{u}\right)}{\partial u_{\sigma_1,\pi_1}\,\partial u_{\sigma_2,\pi_2}\,\cdots\,\partial u_{\sigma_n,\pi_n}}$ in the ring $\mathbf{C}\left[\mathbf{u}\right]$, and then substitute the entries of the matrix $\mathbf{st}$ for the respective entries of $\mathbf{u}$ in the result we obtained). An alternative (and more standard) way to express the same thing is
$$\frac{\partial^n \phi\left(\mathbf{u}\right)}{\partial u_{\sigma_1,\pi_1}\,\partial u_{\sigma_2,\pi_2}\,\cdots\,\partial u_{\sigma_n,\pi_n}}\left(\mathbf{st}\right).$$

125. **page 157, proof of Theorem 4.3.4:** Here is a quick outline of how (4.3.5) can be proved. Namely, (4.3.5) follows by n-times repeated application of the formula
$$\frac{\partial\phi\left(\mathbf{st}\right)}{\partial t_{p,q}}=\sum_{\sigma=1}^{n}\left(\frac{\partial\phi\left(\mathbf{u}\right)}{\partial u_{\sigma,q}}\right)\left(\mathbf{st}\right)\cdot s_{\sigma,p}, \tag{64}$$
which holds for all $p,q\in\left\{1,2,\ldots,n\right\}$. This formula (64), in turn, is a particular case of the multivariate chain rule (since $\phi\left(\mathbf{st}\right)=\phi\left(\psi\left(\mathbf{t}\right)\right)$, where ψ is the function sending each entry $t_{i,j}$ of the matrix $\mathbf{t}$ to the respective entry $\sum\limits_{h=1}^{n}s_{i,h}t_{h,j}$ of $\mathbf{st}$). Alternatively, (64) can be proved by induction on $\deg\phi$, using the fact that both sides of (64) are functions $\omega\left(\phi\right)$ of ϕ that satisfy the recurrence
$$\omega\left(\phi\psi\right)=\phi\left(\mathbf{st}\right)\cdot\omega\left(\psi\right)+\omega\left(\phi\right)\cdot\omega\left(\mathbf{st}\right)$$
(by the Leibniz rule) and send each indeterminate $t_{i,j}$ to $\begin{cases}0, & \text{if } q\neq j;\\ s_{i,p}, & \text{if } q=j\end{cases}$ (this is easy to check directly).

126. **page 157, proof of Theorem 4.3.4:** "antisymmetrize the expression (4.3.5) with respect to $\pi\in S_n$" simply means "multiply (4.3.5) by $\operatorname{sign}\left(\pi\right)$ and sum the result over all $\pi\in S_n$". (No further substitution of variables is required, unlike in a usual antisymmetrization of polynomials.)

127. **page 158, proof of Theorem 4.3.4:** "The proof of the second identity in (4.3.4) is analogous": This needs some caveats. Instead of applying (4.3.3)

directly with each $t_{i,j}$ replaced by $s_{i,j}$, we have to apply the identity

$$\Omega(\phi) = \sum_{\pi \in S_n} \operatorname{sign}(\pi) \cdot \frac{\partial^n \phi}{\partial t_{\pi_1,1}\, \partial t_{\pi_2,2}\, \cdots\, \partial t_{\pi_n,n}}$$

(which follows from (4.3.3) by first reindexing the sum using the bijection $\pi \mapsto \pi^{-1}$ and then rewriting the expression $\dfrac{\partial^n \phi}{\partial t_{1,\pi_1^{-1}}\, \partial t_{2,\pi_2^{-1}}\, \cdots\, \partial t_{n,\pi_n^{-1}}}$ as $\dfrac{\partial^n \phi}{\partial t_{\pi_1,1}\, \partial t_{\pi_2,2}\, \cdots\, \partial t_{\pi_n,n}}$).

