HOPF ALGEBRAS IN COMBINATORICS

DARIJ GRINBERG AND VICTOR REINER

Abstract. Certain Hopf algebras arise in combinatorics because they have bases naturally parametrized by combinatorial objects (partitions, compositions, permutations, tableaux, graphs, trees, posets, polytopes, etc). The rigidity in the structure of a Hopf algebra can lead to enlightening proofs, and many interesting invariants of combinatorial objects turn out to be evaluations of Hopf morphisms.

These are lecture notes for Fall 2012 Math 8680 Topics in Combinatorics at the University of Minnesota taught by the second author. The course was an attempt to focus on examples that we find interesting, but which are hard to find fully explained currently in books or in one paper. Be warned that these notes are highly idiosyncratic in choice of topics, and they steal heavily from the sources in the bibliography.

Contents

Date: April 14, 2014.
Key words and phrases. Hopf algebra, combinatorics.

1. What is a Hopf algebra?

The standard references are Abe [1] and Sweedler [76], and some other good ones are [15, 19, 41, 56]. A reference which we discovered late, having a great deal of overlap with these notes is Hazewinkel, Gubareni, and Kirichenko [29].

Let's build up the definition of Hopf algebra structure bit-by-bit, starting with the more familiar definition of algebras.

Warnings: Unless otherwise specified ...

- $\mathbf{k}$ here usually denotes a field, but sometimes we'll want to take $\mathbf{k} = \mathbb{Z}$,
- all maps between $\mathbf{k}$-modules are $\mathbf{k}$-linear,
- all tensor products are over $\mathbf{k}$, and
- 1 will denote the multiplicative identity in some ring like $\mathbf{k}$ or in some $\mathbf{k}$-algebra, but also denote the identity map on various spaces.
- The symbols $\subset$ (for "subset") and $<$ (for "subgroup") don't imply properness (so $\mathbb{Z} \subset \mathbb{Z}$ and $\mathbb{Z} < \mathbb{Z}$).
- The product of permutations $a \in \mathfrak{S}_n$ and $b \in \mathfrak{S}_n$ is defined by $(ab)(i) = a(b(i))$ for all i.
- *Words* over (or in) an *alphabet* I simply mean finite tuples of elements of a set I. It is customary to write such a word $(a_1, a_2, \ldots, a_k)$ as $a_1 a_2 \ldots a_k$ when this is not likely to be confused for multiplication.

Hopefully context will resolve some of the ambiguities.

1.1. Algebras.

Definition 1.1. An *associative k-algebra* A is a $\mathbf{k}$-vector space with an associative operation $A \otimes A \xrightarrow{m} A$, and a *unit* $\mathbf{k} \xrightarrow{u} A$ sending 1 in $\mathbf{k}$ to the two-sided multiplicative identity element 1 in A. One can rephrase this by saying that these diagrams commute:

(1.1)
$$\begin{tikzcd} & A\otimes A\otimes A \arrow[dl, "m\otimes 1"'] \arrow[dr, "1\otimes m"] & \\ A\otimes A \arrow[dr, "m"] & & A\otimes A \arrow[dl, "m"'] \\ & A & \end{tikzcd}$$

(1.2)
$$\begin{tikzcd} A\otimes \mathbf{k} \arrow[d, "1\otimes u"'] & A \arrow[l] \arrow[r] \arrow[d, "1"] & \mathbf{k}\otimes A \arrow[d, "u\otimes 1"] \\ A\otimes A \arrow[r, "m"] & A & A\otimes A \arrow[l, "m"'] \end{tikzcd}$$

where the maps $A \to A \otimes \mathbf{k}$ and $A \to \mathbf{k} \otimes A$ are the isomorphisms sending $a \mapsto a \otimes 1$ and $a \mapsto 1 \otimes a$.

Well-known examples of $\mathbf{k}$-algebras are *tensor* and *symmetric algebras*, which we can think of as algebras of *words* and *multisets*, respectively.

Example 1.2. The *tensor algebra* $T(V) = \bigoplus_{n\geq 0} V^{\otimes n}$ on a $\mathbf{k}$-vector space V, say with $\mathbf{k}$-basis $\{x_i\}_{i\in I}$, is an associative $\mathbf{k}$-algebra with a $\mathbf{k}$-basis of decomposable tensors $x_{i_1}\cdots x_{i_k} := x_{i_1}\otimes\cdots\otimes x_{i_k}$ indexed by *words* $(i_1,\ldots,i_k)$ in the alphabet I, and multiplication defined $\mathbf{k}$-linearly by concatenation of words:

$$m(x_{i_1}\cdots x_{i_k}\otimes x_{j_1}\cdots x_{j_\ell}) := x_{i_1}\cdots x_{i_k}x_{j_1}\cdots x_{j_\ell}.$$

Recall that in an algebra A, when one has a *two-sided ideal* $J \subset A$, meaning a $\mathbf{k}$-linear subspace with $m(J\otimes A), m(A\otimes J) \subset J$, then one can form a *quotient algebra* A/J.

Example 1.3. The *symmetric algebra* $\mathrm{Sym}(V) = \bigoplus_{n\geq 0}\mathrm{Sym}^n(V)$ is the quotient of $T(V)$ by the two-sided ideal generated by all elements $xy - yx$ with x, y in V. It can be identified with a (commutative) polynomial algebra $\mathbf{k}[x_i]_{i\in I}$, having a $\mathbf{k}$-basis of (commutative) monomials $x_{i_1}\cdots x_{i_k}$ as $(i_1,\ldots,i_k)$ runs through all finite multisubsets of I, and with multiplication defined $\mathbf{k}$-linearly via multiset union.

Topology and group theory give more examples.

Example 1.4. The *cohomology algebra* $H^*(X;\mathbf{k}) = \bigoplus_{i\geq 0} H^i(X;\mathbf{k})$ with coefficients in $\mathbf{k}$ for a topological space X has an associative *cup product*. Its unit $\mathbf{k} = H^*(\mathbf{pt};\mathbf{k}) \stackrel{u}{\to} H^*(X;\mathbf{k})$ is induced from the unique (continuous) map $X \to \mathbf{pt}$, where $\mathbf{pt}$ is a one-point space.

Example 1.5. For a group G, the *group algebra* $\mathbf{k}G$ has $\mathbf{k}$-basis $\{t_g\}_{g\in G}$ and multiplication defined $\mathbf{k}$-linearly by $t_g t_h = t_{gh}$, and unit defined by $u(1) = t_e$, where e is the identity element of G.

1.2. **Coalgebras.** If we are to think of the multiplication $A\otimes A \to A$ in an algebra as *putting together* two basis elements of A to get a sum of basis elements of A, then coalgebra structure should be thought of as *taking basis elements apart*.

Definition 1.6. A *co-associative coalgebra* C is a $\mathbf{k}$-vector space C with a *comultiplication*, that is, a $\mathbf{k}$-linear map $C \stackrel{\Delta}{\to} C\otimes C$, and a *counit* $C \stackrel{\epsilon}{\to} \mathbf{k}$ making commutative the diagrams as in (1.1), (1.2) but with *all arrows reversed*:

(1.3)
$$\begin{tikzcd} & C\otimes C\otimes C & \\ C\otimes C \arrow[ur, "\Delta\otimes 1"] & & C\otimes C \arrow[ul, "1\otimes\Delta"'] \\ & C \arrow[ul, "\Delta"] \arrow[ur, "\Delta"'] & \end{tikzcd}$$

$$\begin{tikzcd} C \otimes \mathbf{k} \arrow[r] & C & \mathbf{k} \otimes C \arrow[l] \\ C \otimes C \arrow[u, "1\otimes\epsilon"] & C \arrow[l, "\Delta"] \arrow[r, "\Delta"'] \arrow[u, "1"] & C \otimes C \arrow[u, "\epsilon\otimes 1"'] \end{tikzcd} \tag{1.4}$$

Here the maps $C \otimes \mathbf{k} \to C$ and $\mathbf{k} \otimes C \to C$ are the isomorphisms sending $c \otimes 1 \mapsto c$ and $1 \otimes c \mapsto c$.

One often uses the *Sweedler notation*

$$\Delta(c) = \sum_{(c)} c_1 \otimes c_2 = \sum c_1 \otimes c_2$$

to abbreviate formulas involving Δ. For example, commutativity of the left square in (1.4) asserts that $\sum_{(c)} c_1 \epsilon(c_2) = c$.

Example 1.7. The *homology* $H_*(X;\mathbf{k}) = \bigoplus_{i\geq 0} H_i(X;\mathbf{k})$ for a topological space X is naturally a coalgebra: the (continuous) *diagonal embedding* $X \to X \times X$ sending $x \mapsto (x,x)$ induces a coassociative map

$$H_*(X;\mathbf{k}) \to H_*(X \times X;\mathbf{k}) \cong H_*(X;\mathbf{k}) \otimes H_*(X;\mathbf{k})$$

in which the last isomorphism comes from the *Künneth theorem* with field coefficients $\mathbf{k}$. As before, the unique (continuous) map $X \to \mathbf{pt}$ induces the counit $H_*(X;\mathbf{k}) \stackrel{\epsilon}{\to} H_*(\mathbf{pt};\mathbf{k}) \cong \mathbf{k}$.

1.3. **Morphisms, tensor products, and bialgebras.**

Definition 1.8. A *morphism* of algebras $A \stackrel{\varphi}{\to} B$ makes these diagrams commute:

$$\begin{tikzcd} A \arrow[r, "\varphi"] & B \\ A \otimes A \arrow[u, "m_A"] \arrow[r, "\varphi\otimes\varphi"] & B \otimes B \arrow[u, "m_B"'] \end{tikzcd} \qquad \begin{tikzcd} A \arrow[rr, "\varphi"] & & B \\ & \mathbf{k} \arrow[ul, "u_A"] \arrow[ur, "u_B"'] & \end{tikzcd} \tag{1.5}$$

Here the subscripts on m_A, m_B, u_A, u_B indicate for which algebra they are part of the structure– we will occasionally use such conventions from now on.

