

Math 701 Fall 2026, Lecture 8: Proof of the Littlewood–Richardson rule

website: <https://www.cip.ifi.lmu.de/~grinberg/t/26fs/>

(GPT-6 was used to polish and detail some of the arguments below. **Unchecked GPT output included currently; to be fixed.**)

Recall:

Definition 1.7.8. Let λ/μ be a skew partition, and let ν be an N -partition. A tableau $T \in \text{SSYT}(\lambda/\mu, N)$ is said to be ν -**Yamanouchi** if it has the property that for each positive integer j , the N -tuple $\nu + \text{cont}(\text{col}_{\geq j} T)$ is an N -partition (i.e., weakly decreasing).

We want to prove:

Theorem 1.7.9 (Zelevinsky’s generalized Littlewood–Richardson rule, in Yamanouchi form for finite N).

Let λ/μ be a skew partition and let ν be an N -partition. Then,

$$s_\nu \cdot s_{\lambda/\mu} = \sum_{\substack{T \in \text{SSYT}(\lambda/\mu, N); \\ T \text{ is } \nu\text{-Yamanouchi}}} s_{\nu + \text{cont } T}.$$

Instead of proving this theorem, we will first show Stembridge’s lemma:

Lemma 1.7.10 (Stembridge’s lemma). Let λ/μ be a skew partition and let ν be an N -partition. Then,

$$a_{\nu+\rho} \cdot s_{\lambda/\mu} = \sum_{\substack{T \in \text{SSYT}(\lambda/\mu, N); \\ T \text{ is } \nu\text{-Yamanouchi}}} a_{\nu + \text{cont } T + \rho}.$$

But before we show this, let me explain why this lemma implies both the LR (Littlewood–Richardson) rule and the $s_\lambda \cdot a_\rho = a_{\lambda+\rho}$ theorem (for $\ell(\lambda) \leq N$) that we have long been trying to prove.

To get the latter, we just specialize Stembridge’s lemma:

Proof of the bialternant identity using Stembridge’s lemma. Let λ be a partition of length $\leq N$, identified with an N -partition by padding it with zeroes. Set $\mu = \emptyset$ and $\nu = \emptyset = (0, 0, \dots, 0)$ in Lemma 1.7.10. We obtain

$$a_\rho \cdot s_\lambda = \sum_{\substack{T \in \text{SSYT}(\lambda, N); \\ T \text{ is } \emptyset\text{-Yamanouchi}}} a_{\text{cont } T + \rho}.$$

But what must a straight tableau $T \in \text{SSYT}(\lambda, N)$ satisfy in order to be $\emptyset$ -Yamanouchi? We can argue column-by-column (starting with the rightmost column) that all entries in each column must be $1, 2, \dots, h$, where h is the length of the column. Here are the details. Suppose that all columns strictly to the right of column j have already been shown to have this form. Let r be the length of column $j+1$ (or 0 if column j is the rightmost column), and let h be the length of column j . Since the shape is straight, $r \leq h$. For each $i \leq r$, strict increase down column j gives $T(i, j) \geq i$, while weak increase along row i gives $T(i, j) \leq T(i, j+1) = i$. Thus, the first r entries in column j are $1, 2, \dots, r$.

All entries in the columns to the right are $\leq r$. Therefore, the remaining entries in column j must be $r+1, r+2, \dots, h$: otherwise the content of $\text{col}_{\geq j} T$ would have a zero coordinate with index $> r$ followed by a nonzero coordinate. This would contradict its being weakly decreasing. This argument also covers the rightmost column, where $r = 0$, and proves the claim by induction. If λ is empty, there are no columns to check.

So T must be the tableau of shape λ whose entries $T(i, j)$ are simply given by $T(i, j) = i$. This is the tableau I call the **minimalistic tableau** of shape λ ; others call it the **highest-weight tableau**, since its content is lexicographically larger than the content of any other SSYT of shape λ with entries in $[N]$ (we will prove this later on). Conversely, this minimalistic tableau really is $\emptyset$ -Yamanouchi: the i -th entry of $\text{cont}(\text{col}_{\geq j} T)$ is $\max\{\lambda_i - j + 1, 0\}$, and these entries are weakly decreasing in i . Its entries belong to $[N]$ since $\ell(\lambda) \leq N$, and its content is $\text{cont } T = \lambda$. Hence,

$$\sum_{\substack{T \in \text{SSYT}(\lambda, N); \\ T \text{ is } \emptyset\text{-Yamanouchi}}} a_{\text{cont } T + \rho} = a_{\lambda + \rho},$$

and thus our above equality shows that $a_\rho \cdot s_\lambda = a_{\lambda + \rho}$, provided that Stembridge's lemma holds. $\square$