Instead of (64), we need the analogous formula

$$\frac{\partial \phi(\mathbf{st})}{\partial s_{p,q}} = \sum_{\sigma=1}^{n} \left(\frac{\partial \phi(\mathbf{u})}{\partial u_{p,\sigma}} \right) (\mathbf{st}) \cdot t_{q,\sigma}. \tag{65}$$

Instead of (4.3.5), we need the formula

$$\frac{\partial^n \phi(\mathbf{st})}{\partial s_{\pi_1,1}\, \partial s_{\pi_2,2}\, \cdots\, \partial s_{\pi_n,n}} = \sum_{\sigma_1,\sigma_2,\ldots,\sigma_n=1}^{n} \frac{\partial^n \phi(\mathbf{u} = \mathbf{st})}{\partial u_{\pi_1,\sigma_1}\, \partial u_{\pi_2,\sigma_2}\, \cdots\, \partial u_{\pi_n,\sigma_n}} t_{1,\sigma_1} t_{2,\sigma_2} \cdots t_{n,\sigma_n}.$$

128. **page 158, Corollary 4.3.6:** It can be shown that this constant c_p equals $p(p+1)(p+2)\cdots(p+n-1)$. Indeed, Corollary 4.3.6 with this particular formula for c_p is known as the *Cayley identity*, and appears (e.g.) as Corollary A.4 in Sergio Caracciolo, Andrea Sportiello, Alan D. Sokal, *Noncommutative determinants, Cauchy-Binet formulae, and Capelli-type identities. I. Generalizations of the Capelli and Turnbull identities*, arXiv:0809.3516v2 (see my errata).

129. **page 159:** "its image $\mathbf{t} \circ f = f(\mathbf{tv})$" should be "its image $f \circ \mathbf{t} = f(\mathbf{tv})$".

130. **page 163, proof of Proposition 4.4.2:** "as polynomial functions" should be "as rational functions".

131. **page 163, proof of Proposition 4.4.2:** In (4.4.5), the "$\partial x_i\, \partial x_j$" on the right hand side should be "$\partial x_k\, \partial x_l$".

132. **page 163, definition of the discriminant of a quadratic form:** You probably want to replace "$\sum_{i=1}^{n} \sum_{j=i}^{n} a_{ij} x_i x_j$" by "$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$ (with $a_{ij} = a_{ji}$)" in order to recover the factors 2 in (4.4.7).

133. **page 164, Theorem 4.4.3:** "ring a quadratic" should be "ring of a quadratic".

134. **page 169, proof of Lemma 4.5.1:** The "if" direction of the lemma can also be proved without any reference to connectedness and Lie groups. Here is an outline of this proof:

Proof of the "if" direction of Lemma 4.5.1. Let $v \in V$ be a vector. Assume that

$$\rho^* \left(E_{i,j}\right) \cdot v = 0 \qquad \text{for all } i, j \in \{1, 2, \ldots, n\} \text{ with } i \neq j,$$

and that

$$\rho^* \left(E_{i,i}\right) \cdot v = g \cdot v \qquad \text{for all } i \in \{1, 2, \ldots, n\}.$$

We can combine these two assumptions into one: namely, for all $i, j \in \{1, 2, \ldots, n\}$, we have

$$\rho^* \left(E_{i,j}\right) \cdot v = \delta_{i,j} g \cdot v, \tag{66}$$

where $\delta_{i,j}$ is the Kronecker delta of i and j. (This covers both the case $i = j$ and the case $i \neq j$.)

Let us set

$$\gamma(T) := (\det T)^{-g} \cdot \rho(T) \in \operatorname{GL}(V) \qquad \text{for each } T \in \Gamma.$$

Thus, γ is a rational map from Γ to $\operatorname{GL}(V)$ (that is, if we choose a basis of V and coordinatize $\operatorname{GL}(V) \subseteq \operatorname{End}(V)$ in the obvious way using this basis, then all the coordinates of $\gamma(T)$ are rational functions in the coordinates of T). By the definition of γ, we have

$$\gamma(\mathbf{1}) = \underbrace{(\det \mathbf{1})^{-g}}_{=1^{-g}=1} \cdot \underbrace{\rho(\mathbf{1})}_{=\operatorname{id}_V} = \operatorname{id}_V.$$