Similarly a *morphism of coalgebras* is a $\mathbf{k}$-linear map $C \stackrel{\varphi}{\to} D$ making the reverse diagrams commute:

$$\begin{tikzcd} C \arrow[r, "\varphi"] \arrow[d, "\Delta_C"'] & D \arrow[d, "\Delta_D"] \\ C \otimes C \arrow[r, "\varphi\otimes\varphi"] & D \otimes D \end{tikzcd} \qquad \begin{tikzcd} C \arrow[rr, "\varphi"] \arrow[dr, "\epsilon_C"'] & & D \arrow[dl, "\epsilon_D"] \\ & \mathbf{k} & \end{tikzcd} \tag{1.6}$$

Example 1.9. Continuous maps $X \stackrel{f}{\to} Y$ of topological spaces induce algebra morphisms $H^*(Y;\mathbf{k}) \to H^*(X;\mathbf{k})$, and coalgebra morphisms $H_*(X;\mathbf{k}) \to H_*(Y;\mathbf{k})$.

Definition 1.10. Given two $\mathbf{k}$-algebras A, B, their tensor product $A \otimes B$ also becomes a $\mathbf{k}$-algebra defining the multiplication bilinearly via

$$m((a \otimes b) \otimes (a' \otimes b')) := aa' \otimes bb'$$

or in other words $m_{A\otimes B}$ is the composite map

$$A \otimes B \otimes A \otimes B \stackrel{1\otimes T\otimes 1}{\longrightarrow} A \otimes A \otimes B \otimes B \stackrel{m_A\otimes m_B}{\longrightarrow} A \otimes B$$

where T is the *twist map* $B \otimes A \to A \otimes B$ that sends $b \otimes a \mapsto a \otimes b$.

Here we are omitting the topologist's sign in the twist map which should be present for graded algebras and coalgebras that come from cohomology and homology: for homogeneous elements a and b the topologist's twist map sends

$$b \otimes a \longmapsto (-1)^{\deg(a)\deg(b)} a \otimes b. \tag{1.7}$$

This means that most of our examples which we later call *graded* should actually be considered to live in only *even* degrees, e.g. by artificially doubling their grading. We will ignore this issue, and hope that it causes no confusion later!

The unit element of $A \otimes B$ is $1_A \otimes 1_B$, meaning that the unit map $\mathbf{k} \stackrel{u_{A \otimes B}}{\to} A \otimes B$ is the composite

$$\mathbf{k} \longrightarrow \mathbf{k} \otimes \mathbf{k} \stackrel{u_A \otimes u_B}{\to} A \otimes B.$$

Similarly, given two coalgebras C, D, one can make $C \otimes D$ a coalgebra in which the comultiplication and counit maps are the composites of

$$C \otimes D \stackrel{\Delta_C \otimes \Delta_D}{\longrightarrow} C \otimes C \otimes D \otimes D \stackrel{1 \otimes T \otimes 1}{\longrightarrow} C \otimes D \otimes C \otimes D$$

and

$$C \otimes D \stackrel{\epsilon_C \otimes \epsilon_D}{\longrightarrow} \mathbf{k} \otimes \mathbf{k} \longrightarrow \mathbf{k}.$$

One of the first signs that these definitions interact nicely is the following straightforward proposition.

Proposition 1.11. *When A is both a $\mathbf{k}$-algebra and a $\mathbf{k}$-coalgebra, the following are equivalent:*

- (Δ, ϵ) *are morphisms for the algebra structure* (m, u).
- (m, u) *are morphisms for the coalgebra structure* (Δ, ϵ).
- *These four diagrams commute:*

$$\begin{tikzcd}
 & A \otimes A \arrow[dl, "\Delta \otimes \Delta"'] \arrow[ddr, "m"] & \\
A \otimes A \otimes A \otimes A \arrow[dd, "1 \otimes T \otimes 1"'] & & \\
 & & A \arrow[ddl, "\Delta"] \\
A \otimes A \otimes A \otimes A \arrow[dr, "m \otimes m"'] & & \\
 & A \otimes A &
\end{tikzcd} \tag{1.8}$$

$$\begin{tikzcd}
A \otimes A \arrow[r, "\epsilon \otimes \epsilon"] \arrow[d, "m"'] & \mathbf{k} \otimes \mathbf{k} \arrow[d, "m"] \\
A \arrow[r, "\epsilon"'] & \mathbf{k}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{k} \arrow[r, "u"] \arrow[d, "\Delta"'] & A \arrow[d, "\Delta"] \\
\mathbf{k} \otimes \mathbf{k} \arrow[r, "u \otimes u"'] & A \otimes A
\end{tikzcd}$$

$$\begin{tikzcd}
\mathbf{k} \arrow[rr, "1"] \arrow[dr, "u"'] & & \mathbf{k} \\
 & A \arrow[ur, "\epsilon"'] &
\end{tikzcd}$$

Definition 1.12. Call the $\mathbf{k}$-vector space A a $\mathbf{k}$-*bialgebra* if it is a $\mathbf{k}$-algebra and $\mathbf{k}$-coalgebra satisfying the three equivalent conditions in Proposition 1.11.

Example 1.13. For a group G, one can make the group algebra $\mathbf{k}G$ a coalgebra with counit $\mathbf{k}G \stackrel{\epsilon}{\to} \mathbf{k}$ mapping $t_g \mapsto 1$ for all g in G, and with comultiplication $\mathbf{k}G \stackrel{\Delta}{\to} \mathbf{k}G \otimes \mathbf{k}G$ given by $\Delta(t_g) := t_g \otimes t_g$. Checking the various diagrams in (1.8) commute is easy. For example, one can check the pentagonal diagram on each basis element $t_g \otimes t_h$:

$$\begin{tikzcd}
 & t_g \otimes t_h \arrow[dl, "\Delta\otimes\Delta"'] \arrow[ddr, "m"] & \\
t_g \otimes t_g \otimes t_h \otimes t_h \arrow[dd, "1\otimes T\otimes 1"'] & & \\
 & & t_{gh} \arrow[ddl, "\Delta"] \\
t_g \otimes t_h \otimes t_g \otimes t_h \arrow[dr, "m\otimes m"'] & & \\
 & t_{gh} \otimes t_{gh} &
\end{tikzcd}$$

Remark 1.14. In fact, one can think of adding a bialgebra structure to a $\mathbf{k}$-algebra A as a way of making A-modules M, N have an A-module structure on their tensor product $M \otimes N$: the algebra $A \otimes A$ already acts naturally on $M \otimes N$, so one can let a in A act via $\Delta(a)$ in $A \otimes A$. In the theory of group representations over $\mathbf{k}$, that is, $\mathbf{k}G$-modules M, this is how one defines the *diagonal action* of G on $M \otimes N$, namely t_g acts as $t_g \otimes t_g$.

Definition 1.15. An element x in a coalgebra for which $\Delta(x) = x \otimes x$ and $\epsilon(x) = 1$ is called *group-like*.

An element x in a bialgebra for which $\Delta(x) = 1 \otimes x + x \otimes 1$ is called *primitive*.

Example 1.16. The *tensor algebra* $T(V) = \bigoplus_{n \geq 0} V^{\otimes n}$ is a coalgebra, with counit ϵ equal to the identity on $V^{\otimes 0} = \mathbf{k}$ and the zero map on $V^{\otimes n}$ for $n > 0$, and with comultiplication defined to make the elements x in $V^{\otimes 1} = V$ all primitive:

$$\Delta(x) := 1 \otimes x + x \otimes 1 \text{ for } x \in V^{\otimes 1}.$$

Since the elements of V generate $T(V)$ as a $\mathbf{k}$-algebra, and since $T(V) \otimes T(V)$ is also an associative $\mathbf{k}$-algebra, the universal property of $T(V)$ as the free associative $\mathbf{k}$-algebra on the generators V allows one to define $T(V) \stackrel{\Delta}{\to} T(V) \otimes T(V)$ arbitrarily on V, and extend it as an algebra morphism.

It may not be obvious that this Δ is coassociative, but one can note that

$$((1 \otimes \Delta) \circ \Delta)(x) = x \otimes 1 \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes 1 \otimes x = ((\Delta \otimes 1) \circ \Delta)(x)$$

for every x in V. Hence the two maps $(1 \otimes \Delta) \circ \Delta$ and $(\Delta \otimes 1) \circ \Delta$, considered as algebra morphisms $T(V) \to T(V) \otimes T(V) \otimes T(V)$, must coincide on every element of $T(V)$ since they coincide on V. We leave it as an exercise to check the map ϵ defined as above satisfies the counit axioms (1.4).

Here is a sample calculation in $T(V)$ when V has basis $\{x, y, z\}$:

$$\begin{aligned}
\Delta(xyz) &= \Delta(x)\Delta(y)\Delta(z) \\
&= (1 \otimes x + x \otimes 1)(1 \otimes y + y \otimes 1)(1 \otimes z + z \otimes 1) \\
&= (1 \otimes xy + x \otimes y + y \otimes x + xy \otimes 1)(1 \otimes z + z \otimes 1) \\
&= 1 \otimes xyz + x \otimes yz + y \otimes xz + z \otimes xy \\
&\qquad + xy \otimes z + xz \otimes y + yz \otimes x + xyz \otimes 1
\end{aligned}$$

This illustrates the idea that comultiplication "takes basis elements apart". Here for any $v_1, v_2, \ldots, v_n$ in V one has

$$\Delta(v_1 v_2 \cdots v_n) = \sum v_{j_1} \cdots v_{j_r} \otimes v_{k_1} \cdots v_{k_{n-r}}$$

where the sum is over ordered pairs $(j_1, j_2, \ldots, j_r), (k_1, k_2, \ldots, k_{n-r})$ of complementary subwords of the word $(1, 2, \ldots, n)$. [1]

Recall one can quotient a $\mathbf{k}$-algebra A by a two-sided ideal J to obtain a quotient algebra A/J.

[1]More formally speaking, the sum is over all permutations $(j_1, j_2, \ldots, j_r, k_1, k_2, \ldots, k_{n-r})$ of $(1, 2, \ldots, n)$ satisfying $j_1 < j_2 < \cdots < j_r$ and $k_1 < k_2 < \cdots < k_{n-r}$.

Definition 1.17. In a coalgebra C, a *two-sided coideal* is a $\mathbf{k}$-linear subspace $J \subset C$ for which

$$\begin{aligned}\Delta(J) &\subset J \otimes C + C \otimes J\\ \epsilon(J) &= 0\end{aligned}$$

The quotient space C/J then inherits a coalgebra structure. Similarly, in a bialgebra A, a subset $J \subset A$ which is both a two-sided ideal and two-sided coideal gives rise to a quotient bialgebra A/J.