Proof of the LR rule. We have assumed that Stembridge's lemma holds, i.e., we have

$$a_{\nu + \rho} \cdot s_{\lambda/\mu} = \sum_{\substack{T \in \text{SSYT}(\lambda/\mu, N); \\ T \text{ is } \nu\text{-Yamanouchi}}} a_{\nu + \text{cont } T + \rho}.$$

By what we just proved, $a_{\nu + \rho} = a_\rho \cdot s_\nu$ and $a_{\nu + \text{cont } T + \rho} = a_\rho \cdot s_{\nu + \text{cont } T}$ (the latter applies because $\nu + \text{cont } T$ is an N -partition whenever T is ν -Yamanouchi). Thus, this equality rewrites as

$$a_\rho \cdot s_\nu \cdot s_{\lambda/\mu} = \sum_{\substack{T \in \text{SSYT}(\lambda/\mu, N); \\ T \text{ is } \nu\text{-Yamanouchi}}} a_\rho \cdot s_{\nu + \text{cont } T}.$$

Cancelling a_ρ would give us precisely the LR rule.

Why can we cancel a_ρ ? Well, if K is an integral domain, then the polynomial ring $\mathcal{P} = K[x_1, x_2, \dots, x_N]$ is an integral domain as well, and the polynomial

$a_\rho = \prod_{1 \leq i < j \leq N} (x_i - x_j)$ (Vandermonde determinant) is nonzero, so it can be cancelled. In the general case, we can argue as follows:

- Either argue that the LR rule is true for $K = \mathbb{Z}$ (which is an integral domain) and therefore for arbitrary K (since there is a functorial morphism $\mathcal{P}_{\mathbb{Z}} \rightarrow \mathcal{P}_K$ that sends skew Schur polynomials in $\mathcal{P}_{\mathbb{Z}}$ to skew Schur polynomials in $\mathcal{P}_K$).
- Or argue that multiplication by $a_\rho = \prod_{1 \leq i < j \leq N} (x_i - x_j)$ is injective for any K . Indeed, each factor $x_i - x_j$ is monic as a polynomial in x_i over the ring of polynomials in the other variables. Multiplication by a monic polynomial is injective even over a ring with zero divisors: the leading nonzero coefficient of a nonzero polynomial stays nonzero in the product. Multiplication by the product of these factors is therefore injective as well.

Either way, the LR rule is proved provided that we can prove Stembridge's lemma. $\square$

So it remains to prove Stembridge's lemma. We prepare with a simple observation about alternants:

Lemma 1.7.11. Let $\alpha \in \mathbb{N}^N$.

- (a) If α has two equal entries, then $a_\alpha = 0$.
- (b) If β is an N -tuple obtained from α by swapping two entries, then $a_\beta = -a_\alpha$.

Proof. Recall that $a_\alpha = \det(x_i^{\alpha_j})$ and use the fact that the determinant of a matrix is alternating in its columns. $\square$

Beginning of the proof of Lemma 1.7.10. This proof is due to Stembridge (2002). Gasharov gave a less elementary proof in 1998 based on some of the same ideas.

We will use the notation β_i for the i -th entry of any tuple β .

Keep in mind that $\sigma \cdot (fg) = (\sigma \cdot f) \cdot (\sigma \cdot g)$ for any $\sigma \in S_N$ and $f, g \in \mathcal{P}$.

For any $\beta \in \mathbb{N}^N$, we have

$$\begin{aligned}
 a_\beta &= \det(x_i^{\beta_j}) = \det(x_j^{\beta_i}) \quad \left(\text{since } \det(A^T) = \det A \right) \\
 &= \sum_{\sigma \in S_N} \underbrace{(-1)^\sigma}_{\text{the sign of } \sigma} \underbrace{x_{\sigma(1)}^{\beta_1} x_{\sigma(2)}^{\beta_2} \cdots x_{\sigma(N)}^{\beta_N}}_{\substack{= \sigma \cdot (x_1^{\beta_1} x_2^{\beta_2} \cdots x_N^{\beta_N}) \\ = \sigma \cdot x^\beta}} \\
 &= \sum_{\sigma \in S_N} (-1)^\sigma \sigma \cdot x^\beta.
 \end{aligned}$$