Moreover, for each $T \in \Gamma$, we have

$$\rho(T) = (\det T)^{g} \cdot \gamma(T) \tag{67}$$

(since $\gamma(T) = (\det T)^{-g} \cdot \rho(T)$). Thus, for each $i, j \in \{1, 2, \ldots, n\}$, we have

$$\begin{aligned}
&\left(\frac{\partial}{\partial t_{i,j}} \rho(T)\right)_{T=\mathbf{1}} \\
&= \left(\frac{\partial}{\partial t_{i,j}} \left((\det T)^{g} \cdot \gamma(T)\right)\right)_{T=\mathbf{1}} \\
&= \underbrace{\left(\frac{\partial}{\partial t_{i,j}} (\det T)^{g}\right)_{T=\mathbf{1}}}_{\substack{=\delta_{i,j} g \\ \text{(this is easy to check)}}} \cdot \underbrace{\gamma(\mathbf{1})}_{=\operatorname{id}_V} + \underbrace{(\det \mathbf{1})^{g}}_{=1^{g}=1} \cdot \left(\frac{\partial}{\partial t_{i,j}} \gamma(T)\right)_{T=\mathbf{1}} \\
&\qquad \text{(by the Leibniz rule for noncommutative products)} \\
&= \delta_{i,j} g \operatorname{id}_V + \left(\frac{\partial}{\partial t_{i,j}} \gamma(T)\right)_{T=\mathbf{1}},
\end{aligned}$$

so that

$$\left(\frac{\partial}{\partial t_{i,j}}\gamma\left(T\right)\right)_{T=\mathbf{1}} = \underbrace{\left(\frac{\partial}{\partial t_{i,j}}\rho\left(T\right)\right)_{T=\mathbf{1}}}_{\substack{=\rho^*\left(E_{i,j}\right) \\ \text{(by (4.5.2))}}} - \delta_{i,j}g\operatorname{id}_V = \rho^*\left(E_{i,j}\right) - \delta_{i,j}g\operatorname{id}_V$$

and therefore

$$\begin{aligned}\left(\frac{\partial}{\partial t_{i,j}}\gamma\left(T\right)\right)_{T=\mathbf{1}} \cdot v &= \left(\rho^*\left(E_{i,j}\right) - \delta_{i,j}g\operatorname{id}_V\right) \cdot v \\ &= \rho^*\left(E_{i,j}\right) \cdot v - \delta_{i,j}g \cdot v = 0 \qquad \text{(by (66))}.\end{aligned}$$

Thus, we have shown that

$$\left(\frac{\partial}{\partial t_{i,j}}\gamma\left(T\right)\right)_{T=\mathbf{1}} \cdot v = 0 \qquad \text{for each } i,j \in \{1,2,\ldots,n\}. \tag{68}$$

Now, fix $S \in \Gamma$. Let $i,j \in \{1,2,\ldots,n\}$. For each $T \in \Gamma$, we have

$$\rho\left(ST\right) = \rho\left(S\right) \cdot \rho\left(T\right)$$

(since ρ is a group action) and thus

$$\begin{aligned}\gamma\left(ST\right) &= \left(\underbrace{\det\left(ST\right)}_{=\det S \cdot \det T}\right)^{-g} \cdot \underbrace{\rho\left(ST\right)}_{=\rho(S)\cdot\rho(T)} \qquad \text{(by the definition of } \gamma\text{)} \\ &= \underbrace{\left(\det S \cdot \det T\right)^{-g}}_{=(\det S)^{-g}\cdot(\det T)^{-g}} \cdot \rho\left(S\right) \cdot \rho\left(T\right) \\ &= \underbrace{\left(\det S\right)^{-g} \cdot \rho\left(S\right)}_{\substack{=\gamma(S) \\ \text{(by the definition of } \gamma\text{)}}} \cdot \underbrace{\left(\det T\right)^{-g} \cdot \rho\left(T\right)}_{\substack{=\gamma(T) \\ \text{(by the definition of } \gamma\text{)}}} \\ &= \gamma\left(S\right) \cdot \gamma\left(T\right)\end{aligned}$$

and therefore

$$\gamma\left(ST\right) \cdot v = \gamma\left(S\right) \cdot \gamma\left(T\right) \cdot v.$$