Example 1.18. The *symmetric algebra* $\mathrm{Sym}(V)$ was the quotient of $T(V)$ by the two-sided ideal J generated by all *commutators* $[x, y] = xy - yx$ for x, y in V. Note that x, y are primitive elements in $T(V)$, and the following very reusable calculation shows that *the commutator of two primitives is primitive*:

$$\begin{aligned}\Delta[x,y] &= \Delta(xy - yx)\\ &= (1 \otimes x + x \otimes 1)(1 \otimes y + y \otimes 1) - (1 \otimes y + y \otimes 1)(1 \otimes x + x \otimes 1)\\ &= 1 \otimes xy - 1 \otimes yx + xy \otimes 1 - yx \otimes 1\\ &\qquad + x \otimes y + y \otimes x - x \otimes y - y \otimes x\\ &= 1 \otimes (xy - yx) + (xy - yx) \otimes 1\\ &= 1 \otimes [x,y] + [x,y] \otimes 1.\end{aligned} \tag{1.9}$$

In particular, the commutators $[x, y]$ have $\Delta[x, y]$ in $J \otimes T(V) + T(V) \otimes J$. They also satisfy $\epsilon([x, y]) = 0$. Since they are generators for J as a two-sided ideal, it is not hard to see this implies $\Delta(J) \subset J \otimes T(V) + T(V) \otimes J$, and $\epsilon(J) = 0$. Thus J is also a two-sided coideal, and $\mathrm{Sym}(V) = T(V)/J$ inherits a bialgebra structure.

In fact we will see in Section 3.1 that symmetric algebras are the universal example of bialgebras which are *graded, connected, commutative, cocommutative*. But first we should define some of these concepts.

Definition 1.19. A *graded* $\mathbf{k}$*-vector space* V is one with a $\mathbf{k}$-vector space direct sum decomposition $V = \bigoplus_{n \geq 0} V_n$. Elements x in V_n are called *homogeneous* of degree n, or $\deg(x) = n$. When we are working with $\mathbf{k} = \mathbb{Z}$, we will always assume that *graded* $\mathbb{Z}$-modules $V = \bigoplus_{n \geq 0} V_n$ have each V_n a *free* $\mathbb{Z}$-module.

One endows tensor products $V \otimes W$ of graded vector spaces V, W with graded vector space structure in which $(V \otimes W)_n := \bigoplus_{i+j=n} V_i \otimes W_j$.

A $\mathbf{k}$-linear map $V \xrightarrow{\varphi} W$ between two graded $\mathbf{k}$-vector spaces is called *graded* if $\varphi(V_n) \subset W_n$ for all n. Say that a $\mathbf{k}$-algebra (coalgebra, bialgebra) is *graded* if it is a graded $\mathbf{k}$-vector space and all of the relevant structure maps (u, ϵ, m, Δ) are graded.

Say that a graded vector space V is *connected* if $V_0 \cong \mathbf{k}$.

Example 1.20. A path-connected space X has its homology and cohomology

$$\begin{aligned}H_*(X; \mathbf{k}) &= \bigoplus_{i \geq 0} H_i(X; \mathbf{k})\\ H^*(X; \mathbf{k}) &= \bigoplus_{i \geq 0} H^i(X; \mathbf{k})\end{aligned}$$

carrying the structure of graded connected coalgebras and algebras, respectively. If in addition, X is a topological group, or even less strongly, an *H-space* (e.g. the *loop space* ΩY on some other space Y), the continuous multiplication map $X \times X \to X$ induces an algebra structure on $H_*(X; \mathbf{k})$ and a coalgebra structure on $H^*(X; \mathbf{k})$, so that each become bialgebras (and these bialgebras are dual to each other in a sense soon to be discussed). This was Hopf's motivation: the (co-)homology of a compact Lie group carries bialgebra structure that explains why it takes a certain form; see Cartier [14, §2].

Example 1.21. Tensor algebras $T(V)$ and symmetric algebras $\mathrm{Sym}(V)$ are graded, once one picks a graded vector space structure for V; then

$$\deg(x_{i_1} \cdots x_{i_k}) = \deg(x_{i_1}) + \cdots + \deg(x_{i_k}).$$

Assuming that $V_0 = 0$, the algebras $T(V)$ and $\mathrm{Sym}(V)$ are connected. For example, we will often say that all elements of V are homogeneous of degree 1, but at other times, it will make sense to have V live in different (positive) degrees.

Exercise 1.22. Check that for a graded connected $\mathbf{k}$-bialgebra A, the gradedness of the unit u and counit ϵ maps, along with commutativity of diagrams (1.2), (1.4), and (1.8) imply

(a) $\mathbf{k}$ lies in A_0,
(b) u is an isomorphism $\mathbf{k} \xrightarrow{u} A_0$, while
(c) the two-sided ideal $\ker \epsilon$ is the space of positive degree elements $I = \bigoplus_{n>0} A_n$.
(d) ϵ restricted to A_0 is the inverse isomorphism $A_0 \xrightarrow{\epsilon} \mathbf{k}$ to u, and
(e) every x in I has comultiplication of the form

$$\Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x)$$

where $\Delta_+(x)$ lies in $I \otimes I$.

1.4. **Antipodes and Hopf algebras.** There is one more piece of structure needed to make a bialgebra a Hopf algebra, although it will come for free in the graded connected case.

Definition 1.23. For any coalgebra C and algebra A, one can endow the $\mathbf{k}$-linear maps $\operatorname{Hom}(C, A)$ with an associative algebra structure called the *convolution algebra*: send f, g in $\operatorname{Hom}(C, A)$ to $f \star g$ defined by $(f \star g)(c) = \sum f(c_1)g(c_2)$, using the *Sweedler notation* $\Delta(c) = \sum c_1 \otimes c_2$. Equivalently, $f \star g$ is the composite

$$C \xrightarrow{\Delta} C \otimes C \xrightarrow{f \otimes g} A \otimes A \xrightarrow{m} A.$$

One sees that $u \circ \epsilon$ is a two-sided identity element for $\star$, meaning that

$$\sum f(c_1)\epsilon(c_2) = f(c) = \sum \epsilon(c_1)f(c_2)$$

by adding a top row to (1.4):

(1.10)
$$\begin{tikzcd}
A \otimes \mathbf{k} \arrow[r] & A & \mathbf{k} \otimes A \arrow[l] \\
C \otimes \mathbf{k} \arrow[r] \arrow[u, "f \otimes 1"] & C \arrow[u, "f"] & \mathbf{k} \otimes C \arrow[l] \arrow[u, "1 \otimes f"'] \\
C \otimes C \arrow[u, "1 \otimes \epsilon"] & C \arrow[l, "\Delta"] \arrow[r, "\Delta"'] \arrow[u, "1"] & C \otimes C \arrow[u, "\epsilon \otimes 1"']
\end{tikzcd}$$

In particular, when one has a bialgebra A, the convolution product $\star$ gives an associative algebra structure on $\operatorname{End}(A) := \operatorname{Hom}(A, A)$.

Definition 1.24. A bialgebra A is called a *Hopf algebra* if there is an element S (called an *antipode* for A) in $\operatorname{End}(A)$ which is a 2-sided inverse under $\star$ for the identity map 1_A. In other words, this diagram commutes:

(1.11)
$$\begin{tikzcd}
 & A \otimes A \arrow[rr, "S \otimes 1_A"] & & A \otimes A \arrow[dr, "m"] & \\
A \arrow[ur, "\Delta"] \arrow[rr, "\epsilon"] \arrow[dr, "\Delta"'] & & \mathbf{k} \arrow[rr, "u"] & & A \\
 & A \otimes A \arrow[rr, "1_A \otimes S"'] & & A \otimes A \arrow[ur, "m"'] &
\end{tikzcd}$$

Or equivalently, if $\Delta(a) = \sum a_1 \otimes a_2$, then

(1.12)
$$\sum_{(a)} S(a_1)a_2 = u(\epsilon(a)) = \sum_{(a)} a_1 S(a_2).$$

Example 1.25. For a group algebra $\mathbf{k}G$, one can define an antipode $\mathbf{k}$-linearly via $S(t_g) = t_{g^{-1}}$. The top pentagon in the above diagram commutes because

$$(S \star 1)(t_g) = m((S \otimes 1)(t_g \otimes t_g)) = S(t_g)t_g = t_{g^{-1}}t_g = t_e = (u \circ \epsilon)(t_g).$$

Note that when it exists, the antipode S is unique, as with all 2-sided inverses in associative algebras: if S, S' are both 2-sided $\star$-inverses to 1_A then

$$S' = (u \circ \epsilon) \star S' = (S \star 1_A) \star S' = S \star (1_A \star S') = S \star (u \circ \epsilon) = S.$$

On the other hand, the next property is not quite as obvious, but is useful when one wants to check that a certain map *is* the antipode in a particular Hopf algebra, by checking it on an algebra generating set.

Proposition 1.26. *The antipode S in a Hopf algebra A is an algebra anti-endomorphism: $S(1) = 1$, and $S(ab) = S(b)S(a)$ for all a, b in A.*

Proof. (see [76, Chap. 4]) Since Δ is an algebra map, one has $\Delta(1) = 1 \otimes 1$, and therefore $1 = u\epsilon(1) = S(1) \cdot 1 = S(1)$.