So

$$a_{\nu+\rho} = \sum_{\sigma \in S_N} (-1)^\sigma \sigma \cdot x^{\nu+\rho}.$$

Multiplying this by $s_{\lambda/\mu}$, we find

$$\begin{aligned} a_{\nu+\rho} \cdot s_{\lambda/\mu} &= \sum_{\sigma \in S_N} (-1)^\sigma (\sigma \cdot x^{\nu+\rho}) \cdot \underbrace{s_{\lambda/\mu}}_{=\sigma \cdot s_{\lambda/\mu}} \\ &\quad \text{(since } s_{\lambda/\mu} \text{ is symmetric)} \\ &= \sum_{\sigma \in S_N} (-1)^\sigma \underbrace{(\sigma \cdot x^{\nu+\rho}) \cdot (\sigma \cdot s_{\lambda/\mu})}_{=\sigma \cdot (x^{\nu+\rho} s_{\lambda/\mu})} \\ &= \sum_{\sigma \in S_N} (-1)^\sigma \sigma \cdot (x^{\nu+\rho} s_{\lambda/\mu}). \end{aligned}$$

Since

$$x^{\nu+\rho} s_{\lambda/\mu} = x^{\nu+\rho} \sum_{T \in \text{SSYT}(\lambda/\mu, N)} x^{\text{cont } T} = \sum_{T \in \text{SSYT}(\lambda/\mu, N)} x^{\nu+\text{cont } T+\rho},$$

we can rewrite this as

$$\begin{aligned} a_{\nu+\rho} \cdot s_{\lambda/\mu} &= \sum_{\sigma \in S_N} (-1)^\sigma \sigma \cdot \sum_{T \in \text{SSYT}(\lambda/\mu, N)} x^{\nu+\text{cont } T+\rho} \\ &= \sum_{T \in \text{SSYT}(\lambda/\mu, N)} \underbrace{\sum_{\sigma \in S_N} (-1)^\sigma \sigma \cdot x^{\nu+\text{cont } T+\rho}}_{\substack{=a_{\nu+\text{cont } T+\rho} \\ \text{(by the above } a_\beta = \sum_{\sigma \in S_N} (-1)^\sigma \sigma \cdot x^\beta \\ \text{formula, applied backwards)}}} \\ &= \sum_{T \in \text{SSYT}(\lambda/\mu, N)} a_{\nu+\text{cont } T+\rho}. \end{aligned}$$

This is **almost** the sum we want; we want

$$\sum_{\substack{T \in \text{SSYT}(\lambda/\mu, N); \\ T \text{ is } \nu\text{-Yamanouchi}}} a_{\nu+\text{cont } T+\rho}.$$

All that remains is to prove that these two sums are equal, i.e., that the non- ν -Yamanouchi tableaux in the above sum cancel out. To be precise, we must show that

$$\sum_{\substack{T \in \text{SSYT}(\lambda/\mu, N); \\ T \text{ is not } \nu\text{-Yamanouchi}}} a_{\nu+\text{cont } T+\rho} = 0.$$

But we will achieve this by cancellation and throwing out zero addends.

Let $\mathcal{X}$ be the set of all $T \in \text{SSYT}(\lambda/\mu, N)$ that are not ν -Yamanouchi. We must find an involution

$$f : \mathcal{X} \rightarrow \mathcal{X}$$

with the properties that

1. if $f(T) = T$, then $a_{\nu + \text{cont } T + \rho} = 0$;
2. for all $T \in \mathcal{X}$, we have $a_{\nu + \text{cont } T + \rho} + a_{\nu + \text{cont}(f(T)) + \rho} = 0$.

If such an involution f is found, then the above sum will break up into addends that are simply 0 (the ones corresponding to T that satisfy $f(T) = T$) and pairs of addends that together sum up to 0 (the ones corresponding to all the other T 's and their respective $f(T)$'s).

To define our involution f , let us analyze the non- ν -Yamanouchi tableaux $T \in \mathcal{X}$ more closely. Let $T \in \mathcal{X}$ be such a tableau. Then, there exists at least one $j \geq 1$ such that $\nu + \text{cont}(\text{col}_{\geq j} T)$ is **not** a partition. Any such j will be called a **violator** of T . Note that T has only finitely many violators (since $\text{col}_{\geq j} T$ is empty for all sufficiently large j , so its content is the zero N -tuple and $\nu + \text{cont}(\text{col}_{\geq j} T) = \nu$, which is a partition).