Applying the operator $\frac{\partial}{\partial t_{i,j}} |_{T=\mathbf{1}}$ to this equality (i.e., taking the derivative with respect to $t_{i,j}$ at the point $T = \mathbf{1}$), we obtain

$$\begin{aligned}\frac{\partial}{\partial t_{i,j}}\left(\gamma\left(ST\right) \cdot v\right)_{T=\mathbf{1}} &= \frac{\partial}{\partial t_{i,j}}\left(\gamma\left(S\right) \cdot \gamma\left(T\right) \cdot v\right)_{T=\mathbf{1}} = \gamma\left(S\right) \cdot \underbrace{\left(\frac{\partial}{\partial t_{i,j}}\gamma\left(T\right)\right)_{T=\mathbf{1}} \cdot v}_{\substack{=0 \\ \text{(by (68))}}} \\ &= 0.\end{aligned} \tag{69}$$

On the other hand, the map $\Gamma \to V,\ T \mapsto \gamma(ST) \cdot v$ can be written as the composition of the maps $\Gamma \to \Gamma,\ T \mapsto ST$ and $\Gamma \to V,\ U \mapsto \gamma(U) \cdot v$. Thus, the chain rule for multivariate functions yields

$$\begin{aligned}
&\frac{\partial}{\partial t_{i,j}} \left(\gamma(ST) \cdot v\right)_{T=\mathbf{1}} \\
&= \sum_{k,l=1}^{n} \underbrace{\left(\frac{\partial}{\partial t_{i,j}} \left((ST)_{k,l}\right)\right)_{T=\mathbf{1}}}_{\substack{=\begin{cases} s_{k,i}, & \text{if } l=j; \\ 0, & \text{if } l \neq j \end{cases} \\ \left(\text{since } (ST)_{k,l} = \sum_{h=1}^{n} s_{k,h} t_{h,l}\right)}} \cdot \underbrace{\left(\frac{\partial}{\partial u_{k,l}} \left(\gamma(U) \cdot v\right)\right)_{U=S\mathbf{1}}}_{\substack{=\left(\frac{\partial}{\partial u_{k,l}}\gamma(U)\right)_{U=S\mathbf{1}} \cdot v \\ =\left(\frac{\partial}{\partial u_{k,l}}\gamma(U)\right)_{U=S} \cdot v \\ (\text{since } S\mathbf{1}=S)}} \\
&\qquad \left(\begin{array}{c}\text{where } u_{k,l} \text{ denotes the } (k,l)\text{-th entry of the matrix } U \in \Gamma, \\ \text{and where } (ST)_{k,l} \text{ denotes the } (k,l)\text{-th entry of the matrix } ST\end{array}\right) \\
&= \sum_{k,l=1}^{n} \begin{cases} s_{k,i}, & \text{if } l=j; \\ 0, & \text{if } l \neq j \end{cases} \cdot \left(\frac{\partial}{\partial u_{k,l}}\gamma(U)\right)_{U=S} \cdot v \\
&= \sum_{k=1}^{n} s_{k,i} \cdot \left(\frac{\partial}{\partial u_{k,j}}\gamma(U)\right)_{U=S} \cdot v \\
&\qquad \left(\begin{array}{c}\text{here, we have removed all addends with } l \neq j \\ \text{from the sum, since these addends are } 0\end{array}\right).
\end{aligned}$$

Comparing this with (69), we obtain

$$\sum_{k=1}^{n} s_{k,i} \cdot \left(\frac{\partial}{\partial u_{k,j}}\gamma(U)\right)_{U=S} \cdot v = 0. \tag{70}$$

Now forget that we fixed i. We thus have proved the equality (70) for all $i \in \{1, 2, \ldots, n\}$.

But the matrix S is invertible (since $S \in \Gamma = \mathrm{GL}(\mathbf{C}^n)$). Let $S^{-1} = \left(s'_{p,q}\right)_{p,q \in \{1,2,\ldots,n\}}$ be its inverse matrix. Then, $SS^{-1} = \mathbf{1}$, so that all $k, q \in \{1, 2, \ldots, n\}$ satisfy

$$\sum_{i=1}^{n} s_{k,i} s'_{i,q} = \delta_{k,q} \tag{71}$$

(by comparing the (k,q)-th entries of the matrices SS^{-1} and $\mathbf{1}$). Now, for any $q \in \{1, 2, \ldots, n\}$, we have

$$\sum_{i=1}^{n} s'_{i,q} \cdot \underbrace{\sum_{k=1}^{n} s_{k,i} \cdot \left(\frac{\partial}{\partial u_{k,j}}\gamma(U)\right)_{U=S} \cdot v}_{\substack{=0 \\ (\text{by } (70))}} = 0$$