To show $S(ab) = S(b)S(a)$, consider $A \otimes A$ as a coalgebra and A as an algebra. Then $\mathrm{Hom}(A \otimes A, A)$ is an associative algebra with a convolution product $\circledast$ (to be distinguished from the convolution $\star$ on $\mathrm{End}(A)$), having two-sided identity element $u_A \epsilon_{A \otimes A}$. We will show below that these three elements of $\mathrm{Hom}(A \otimes A, A)$

$$\begin{aligned} f(a \otimes b) &= ab \\ g(a \otimes b) &= S(b)S(a) \\ h(a \otimes b) &= S(ab) \end{aligned}$$

have the property that

$$h \circledast f = u_A \epsilon_{A \otimes A} = f \circledast g \tag{1.13}$$

which would then show the desired equality $h = g$ via associativity:

$$h = h \circledast (u_A \epsilon_{A \otimes A}) = h \circledast (f \circledast g) = (h \circledast f) \circledast g = (u_A \epsilon_{A \otimes A}) \circledast g = g.$$

So we evaluate the three elements in (1.13) on $a \otimes b$, assuming $\Delta(a) = \sum_{(a)} a_1 \otimes a_2$ and $\Delta(b) = \sum_{(b)} b_1 \otimes b_2$, and hence $\Delta(ab) = \sum_{(a),(b)} a_1 b_1 \otimes a_2 b_2$. One has

$$(u_A \epsilon_{A \otimes A})(a \otimes b) = u_A(\epsilon_A(a)\epsilon_A(b)) = u_A(\epsilon_A(ab)).$$

$$\begin{aligned} (h \circledast f)(a \otimes b) &= \sum_{(a),(b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2) \\ &= \sum_{(a),(b)} S(a_1 b_1) a_2 b_2 \\ &= (S \star 1_A)(ab) = u_A(\epsilon_A(ab)). \end{aligned}$$

$$\begin{aligned} (f \circledast g)(a \otimes b) &= \sum_{(a),(b)} f(a_1 \otimes b_1) g(a_2 \otimes b_2) \\ &= \sum_{(a),(b)} a_1 b_1 S(b_2) S(a_2) \\ &= \sum_{(a)} a_1 \cdot (1_A \star S)(b) \cdot S(a_2) \\ &= u_A(\epsilon_A(b)) \sum_{(a)} a_1 S(a_2) = u_A(\epsilon_A(b)) u_A(\epsilon_A(a)) = u_A(\epsilon_A(ab)) \end{aligned}$$

□

Remark 1.27. Recall from Remark 1.14 that the comultiplication on a bialgebra A allows one to define an A-module structure on the tensor product $M \otimes N$ of two A-modules M, N. Similarly, the anti-endomorphism S in a Hopf algebra allows one to turn *left* A-modules into *right* A-modules, or vice-versa. E.g., left A-modules M naturally have a right A-module structure on the dual space $M^* := \mathrm{Hom}(M, \mathbf{k})$, defined via $(fa)(m) := f(am)$ for f in M^* and a in A. The antipode S can be used to turn this back into a left A-module M^*, via $(af)(m) = f(S(a)m)$.

For groups G and left $\mathbf{k}G$-modules (group representations) M, this is how one defines the *contragredient action* of G on M^*, namely t_g acts as $(t_g f)(m) = f(t_{g^{-1}} m)$.

Along the same lines, we are supposed to think of the counit $A \xrightarrow{\epsilon} \mathbf{k}$ as giving a way to make $\mathbf{k}$ into a *trivial* A-module.

Corollary 1.28. *Commutativity of A implies the antipode is an involution: $S^2 = 1_A$.*

Proof. One checks that $S^2 = S \circ S$ is a right $\star$-inverse to S, and hence $S^2 = 1_A$:

$$\begin{aligned}(S \star S^2)(a) &= \sum_{(a)} S(a_1) S^2(a_2) \\ &= S\left(\sum_{(a)} S(a_2) a_1\right) \text{ by Proposition 1.26} \\ &= S\left(u(\epsilon(a))\right) \\ &= u(\epsilon(a)) \text{ since } S(1) = 1 \text{ by Proposition 1.26.}\end{aligned}$$

□

Remark 1.29. We won't need it, but it is easy to adapt the above proof to show that $S^2 = 1_A$ also holds for *cocommutative* Hopf algebras; see [56, Corollary 1.5.12] or [76, Chapter 4]. For a general Hopf algebra which is not finite-dimensional over $\mathbf{k}$, the antipode S may not even have finite order, even in the graded connected setting. E.g., Aguiar and Sottile [6] show that the Malvenuto-Reutenauer Hopf algebra of permutations has antipode of infinite order. In general, antipodes need not even be invertible [77].

In our frequent setting of graded connected bialgebras, antipodes come for free.

Proposition 1.30. *A graded connected bialgebra A has a unique antipode S, which is a graded map $A \xrightarrow{S} A$, endowing it with a Hopf structure.*

Proof. Let us try to define a ($\mathbf{k}$-linear) left $\star$-inverse S to 1_A on each homogeneous component A_n, via induction on n.

In the base case $n = 0$, Proposition 1.26 and its proof show that one must define $S(1) = 1$ so S is the identity on $A_0 = \mathbf{k}$.

In the inductive step, recall from Exercise 1.22 that a homogeneous element a of degree $n > 0$ has $\Delta(a) = a \otimes 1 + \sum a_1 \otimes a_2$, with each $\deg(a_1) < n$. Hence in order to have $S \star 1_A = u\epsilon$, one must define $S(a)$ in such a way that $S(a) \cdot 1 + \sum S(a_1) a_2 = u\epsilon(a) = 0$ and hence $S(a) := -\sum S(a_1) a_2$, where $S(a_1)$ have already been uniquely defined by induction. This does indeed define such a left $\star$-inverse S to 1_A, by induction. It is also a graded map by induction.

The same argument shows how to define a right $\star$-inverse S' to 1_A. Then $S = S'$ is a two-sided $\star$-inverse to 1_A by the associativity of $\star$. □

Here is another consequence of the fact that $S(1) = 1$.

Proposition 1.31. *In bialgebras, primitive elements x have $\epsilon(x) = 0$, and in Hopf algebras, they have $S(x) = -x$.*

Proof. In a bialgebra, $\epsilon(1) = 1$. Hence $\Delta(x) = 1 \otimes x + x \otimes 1$ implies via (1.4) that $1 \cdot \epsilon(x) + \epsilon(1) x = x$, so $\epsilon(x) = 0$. It also implies via (1.11) that $S(x)1 + S(1)x = u\epsilon(x) = u(0) = 0$, so $S(x) = -x$. □

Thus whenever A is a Hopf algebra generated as an algebra by its primitive elements, S is the unique anti-endomorphism that negates all primitive elements.

Example 1.32. The tensor and symmetric algebras $T(V)$ and $\mathrm{Sym}(V)$ are each generated by V, which contains only primitive elements in either case. Hence one has in $T(V)$ that

$$S(x_{i_1} x_{i_2} \cdots x_{i_k}) = (-x_{i_k}) \cdots (-x_{i_2})(-x_{i_1}) = (-1)^k x_{i_k} \cdots x_{i_2} x_{i_1} \tag{1.14}$$

for each word $(i_1, \ldots, i_k)$ in the alphabet I. The same holds in $\mathrm{Sym}(V)$ for each multiset $(i_1, \ldots, i_k)$, recalling that the monomials are now commutative. In other words, for a commutative polynomial $f(\mathbf{x})$ in $\mathrm{Sym}(V)$, the antipode S sends f to $f(-\mathbf{x})$, negating all the variables.

The antipode for a graded connected Hopf algebra has an interesting formula due to Takeuchi [77], reminiscent of P. Hall's formula for the Möbius function of a poset[2]. For the sake of stating this, define the k-fold *tensor power* $A^{\otimes k} = A \otimes \cdots \otimes A$ and define *iterated multiplication and comultiplication* maps

$$\begin{array}{ccc} A^{\otimes k} & \stackrel{m^{(k-1)}}{\longrightarrow} & A \\ A & \stackrel{\Delta^{(k-1)}}{\longrightarrow} & A^{\otimes k} \end{array}$$

by $m^{(-1)} = u$, $\Delta^{(-1)} = \epsilon$, $m^{(0)} = \Delta^{(0)} = 1_A$, and

$$\begin{array}{lll} m^{(k)} & = m \circ (1_A \otimes m^{(k-1)}) & = m \circ (m^{(k-1)} \otimes 1_A) \\ \Delta^{(k)} & = (1_A \otimes \Delta^{(k-1)}) \circ \Delta & = (\Delta^{(k-1)} \otimes 1_A) \circ \Delta \end{array}$$

using associativity, coassociativity to show that these are well-defined. They are how one gives meaning to the right sides of these equations:

$$m^{(k)}(a^{(1)} \otimes \cdots \otimes a^{(k+1)}) = a^{(1)} \cdots a^{(k+1)}$$

$$\Delta^{(k)}(b) = \sum b_1 \otimes \cdots \otimes b_{k+1} \text{ in Sweedler notation.}$$

Proposition 1.33. *In a graded connected Hopf algebra A, the antipode has formula*

$$\begin{aligned} S &= \sum_{k \geq 0} (-1)^k m^{(k-1)} f^{\otimes k} \Delta^{(k-1)} \\ &= u\epsilon - f + m \circ f^{\otimes 2} \circ \Delta - m^{(2)} \circ f^{\otimes 3} \circ \Delta^{(2)} + \cdots \end{aligned} \tag{1.15}$$

where $f := 1_A - u\epsilon$ in $\mathrm{End}(A)$.

Proof. We argue as in [77, Lemma 14] or [6, §5]. For any f in $\mathrm{End}(A)$ one has this explicit formula for its k-fold convolution power $f^{\star k} := f \star \cdots \star f$ in terms of its tensor powers $f^{\otimes k} := f \otimes \cdots \otimes f$:

$$f^{\star k} = m^{(k-1)} \circ f^{\otimes k} \circ \Delta^{(k-1)}.$$

Therefore any f annihilating A_0 will be *locally $\star$-nilpotent* on A, meaning that for each n one has that A_n is annihilated by $f^{\star m}$ for every $m > n$: homogeneity forces that for a in A_n, every summand of $\Delta^{(m-1)}(a)$ must contain among its m tensor factors at least one factor lying in A_0, so each summand is annihilated by $f^{\otimes m}$, and $f^{\star m}(a) = 0$.

In particular such f have the property that $u\epsilon + f$ has as two-sided $*$-inverse

$$\begin{aligned} (u\epsilon + f)^{\star(-1)} &= u\epsilon - f + f \star f - f \star f \star f + \cdots \\ &= \sum_{k \geq 0} (-1)^k f^{\star k} = \sum_{k \geq 0} (-1)^k m^{(k-1)} \circ f^{\otimes k} \circ \Delta^{(k-1)}. \end{aligned}$$

The proposition follows upon taking $f := 1_A - u\epsilon$, which annihilates A_0. ☐

Remark 1.34. In fact, one can see that Takeuchi's formula applies more generally to define an antipode $A \xrightarrow{S} A$ in any (not necessarily graded) bialgebra A where the map $1_A - u\epsilon$ is locally $\star$-nilpotent.

It is also worth noting that the proof of Proposition 1.33 gives an alternate proof of Proposition 1.30

To finish our discussion of antipodes, we mention some properties (taken from [76, Chap. 4]) relating antipodes to convolutional inverses. It also shows that a bialgebra morphism between Hopf algebras automatically respects the antipodes.