Let j be the **largest** violator of T . Set

$$\begin{aligned} b &:= \nu + \text{cont}(\text{col}_{\geq j} T) && \text{(not a partition)} && \text{and} \\ c &:= \nu + \text{cont}(\text{col}_{\geq j+1} T) && \text{(a partition, since } j \text{ is largest)}. \end{aligned}$$

Since b is not a partition, there exists a $k \in [N-1]$ such that $b_k < b_{k+1}$. Call such a k a **misstep** of T . Let k be the **smallest** misstep of T . Thus,

$$\begin{aligned} b_1 &\geq b_2 \geq \cdots \geq b_k < b_{k+1} && \text{whereas} \\ c_1 &\geq c_2 \geq \cdots \geq c_k \geq c_{k+1} \geq \cdots \geq c_N. \end{aligned}$$

We want to define $f(T)$ to be again a non- ν -Yamanouchi SSYT of shape λ/μ with the same largest violator j and the same smallest misstep k , but we want it to satisfy $a_{\nu + \text{cont } T + \rho} + a_{\nu + \text{cont}(f(T)) + \rho} = 0$. The simplest way to achieve the latter equality is to have $\nu + \text{cont } T + \rho$ and $\nu + \text{cont}(f(T)) + \rho$ differ by a switch of two entries – let's even make it two consecutive entries.

It stands to reason that $f(T)$ should agree with T in all columns $j, j+1, j+2, \dots$. Indeed, if this is the case, then it is automatic that $f(T)$ still has largest violator j and smallest misstep k (since nothing has changed that would affect these conditions). So let us try to obtain $f(T)$ from T by only changing the first $j-1$ columns. We consider the tableau $\text{col}_{< j} T$ obtained from T by removing all the columns $j, j+1, j+2, \dots$, without moving the remaining cells. This is again an SSYT of skew shape: its outer and inner partitions have entries $\min(\lambda_i, j-1)$ and $\min(\mu_i, j-1)$, respectively. If $j=1$, this is the empty tableau.

Before we define $f(T)$, let us make some observations. The N -tuples c and b have the property that for each $i \in [N]$, we have

$$c_i \leq b_i \leq c_i + 1$$

(since b differs from c only in that it includes the j -th column of T , but this column can contain at most 1 copy of i). So in particular,

$$c_k \leq b_k \leq c_k + 1 \quad \text{and} \quad c_{k+1} \leq b_{k+1} \leq c_{k+1} + 1.$$

But

$$c_k \geq c_{k+1} \quad \text{and} \quad b_k < b_{k+1}.$$

These six inequalities (actually four of them suffice) give the chain

$$c_k \leq b_k < b_{k+1} \leq c_{k+1} + 1 \leq c_k + 1.$$

Since all entries are integers, this entails that

$$c_k = c_{k+1}, \quad b_k = c_k, \quad \text{and} \quad b_{k+1} = c_{k+1} + 1.$$

The latter two equalities show that the j -th column of T has no k and one $k + 1$. In particular, $b_{k+1} = b_k + 1$.

Now, we define $f(T)$ by keeping columns $j, j + 1, j + 2, \dots$ of T unchanged while replacing $\text{col}_{<j} T$ by $\beta_k(\text{col}_{<j} T)$, where β_k is the Bender–Knuth involution. We must prove four claims:

1. The resulting tableau $f(T)$ is still semistandard.
2. We have $a_{\nu + \text{cont } T + \rho} + a_{\nu + \text{cont}(f(T)) + \rho} = 0$.
3. If $f(T) = T$, then $a_{\nu + \text{cont } T + \rho} = 0$.
4. This is an involution, i.e., we have $f(f(T)) = T$.

Proof of Claim 1. If $j = 1$, no entries change, so the claim is immediate. Thus, assume $j > 1$. If $f(T)$ were not semistandard, then we would have $(f(T))(i, j - 1) > (f(T))(i, j)$ for some i for which both cells $(i, j - 1)$ and (i, j) belong to the skew diagram (because all the other inequalities are guaranteed either by the semistandardness of T or by the semistandardness of $\beta_k(\text{col}_{<j} T)$). Since $T(i, j - 1) \leq T(i, j)$ and $(f(T))(i, j) = T(i, j)$, this can only happen if the $(i, j - 1)$ -th entry has increased from T to $f(T)$. Since this entry is affected only by β_k applied to $\text{col}_{<j} T$, this means that it would have increased from k to $k + 1$. But the j -th column of T has no k , so $T(i, j) \neq k$. Therefore, the inequality $T(i, j - 1) \leq T(i, j)$ could not have become false as $T(i, j - 1)$ increased from k to $k + 1$. Thus, $f(T)$ is semistandard. Since its columns $j, j + 1, \dots$ are unchanged, its largest violator is still j and its smallest misstep is still k . In particular, $f(T) \in \mathcal{X}$, so the map $f : \mathcal{X} \rightarrow \mathcal{X}$ is well-defined.

I will finish the other three claims next time.