and thus

$$\begin{aligned} 0 &= \sum_{i=1}^{n} s'_{i,q} \cdot \sum_{k=1}^{n} s_{k,i} \cdot \left(\frac{\partial}{\partial u_{k,j}} \gamma\left(U\right) \right)_{U=S} \cdot v \\ &= \sum_{k=1}^{n} \underbrace{\sum_{i=1}^{n} s_{k,i} s'_{i,q}}_{\substack{=\delta_{k,q} \\ \text{(by (71))}}} \cdot \underbrace{\left(\frac{\partial}{\partial u_{k,j}} \gamma\left(U\right) \right)_{U=S} \cdot v}_{= \left(\frac{\partial}{\partial u_{k,j}} (\gamma(U) \cdot v) \right)_{U=S}} \\ &= \sum_{k=1}^{n} \delta_{k,q} \cdot \left(\frac{\partial}{\partial u_{k,j}} \left(\gamma\left(U\right) \cdot v\right) \right)_{U=S} = \left(\frac{\partial}{\partial u_{q,j}} \left(\gamma\left(U\right) \cdot v\right) \right)_{U=S} \end{aligned}$$

(since all the addends of the sum are 0 except for the addend for $k = q$). In other words,

$$\left(\frac{\partial}{\partial u_{q,j}} \left(\gamma\left(U\right) \cdot v\right) \right)_{U=S} = 0 \qquad \text{for all } q \in \{1, 2, \ldots, n\}.$$

Now, forget that we fixed j. We thus have shown that

$$\left(\frac{\partial}{\partial u_{q,j}} \left(\gamma\left(U\right) \cdot v\right) \right)_{U=S} = 0 \qquad \text{for all } q, j \in \{1, 2, \ldots, n\}.$$

In other words, all the partial derivatives of $\gamma(U) \cdot v$ (as a function in $U \in \Gamma$) at the point $U = S$ are 0. Since $S \in \Gamma$ was also chosen arbitrary, this shows that all the partial derivatives of $\gamma(U) \cdot v$ (as a function in $U \in \Gamma$) at every point are 0. Therefore, this function $\gamma(U) \cdot v$ must be constant (since a rational function whose all partial derivatives at every point are 0 must be constant[21]). In particular, every $T \in \Gamma$ thus satisfies

$$\gamma\left(T\right) \cdot v = \underbrace{\gamma\left(\mathbf{1}\right)}_{=\mathrm{id}_V} \cdot v = v$$

and therefore

$$\begin{aligned} \rho\left(T\right) \cdot v &= \left(\det T\right)^{g} \cdot \underbrace{\gamma\left(T\right) \cdot v}_{=v} \qquad \text{(by (67))} \\ &= \left(\det T\right)^{g} \cdot v. \end{aligned}$$

[21]For those who want to prove this without the use of analysis: First, reduce the result to the case of a univariate rational function with values in **C** (by focussing on a single coordinate of the output and on a change in a single coordinate of the input); then express this rational function as a ratio f/g of two coprime polynomials f and g, and argue that $(f/g)' = 0$ entails $f'g = fg'$, which can be combined with the coprimality of f and g to yield $f \mid f'$ and $g \mid g'$, which is absurd unless f and g are constant.

In other words, v is a Γ-invariant of index g. This proves the "if" direction of Lemma 4.5.1.

135. **page 170, Theorem 4.5.2:** In the first equality in the display (4.5.7), replace "for $i = j, 2, \ldots, n$" by "for $j = 1, 2, \ldots, n$".

136. **page 174, proof of Lemma 4.5.4:** On the right hand side of the display, replace "$x_{\sigma(i),j}^{v_{\sigma(i),j}}$" by "$x_{\sigma(i),j}^{v_{i,j}}$" (otherwise, we don't get an S_m-invariant sum).

137. **page 185, Lemma 4.7.2:** Remark: If $t \geq 1$, then the bound "$\leq s\left(t+1\right)^{s-1}$" can be replaced by the (slightly less horrible) bound "$\leq s\left(t^{s-1}-1\right)$" (since $\dfrac{r+k}{rt+k} \geq \dfrac{1}{t}$ for each $k \geq 0$).