Proposition 1.35. *Let H be a Hopf algebra with antipode S.*

(a) *For any algebra A and algebra morphism $H \xrightarrow{\alpha} A$, one has $\alpha \circ S = \alpha^{\star -1}$, the convolutional inverse to α in* $\mathrm{Hom}(H, A)$.

(b) *For any coalgebra C and coalgebra morphism $C \xrightarrow{\gamma} H$, one has $S \circ \gamma = \gamma^{\star -1}$, the convolutional inverse to γ in* $\mathrm{Hom}(C, H)$.

(c) *If H_1, H_2 are Hopf algebras with antipodes S_1, S_2, then any bialgebra morphism $H_1 \xrightarrow{\beta} H_2$ is a Hopf morphism, that is, it commutes with the antipodes, since $\beta \circ S_1 \stackrel{(a)}{=} \beta^{\star -1} \stackrel{(b)}{=} S_2 \circ \beta$.*

[2] In fact, for incidence Hopf algebras, Takeuchi's formula generalizes Hall's formula– see Corollary 6.11.

Proof. We prove (a); the proof of (b) is similar, and (c) follows immediately from (a),(b) as indicated in its statement. Begin with the following: given an algebra morphism $A \xrightarrow{\alpha} A'$ and coalgebra morphism $C \xrightarrow{\gamma} C'$, the pre- and post-composition $f \longmapsto \alpha \circ f \circ \gamma$ is a convolution algebra morphism $\mathrm{Hom}(C', A) \xrightarrow{\varphi} \mathrm{Hom}(C, A')$, as one can check that $\varphi(u_A \circ \epsilon_{C'}) = \alpha \circ u_A \circ \epsilon_{C'} \circ \gamma = u_{A'} \circ \epsilon_C$, and

$$\begin{aligned} \varphi(f \star g) &= \alpha \circ (f \star g) \circ \gamma \\ &= \alpha \circ m \circ (f \otimes g) \circ \Delta \circ \gamma \\ &= m \circ (\alpha \otimes \alpha) \circ (f \otimes g) \circ (\gamma \otimes \gamma) \circ \Delta \\ &= m \circ ((\alpha \circ f \circ \gamma) \otimes (\alpha \circ g \circ \gamma)) \circ \Delta \\ &= \varphi(f) \star \varphi(g). \end{aligned} \tag{1.16}$$

For assertion (a), note that a special case of the above observation shows $f \longmapsto \alpha \circ f$ gives a convolution-algebra morphism $\mathrm{Hom}(H, H) \xrightarrow{\varphi} \mathrm{Hom}(H, A)$, and hence

$$\alpha \circ S = \varphi(S) = \varphi\left((1_H)^{\star -1}\right) = (\varphi(1_H))^{\star -1} = (\alpha \circ 1_H)^{\star -1} = \alpha^{\star -1}.$$

$\square$

1.5. **Commutativity, cocommutativity.**

Definition 1.36. Say that the $\mathbf{k}$-algebra A is *commutative* if $ab = ba$, that is, this diagram commutes:

$$\begin{tikzcd} A \otimes A \arrow[rr, "T"] \arrow[dr, "m"'] & & A \otimes A \arrow[dl, "m"] \\ & A & \end{tikzcd} \tag{1.17}$$

Say that the $\mathbf{k}$-coalgebra C is *cocommutative* if this diagram commutes:

$$\begin{tikzcd} C \otimes C \arrow[rr, "T"] & & C \otimes C \\ & C \arrow[ul, "\Delta"] \arrow[ur, "\Delta"'] & \end{tikzcd} \tag{1.18}$$

Example 1.37. Group algebras $\mathbf{k}G$ are always cocommutative, but commutative if and only if G is abelian.

Tensor algebras $T(V)$ are always cocommutative, but commutative if and only if $\dim_{\mathbf{k}} V \leq 1$.

Symmetric algebras $\mathrm{Sym}(V)$ are always cocommutative and commutative.

Homology and cohomology of H-spaces are always cocommutative and commutative *in the topologist's sense* where one reinterprets that twist map $A \otimes A \xrightarrow{T} A \otimes A$ to have the extra sign as in (1.7).

Note how the cocommutative Hopf algebras $T(V), \mathrm{Sym}(V)$ have much of their structure controlled by their subspace V of primitive elements. This is not far from the truth in general, and closely related to Lie algebras.

Exercise 1.38. Recall that a *Lie algebra* over $\mathbf{k}$ is a $\mathbf{k}$-vector space $\mathfrak{g}$ with a $\mathbf{k}$-bilinear map $[\cdot, \cdot]$ that satisfies $[x, x] = 0$ for x in $\mathfrak{g}$, and the *Jacobi identity*

$$\begin{aligned} &[x, [y, z]] = [[x, y], z] + [y, [x, z]], \text{ or equivalently} \\ &[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0. \end{aligned}$$

(a) Check that any associative algebra A gives rise to a Lie algebra by means of the commutator operation $[a, b] := ab - ba$.

(b) If A is also a bialgebra, show that the $\mathbf{k}$-subspace of primitive elements $\mathfrak{p} \subset A$ is closed under the Lie bracket, that is, $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{p}$, and hence forms a Lie subalgebra.

Conversely, given a Lie algebra $\mathfrak{p}$, one constructs the *universal enveloping algebra* $\mathcal{U}(\mathfrak{p}) := T(\mathfrak{p})/J$ as the quotient of the tensor algebra $T(\mathfrak{p})$ by the two-sided ideal J generated by all elements $xy - yx - [x, y]$ for x, y in $\mathfrak{p}$.

(c) Show that J is also a two-sided coideal in $T(\mathfrak{p})$ for its usual coalgebra structure, and hence the quotient $\mathcal{U}(\mathfrak{p})$ inherits the structure of a cocommutative bialgebra.

(d) Show that the antipode S on $T(\mathfrak{p})$ preserves J, meaning that $S(J) \subset J$, and hence $\mathcal{U}(\mathfrak{p})$ inherits the structure of a (cocommutative) Hopf algebra.

There are theorems, discussed in [14, §3.8], [56, Chap. 5], giving various mild hypotheses in addition to cocommutativity which imply that the inclusion of the space $\mathfrak{p}$ of primitives in a Hopf algebra A extends to a Hopf isomorphism $\mathcal{U}(\mathfrak{p}) \cong A$.

1.6. **Duals.** Recall that for finite dimensional $\mathbf{k}$-vector spaces V, taking the dual space $V^* := \operatorname{Hom}(V, \mathbf{k})$ reverses $\mathbf{k}$-linear maps. That is, $V \overset{\varphi}{\to} W$ induces $W^* \overset{\varphi^*}{\to} V^*$ defined uniquely by

$$(f, \varphi(v)) = (\varphi^*(f), v)$$

in which (f, v) is the bilinear pairing $V^* \times V \to \mathbf{k}$ sending $(f, v) \mapsto f(v)$. When φ is expressed in terms of a basis $\{v_i\}_{i \in I}$ for V, the map φ^* is expressed by the transpose matrix in terms of the dual basis $\{f_i\}_{i \in I}$ for V^* that satisfies $(f_i, v_j) = \delta_{i,j}$, where $\delta_{i,j}$ is the *Kronecker delta*: $\delta_{i,j} = 1$ if $i = j$ and 0 else.

When discussing graded vector spaces $V = \bigoplus_{n \geq 0} V_n$ *of finite type*, meaning that each V_n is finite-dimensional, note that V^* can contain functionals f supported on infinitely many V_n. Instead we will consider the subspace $V^o := \bigoplus_{n \geq 0} (V_n)^* \subset V^*$, sometimes called the *restricted dual*, consisting of the functions f that vanish on all but finitely many V_n, which is again a graded vector space of finite type.

Reversing the diagrams should then make it clear that, in the finite-dimensional or finite-type situation, duals of algebras are coalgebras, and vice-versa, and duals of bialgebras or Hopf algebras are bialgebras or Hopf algebras. For example, the product in a Hopf algebra A uniquely defines the coproduct of A^o via adjointness:

$$(\Delta_{A^o}(f), a \otimes b)_{A \otimes A} = (f, ab)_A.$$

Thus if A has a basis $\{a_i\}_{i \in I}$ with *product structure constants* $\{c^i_{j,k}\}$, meaning

$$a_j a_k = \sum_{i \in I} c^i_{j,k} a_i,$$

then the dual basis $\{f_i\}_{i \in I}$ has the same $\{c^i_{j,k}\}$ as its *coproduct structure constants*:

$$\Delta_{A^o}(f_i) = \sum_{(j,k) \in I \times I} c^i_{j,k} f_j \otimes f_k.$$

Another example of a Hopf algebra is provided by the so-called shuffle algebra. Before we introduce it, let us define the *shuffles* of two words:

Definition 1.39. Given two words $a = (a_1, a_2, ..., a_n)$ and $b = (b_1, b_2, ..., b_m)$, the *multiset of shuffles of a and b* is defined as the multiset

$$\left\{ \left(c_{w(1)}, c_{w(2)}, ..., c_{w(n+m)} \right) \ \mid \ w \in \operatorname{Sh}_{n,m} \right\}_{\text{multiset}},$$

where $(c_1, c_2, ..., c_{n+m})$ is the concatenation $a \cdot b = (a_1, a_2, ..., a_n, b_1, b_2, ..., b_m)$, and where Sh_{n+m} is the subset

$$\left\{ w \in \mathfrak{S}_{n+m} \ \mid \ w^{-1}(1) < w^{-1}(2) < \cdots < w^{-1}(n); \ w^{-1}(n+1) < w^{-1}(n+2) < \cdots < w^{-1}(n+m) \right\}$$

of the symmetric group $\mathfrak{S}_{n+m}$. Informally speaking, the shuffles of the two words a and b are the words obtained by overlaying the words a and b, after first moving their letters apart so that no letters get superimposed when the words are overlayed. In particular, any shuffle of a and b contains a and b as subsequences. The multiset of shuffles of a and b has $\binom{m+n}{n}$ elements (counted with multiplicity) and is denoted by $a ⧢ b$. For instance, the shuffles of $(1, 2, 1)$ and $(3, 2)$ are

$$\begin{gathered}
(\underline{1}, \underline{2}, \underline{1}, 3, 2), (\underline{1}, \underline{2}, 3, \underline{1}, 2), (\underline{1}, \underline{2}, 3, 2, \underline{1}), (\underline{1}, 3, \underline{2}, \underline{1}, 2), (\underline{1}, 3, \underline{2}, 2, \underline{1}), \\
(\underline{1}, 3, 2, \underline{2}, \underline{1}), (3, \underline{1}, \underline{2}, \underline{1}, 2), (3, \underline{1}, \underline{2}, 2, \underline{1}), (3, \underline{1}, 2, \underline{2}, \underline{1}), (3, 2, \underline{1}, \underline{2}, \underline{1}),
\end{gathered}$$

listed here with the multiplicities with which they appear in the multiset $(1, 2, 1) ⧢ (3, 2)$. Here we have underlined the letters taken from a – that is, the letters at positions $w^{-1}(1)$, $w^{-1}(2)$, ..., $w^{-1}(n)$.

Example 1.40. When $A = T(V)$ is the tensor algebra for a finite-dimensional $\mathbf{k}$-vector space V, having $\mathbf{k}$-basis $\{x_i\}_{i \in I}$, its restricted dual A^o is another Hopf algebra whose basis $\{y_{(i_1,\ldots,i_\ell)}\}$ is indexed by words in the alphabet I, called the *shuffle algebra* of V^*. Duality shows that the *cut* coproduct in A^o is defined by

$$\Delta y_{(i_1,\ldots,i_\ell)} = \sum_{j=0}^{\ell} y_{(i_1,\ldots,i_j)} \otimes y_{(i_{j+1},i_{j+2},\ldots,i_\ell)}.$$

For example,

$$\Delta y_{abcb} = y_\varnothing \otimes y_{abcb} + y_a \otimes y_{bcb} + y_{ab} \otimes y_{cb} + y_{abc} \otimes y_b + y_{abcb} \otimes y_\varnothing$$

Duality also shows that the *shuffle* product in A^o will be given by

$$y_{(i_1,\ldots,i_\ell)} y_{(j_1,\ldots,j_m)} = \sum_{\mathbf{k}=(k_1,\ldots,k_{\ell+m}) \in \mathbf{i} \sqcup\!\sqcup \mathbf{j}} y_{(k_1,\ldots,k_{\ell+m})}$$

where $\mathbf{i} \sqcup\!\sqcup \mathbf{j}$ (as in Definition 1.39) denotes the multiset of the $\binom{\ell+m}{\ell}$ words obtained as *shuffles* of the two words $\mathbf{i} = (i_1, \ldots, i_\ell)$ and $\mathbf{j} = (j_1, \ldots, j_m)$. For example,

$$\begin{aligned} y_{ab} y_{cb} &= y_{abcb} + y_{acbb} + y_{cabb} + y_{cabb} + y_{acbb} + y_{cbab} \\ &= y_{abcb} + 2y_{acbb} + 2y_{cabb} + y_{cbab} \end{aligned}$$

Equivalently, one has

$$y_{(i_1,i_2,\ldots,i_\ell)} y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})} = \sum_{\substack{w \in \mathfrak{S}_{\ell+m}: \\ w(1)<\cdots<w(\ell), \\ w(\ell+1)<\cdots<w(\ell+m)}} y_{(i_{w^{-1}(1)}, i_{w^{-1}(2)}, \ldots, i_{w^{-1}(\ell+m)})}.$$

Lastly, the antipode S of A^o is the adjoint of the antipode of $A = T(V)$ described in (1.14):

$$S y_{(i_1,i_2,\ldots,i_\ell)} = (-1)^\ell y_{(i_\ell,\ldots,i_2,i_1)}.$$

Exercise 1.41. Let V be a 1-dimensional vector space with basis element x, so $\mathrm{Sym}(V) \cong \mathbf{k}[x]$, with $\mathbf{k}$-basis $\{1 = x^0, x^1, x^2, \ldots\}$.
(a) Check that the powers x^i satisfy

$$\begin{aligned} x^i \cdot x^j &= x^{i+j} \\ \Delta(x^n) &= \sum_{i+j=n} \binom{n}{i} x^i \otimes x^j \\ S(x^n) &= (-1)^n x^n \end{aligned}$$

(b) Check that the dual basis elements $\{f^{(0)}, f^{(1)}, f^{(2)}, \ldots\}$ for $\mathrm{Sym}(V)^o$, defined by $f^{(i)}(x^j) = \delta_{i,j}$, satisfy

$$\begin{aligned} f^{(i)} f^{(j)} &= \binom{i+j}{i} f^{(i+j)} \\ \Delta(f^{(n)}) &= \sum_{i+j=n} f^{(i)} \otimes f^{(j)} \\ S(f^{(n)}) &= (-1)^n f^{(n)} \end{aligned}$$

(c) Show that if $\mathbf{k}$ has characteristic zero, then the map $\mathrm{Sym}(V)^o \to \mathrm{Sym}(V)$ sending $f^{(n)} \mapsto \frac{x^n}{n!}$ is a graded Hopf isomorphism.

For this reason, the Hopf structure on $\mathrm{Sym}(V)^o$ is called a *divided power algebra*.
(d) Show that when $\mathbf{k}$ has characteristic $p > 0$, one has $(f^{(1)})^p = 0$, and hence why there can be no Hopf isomorphism $\mathrm{Sym}(V)^o \to \mathrm{Sym}(V)$.

Exercise 1.42. Let V have $\mathbf{k}$-basis $\{x_1, \ldots, x_n\}$, and let $V \oplus V$ have $\mathbf{k}$-basis $\{x_1, \ldots, x_n, y_1, \ldots, y_n\}$, so that one has isomorphisms

$$\mathrm{Sym}(V \oplus V) \cong \mathbf{k}[\mathbf{x}, \mathbf{y}] \cong \mathbf{k}[\mathbf{x}] \otimes \mathbf{k}[\mathbf{y}] \cong \mathrm{Sym}(V) \otimes \mathrm{Sym}(V).$$

(a) Show that our usual coproduct on $\mathrm{Sym}(V)$ can be re-expressed as follows:

$$\begin{array}{ccc} \mathrm{Sym}(V) & & \mathrm{Sym}(V)\otimes\mathrm{Sym}(V) \\ \| & & \| \\ \mathbf{k}[\mathbf{x}] & \stackrel{\Delta}{\longrightarrow} & \mathbf{k}[\mathbf{x},\mathbf{y}] \\ f(x_1,\ldots,x_n) & \longmapsto & f(x_1+y_1,\ldots,x_n+y_n) \end{array}$$

In other words, it is induced from the diagonal map

$$\begin{array}{ccc} V & \longrightarrow & V\oplus V \\ x_i & \longmapsto & x_i+y_i \end{array} \tag{1.19}$$

(b) One can similarly define a coproduct on the *exterior algebra* $\wedge V$, which is the quotient $T(V)/J$ where J is the two-sided ideal generated by the elements $\{x^2(=x\otimes x)\}_{x\in V}$ in $T^2(V)$. This becomes a graded commutative algebra

$$\wedge V = \bigoplus_{d=0}^{n} \wedge^d V \left(= \bigoplus_{d=0}^{\infty} \wedge^d V \right),$$

if one views the elements of $V=\wedge^1 V$ as having *odd* degree, and uses the topologist's sign conventions (as in (1.7)). One again has $\wedge(V\oplus V)=\wedge V\otimes\wedge V$ as graded algebras. Show that one can again let the diagonal map (1.19) induce a map

$$\begin{array}{ccc} \wedge(V) & \stackrel{\Delta}{\longrightarrow} & \wedge V\otimes\wedge V \\ f(x_1,\ldots,x_n) & \longmapsto & f(x_1+y_1,\ldots,x_n+y_n) \\ \| & & \| \\ \sum c_{i_1,\ldots,i_d}\cdot x_{i_1}\wedge\cdots\wedge x_{i_d} & & \sum c_{i_1,\ldots,i_d}\cdot(x_{i_1}+y_{i_1})\wedge\cdots\wedge(x_{i_d}+y_{i_d}) \end{array}$$

which makes $\wedge V$ into a graded connected Hopf algebra.
(c) Show that in the tensor algebra $T(V)$, if one views the elements of $V=V^{\otimes 1}$ as having odd degree, then for any x in V one has $\Delta(x^2)=1\otimes x^2+x^2\otimes 1$.
(Hint: Make sure you use the convention (1.7) in the twist map!)
(d) Use part (d) to show that the two-sided ideal $J\subset T(V)$ generated by $\{x^2\}_{x\in V}$ is also a two-sided coideal, and hence the quotient $\wedge V=T(V)/J$ inherits the structure of a bialgebra. Check that the coproduct on $\wedge V$ inherited from $T(V)$ is the same as the one defined in part (b).

2. Review of symmetric functions Λ as Hopf algebra

Here we review the ring of symmetric functions, borrowing heavily from standard treatments, such as Macdonald [49, Chap. I], Sagan [62, Chap. 4], and Stanley [72, Chap. 7], but emphasizing the Hopf structure early on.

2.1. **Definition of Λ.** As before, $\mathbf{k}$ here could either be a field or the integers $\mathbb{Z}$. Given an infinite variable set $\mathbf{x} = (x_1, x_2, \ldots)$, a monomial $\mathbf{x}^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots$ is indexed by an element $\alpha = (\alpha_1, \alpha_2, \ldots)$ in $\mathbb{N}^\infty$ having finite support; such α are called *weak compositions*. The nonzero ones among the integers $\alpha_1, \alpha_2, \ldots$ are called the *parts* of the weak composition α. We will consider the ring $R(\mathbf{x})$ of formal power series $f(\mathbf{x}) = \sum_\alpha c_\alpha \mathbf{x}^\alpha$ with c_α in $\mathbf{k}$ *of bounded degree*, that is, where there exists some bound $d = d(f)$ for which $\deg(\mathbf{x}^\alpha) := \sum_i \alpha_i > d$ implies $c_\alpha = 0$. It is easy to see that the product of two such power series is well-defined, and also has bounded degree.

The symmetric group $\mathfrak{S}_n$ permuting the first n variables $x_1, \ldots, x_n$ acts as a group of automorphisms on $R(\mathbf{x})$, as does the union $\mathfrak{S}_{(\infty)} = \bigcup_{n \geq 1} \mathfrak{S}_n$ of the infinite ascending chain $\mathfrak{S}_1 \subset \mathfrak{S}_2 \subset \cdots$ of symmetric groups.

Definition 2.1. The *ring of symmetric functions in* $\mathbf{x}$ *with coefficients in* $\mathbf{k}$, denoted $\Lambda = \Lambda_{\mathbf{k}} = \Lambda(\mathbf{x}) = \Lambda_{\mathbf{k}}(\mathbf{x})$, is the $\mathfrak{S}_{(\infty)}$-invariant subalgebra $R(\mathbf{x})^{\mathfrak{S}_{(\infty)}}$ of $R(\mathbf{x})$:

$$\Lambda := \left\{ f = \sum_\alpha c_\alpha \mathbf{x}^\alpha \in R(\mathbf{x}) : c_\alpha = c_\beta \text{ if } \alpha, \beta \text{ lie in the same } \mathfrak{S}_{(\infty)}\text{-orbit} \right\}.$$

Note that Λ is a graded $\mathbf{k}$-algebra, since $\Lambda = \bigoplus_{n \geq 0} \Lambda_n$ where Λ_n are the symmetric functions $f = \sum_\alpha c_\alpha \mathbf{x}^\alpha$ which are *homogeneous of degree* n, meaning $\deg(\mathbf{x}^\alpha) = n$ for all $c_\alpha \neq 0$.

Definition 2.2. A *partition* $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell, 0, 0, \ldots)$ is a weak composition whose parts weakly decrease: $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$. One sometimes omits trailing zeroes from a partition. The (uniquely defined) ℓ is said to be the *length* of the partition λ. The sum $\lambda_1 + \lambda_2 + \cdots + \lambda_\ell = \lambda_1 + \lambda_2 + \cdots$ of all parts of λ is called the *size* of λ and denoted by $|\lambda|$; for a given integer n, the partitions of size n are referred to as the *partitions of* n. The empty partition $()$ is denoted by $\varnothing$.

Every weak composition α lies in the $\mathfrak{S}_{(\infty)}$-orbit of a unique partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell, 0, 0, \ldots)$ with $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$. For any partition λ, define the *monomial symmetric function*

$$m_\lambda := \sum_{\alpha \in \mathfrak{S}_{(\infty)} \lambda} \mathbf{x}^\alpha. \tag{2.1}$$

Letting λ run through the set Par of all partitions, this gives the *monomial* $\mathbf{k}$*-basis* $\{m_\lambda\}$ of Λ. Letting λ run only through the set Par_n of partitions of n gives the monomial $\mathbf{k}$-basis for Λ_n.

Example 2.3. For $n = 3$, one has

$$\begin{aligned} m_{(3)} &= x_1^3 + x_2^3 + x_3^3 + \cdots \\ m_{(2,1)} &= x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + \cdots \\ m_{(1,1,1)} &= x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4 + x_1 x_2 x_5 + \cdots \end{aligned}$$

Remark 2.4. It is sometimes convenient to work with finite variable sets $x_1, \ldots, x_n$, which one justifies as follows. Note that the algebra homomorphism

$$R(\mathbf{x}) \to R(x_1, \ldots, x_n) = \mathbf{k}[x_1, \ldots, x_n]$$

which sends $x_{n+1}, x_{n+2}, \ldots$ to 0 restricts to an algebra homomorphism

$$\Lambda_{\mathbf{k}}(\mathbf{x}) \to \Lambda_{\mathbf{k}}(x_1, \ldots, x_n) = \mathbf{k}[x_1, \ldots, x_n]^{\mathfrak{S}_n}.$$

Furthermore, this last homomorphism is a $\mathbf{k}$-linear isomorphism when restricted to Λ_i for $0 \leq i \leq n$, since it sends the monomial basis elements $m_\lambda(\mathbf{x})$ to the monomial basis elements $m_\lambda(x_1, \ldots, x_n)$. Thus when one proves identities in $\Lambda_{\mathbf{k}}(x_1, \ldots, x_n)$ for all n, they are valid in Λ, that is, Λ is the inverse limit of the $\Lambda(x_1, \ldots, x_n)$ in the category of graded $\mathbf{k}$-algebras.

One can also define a comultiplication on Λ as follows. Note that when one decomposes the variables into two sets $(\mathbf{x},\mathbf{y}) = (x_1, x_2, \ldots, y_1, y_2, \ldots)$, one has a ring homomorphism

$$\begin{array}{rcl} R(\mathbf{x}) \otimes R(\mathbf{x}) & \longrightarrow & R(\mathbf{x},\mathbf{y}) \\ f(\mathbf{x}) \otimes g(\mathbf{x}) & \longmapsto & f(\mathbf{x})g(\mathbf{y}). \end{array}$$

This restricts to an isomorphism

$$\Lambda \otimes \Lambda = R(\mathbf{x})^{\mathfrak{S}_{(\infty)}} \otimes R(\mathbf{x})^{\mathfrak{S}_{(\infty)}} \longrightarrow R(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}} \tag{2.2}$$

where $\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}$ denotes permutations of (finite subsets of) the $\mathbf{x}$ and separate permutations of (finite subsets of) the $\mathbf{y}$, because $R(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}}$ has $\mathbb{Z}$-basis $\{m_\lambda(\mathbf{x}) m_\mu(\mathbf{y})\}_{\lambda,\mu \in \mathrm{Par}}$. As $\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}$ is a subgroup of the group $\mathfrak{S}_{(\infty)}$ acting on all of $(\mathbf{x},\mathbf{y})$, one gets an inclusion of rings

$$\Lambda(\mathbf{x},\mathbf{y}) = R(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)}} \hookrightarrow R(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}} \cong \Lambda \otimes \Lambda$$

where the last isomorphism is the inverse of the one in (2.2). This gives a comultiplication

$$\begin{array}{rcl} \Lambda = \Lambda(\mathbf{x}) & \stackrel{\Delta}{\longrightarrow} & \Lambda(\mathbf{x},\mathbf{y}) \hookrightarrow \Lambda \otimes \Lambda \\ f(\mathbf{x}) = f(x_1, x_2, \ldots) & \longmapsto & f(\mathbf{x},\mathbf{y}) = f(x_1, x_2, \ldots, y_1, y_2, \ldots). \end{array}$$

Example 2.5. One has

$$\begin{aligned} \Delta m_{(2,1)} &= m_{(2,1)}(x_1, x_2, \ldots, y_1, y_2, \ldots) \\ &= x_1^2 x_2 + x_1 x_2^2 + \cdots \\ &\quad + x_1^2 y_1 + x_1^2 y_2 + \cdots \\ &\quad + x_1 y_1^2 + x_1 y_2^2 + \cdots \\ &\quad + y_1^2 y_2 + y_1 y_2^2 + \cdots \\ &= m_{(2,1)}(\mathbf{x}) + m_{(2)}(\mathbf{x}) m_{(1)}(\mathbf{y}) + m_{(1)}(\mathbf{x}) m_{(2)}(\mathbf{y}) + m_{(2,1)}(\mathbf{y}) \\ &= m_{(2,1)} \otimes 1 + m_{(2)} \otimes m_{(1)} + m_{(1)} \otimes m_{(2)} + 1 \otimes m_{(2,1)}. \end{aligned}$$

This example generalizes easily to the following formula

$$\Delta m_\lambda = \sum_{\substack{(\mu,\nu): \\ \mu \sqcup \nu = \lambda}} m_\mu \otimes m_\nu. \tag{2.3}$$

in which $\mu \sqcup \nu$ is the partition obtained by taking the multiset union of the parts of μ and ν, and then reordering them to make them weakly decreasing.

Checking that Δ is coassociative amounts to checking that

$$(\Delta \otimes 1) \circ \Delta f = f(\mathbf{x},\mathbf{y},\mathbf{z}) = (1 \otimes \Delta) \circ \Delta f$$

inside $\Lambda(\mathbf{x},\mathbf{y},\mathbf{z})$ as a subring of $\Lambda \otimes \Lambda \otimes \Lambda$. The counit $\Lambda \stackrel{\epsilon}{\to} \mathbf{k}$ is defined in the usual fashion for graded connected coalgebras, namely ϵ annihilates $I = \bigoplus_{n>0} \Lambda_n$, and ϵ is the identity on $\Lambda_0 = \mathbf{k}$; alternatively ϵ sends a symmetric function $f(\mathbf{x})$ to its constant term $f(0,0,\ldots)$.

Note that Δ is an algebra morphism $\Lambda \to \Lambda \otimes \Lambda$ because it is a composition of maps which are all algebra morphisms. As the unit and counit axioms are easily checked, Λ becomes a graded connected $\mathbf{k}$-bialgebra of finite type, and hence also a Hopf algebra by Proposition 1.30. We will identify its antipode more explicitly in Section 2.4 below.

2.2. **Other Bases.** We introduce the usual other bases of Λ, and explain their significance later.

Definition 2.6. Define the families of *power sum symmetric functions* p_n, *elementary symmetric functions* e_n, and *complete homogeneous symmetric functions* h_n, for $n = 1, 2, 3, \ldots$ by

$$\begin{array}{lll} p_n & := x_1^n + x_2^n + \cdots & = m_{(n)} \\ e_n & := \sum_{i_1 < \cdots < i_n} x_{i_1} \cdots x_{i_n} & = m_{(1^n)} \\ h_n & := \sum_{i_1 \leq \cdots \leq i_n} x_{i_1} \cdots x_{i_n} & = \sum_{\lambda \in \mathrm{Par}_n} m_\lambda \end{array}$$

where $(1^n) = (1, 1, \ldots, 1)$, using a multiplicative notation $\lambda = (1^{m_1} 2^{m_2} \cdots)$ if the multiplicity of the part i in λ is m_i. By convention, also define $h_0 = e_0 = 1$, and $h_n = e_n = 0$ if $n < 0$. Extend these multiplicatively to partitions $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$:

$$\begin{aligned} p_\lambda &:= p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell} \\ e_\lambda &:= e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_\ell} \\ h_\lambda &:= h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_\ell} \end{aligned}$$

Also define the *Schur function*

$$s_\lambda := \sum_T \mathbf{x}^{\mathrm{cont}(T)} \tag{2.4}$$

where T runs through all *column-strict tableaux* of shape λ, that is, T is an assignment of entries in $\{1, 2, 3, \ldots\}$ to the cells of the *Ferrers diagram* for λ, weakly increasing left-to-right in rows, and strictly increasing top-to-bottom in columns. Here $\mathbf{x}^{\mathrm{cont}(T)} = \prod_i x_i^{|T^{-1}(i)|}$. For example,

$$T = \begin{array}{ccccc} 1 & 1 & 1 & 4 & 7 \\ 2 & 3 & 3 & & \\ 4 & 4 & 6 & & \\ 6 & 7 & & & \end{array}$$

is a column-strict tableau of shape $\lambda = (5, 3, 3, 2)$ with $\mathbf{x}^{\mathrm{cont}(T)} = x_1^3 x_2^1 x_3^2 x_4^3 x_5^0 x_6^2 x_7^2$.

Example 2.7. One has

$$\begin{aligned} m_{(1)} = p_{(1)} = e_{(1)} = h_{(1)} = s_{(1)} &= x_1 + x_2 + x_3 + \cdots \\ s_{(n)} &= h_n \\ s_{(1^n)} &= e_n \end{aligned}$$

Example 2.8. One has for $\lambda = (2, 1)$ that

$$\begin{aligned} p_{(2,1)} &= p_2 p_1 = (x_1^2 + x_2^2 + \cdots)(x_1 + x_2 + \cdots) \\ &= m_{(2,1)} + m_{(3)} \\ \\ e_{(2,1)} &= e_2 e_1 = (x_1 x_2 + x_1 x_3 + \cdots)(x_1 + x_2 + \cdots) \\ &= m_{(2,1)} + 3m_{(1,1,1)} \\ \\ h_{(2,1)} &= h_2 h_1 = (x_1^2 + x_2^2 + \cdots + x_1 x_2 + x_1 x_3 + \cdots)(x_1 + x_2 + \cdots) \\ &= m_{(3)} + 2m_{(2,1)} + 3m_{(1,1,1)} \end{aligned}$$

and

$$\begin{array}{cccccccccc} s_{(2,1)} & = x_1^2 x_2 & + x_1^2 x_3 & + x_1 x_2^2 & + x_1 x_3^2 & + x_1 x_2 x_3 & + x_1 x_2 x_3 & + x_1 x_2 x_4 & + \cdots \\ & 11 & 11 & 12 & 13 & 12 & 13 & 12 & \\ & 2 & 3 & 2 & 3 & 3 & 2 & 4 & \\ & = m_{(2,1)} + 2m_{(1,1,1)} & & & & & & & \end{array}$$

In fact, one has these transition matrices for $n = 3$ expressing elements in terms of the monomial basis m_λ:

$$\begin{array}{c} \\ m_{(3)} \\ m_{(2,1)} \\ m_{(1,1,1)} \end{array} \begin{array}{c} \begin{array}{ccc} p_{(3)} & p_{(2,1)} & p_{(1,1,1)} \end{array} \\ \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 6 \end{pmatrix} \end{array} \qquad \begin{array}{c} \\ m_{(3)} \\ m_{(2,1)} \\ m_{(1,1,1)} \end{array} \begin{array}{c} \begin{array}{ccc} e_{(3)} & e_{(2,1)} & e_{(1,1,1)} \end{array} \\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 3 & 6 \end{pmatrix} \end{array}$$

$$\begin{array}{c} \\ m_{(3)} \\ m_{(2,1)} \\ m_{(1,1,1)} \end{array} \begin{array}{c} \begin{array}{ccc} h_{(3)} & h_{(2,1)} & h_{(1,1,1)} \end{array} \\ \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix} \end{array} \qquad \begin{array}{c} \\ m_{(3)} \\ m_{(2,1)} \\ m_{(1,1,1)} \end{array} \begin{array}{c} \begin{array}{ccc} s_{(3)} & s_{(2,1)} & s_{(1,1,1)} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \end{array}$$

Our next goal is to show that $p_\lambda, e_\lambda, s_\lambda, h_\lambda$ all give bases for Λ. However at the moment it is not yet even clear that s_λ are symmetric!

Proposition 2.9. *Schur functions s_λ are symmetric, that is, they lie in Λ.*

Proof. It suffices to show s_λ is symmetric under swapping the variables x_i, x_{i+1}, by providing an involution ι on the set of all column-strict tableaux T of shape λ which switches the $\mathrm{cont}(T)$ for $(i, i+1)\,\mathrm{cont}(T)$. Restrict attention to the entries $i, i+1$ in T, which must look something like this:

$$\begin{array}{cccccccccccccc}
 & & & & & & & & i & i & i & i & i+1 & i+1 \\
 & i & i & i & i & i & i+1 & i+1 & i+1 & i+1 & i+1 & & & \\
i+1 & i+1 & i+1 & & & & & & & & & & &
\end{array}$$

One finds several vertically aligned pairs $\begin{matrix} i \\ i+1 \end{matrix}$. If one were to remove all such pairs, the remaining entries would be a sequence of rows, each looking like this:

$$\underbrace{i, i, \ldots, i}_{r \text{ occurrences}}, \underbrace{i+1, i+1, \ldots, i+1}_{s \text{ occurrences}} \tag{2.5}$$

An involution due to Bender and Knuth tells us to leave fixed all the vertically aligned pairs $\begin{matrix} i \\ i+1 \end{matrix}$, but change each sequence as in (2.5) to this:

$$\underbrace{i, i, \ldots, i}_{s \text{ occurrences}}, \underbrace{i+1, i+1, \ldots, i+1}_{r \text{ occurrences}}$$

For example, the above configuration in T would change to

$$\begin{array}{cccccccccccccc}
 & & & & & & & & i & i & i & i & i & i+1 \\
 & i & i & i & i & i+1 & i+1 & i+1 & i+1 & i+1 & i+1 & & & \\
i & i+1 & i+1 & & & & & & & & & & &
\end{array}$$

It is easily checked that this map is an involution, and that it has the effect of swapping $(i, i+1)$ in $\mathrm{cont}(T)$. $\square$

Remark 2.10. The symmetry of Schur functions allows one to reformulate them via column-strict tableaux defined with respect to *any* total ordering $\mathcal{L}$ on the positive integers, rather than the usual $1 < 2 < 3 < \cdots$. For example, one can use the *reverse order*[3] $\cdots < 3 < 2 < 1$, or even more exotic orders, such as

$$1 < 3 < 5 < 7 < \cdots < 2 < 4 < 6 < 8 < \cdots.$$

Say that an assignment T of entries in $\{1, 2, 3, \ldots\}$ to the cells of the Ferrers diagram of λ is an $\mathcal{L}$-*column-strict tableau* if it is weakly $\mathcal{L}$-increasing left-to-right in rows, and strictly $\mathcal{L}$-increasing top-to-bottom in columns.

Proposition 2.11. *For any total order $\mathcal{L}$ on the positive integers,*

$$s_\lambda = \sum_T \mathbf{x}^{\mathrm{cont}(T)} \tag{2.6}$$

as T runs through all $\mathcal{L}$-column-strict tableaux of shape λ.

Proof. Given a weak composition $\alpha = (\alpha_1, \alpha_2, \ldots)$ with $\alpha_{n+1} = \alpha_{n+2} = \cdots = 0$, assume that the integers $1, 2, \ldots, n$ are totally ordered by $\mathcal{L}$ as $w(1) <_{\mathcal{L}} \cdots <_{\mathcal{L}} w(n)$ for some w in $\mathfrak{S}_n$. Then the coefficient of $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ on the right side of (2.6) is the same as the coefficient of $\mathbf{x}^{w^{-1}(\alpha)}$ on the right side of (2.4) defining s_λ, which by symmetry of s_λ is the same as the coefficient of $\mathbf{x}^\alpha$ on the right side of (2.4). $\square$

It is now not hard to show that $p_\lambda, e_\lambda, s_\lambda$ give bases by a triangularity argument. For this purpose, let us introduce a useful partial order on partitions.

Definition 2.12. The *dominance* or *majorization* partial order on Par_n, written $\lambda \triangleright \mu$, is defined by

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k \quad \text{for } k = 1, 2, \ldots, n.$$

---

[3]This reverse order is what one uses when one defines a Schur function as a generating function for *reverse semistandard tableaux* or *column-strict plane partitions*; see Stanley [72, Proposition 7.10.4].

Definition 2.13. For a partition λ, its *conjugate* or *transpose* partition λ^t, is the one whose Ferrers diagram is obtained from that of λ by exchanging rows for columns. Alternatively, one has this formula for its i^{th} part:

$$(\lambda^t)_i := |\{j : \lambda_j \geq i\}|.$$

It is an interesting (and useful) exercise to check that $\lambda \triangleright \mu$ if and only if $\mu^t \triangleright \lambda^t$.

Proposition 2.14. *The sets $\{e_\lambda\}, \{s_\lambda\}$ as λ runs through all partitions give $\mathbf{k}$-bases for $\Lambda_{\mathbf{k}}$ for any field $\mathbf{k}$ or $\mathbf{k} = \mathbb{Z}$. The same holds for $\{p_\lambda\}$ when $\mathbf{k}$ is a field of characteristic zero.*

Proof. One can restrict attention to each homogeneous component Λ_n and partitions λ of n. We check that in each case, the proposed basis expands triangularly in the $\{m_\lambda\}$ with some choice of orderings on Par_n indexing the rows and columns, as illustrated in Example 2.8.

One has $s_\lambda = \sum_\mu K_{\lambda,\mu} m_\mu$ where the coefficient $K_{\lambda,\mu}$ is the *Kostka number* counting the column-strict tableaux T of shape λ having $\mathrm{cont}(T) = \mu$; this follows because both sides are symmetric functions, and $K_{\lambda,\mu}$ is the coefficient of $\mathbf{x}^\mu$ on both sides. Since for each positive integer k, the entries $1, 2, \ldots, k$ in T must all lie within the first k rows of λ, one has that $\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k$, that is, $\lambda \triangleright \mu$, so $s_\lambda = \sum_{\mu : \lambda \triangleright \mu} K_{\lambda,\mu} m_\mu$. One can also check that $K_{\lambda,\lambda} = 1$, so this expansion is unitriangular with appropriate ordering of rows and columns.

One has $e_\lambda = \sum_\mu a_{\lambda,\mu} m_\mu$ where $a_{\lambda,\mu}$ counts matrices with $\{0,1\}$ entries having row sum λ and column-sum μ: when one expands $e_{\lambda_1} e_{\lambda_2} \cdots$, choosing the monomial $x_{j_1} \ldots x_{j_{\lambda_i}}$ in the e_{λ_i} factor corresponds to putting 1's in the i^{th} row and columns $j_1, \ldots, j_{\lambda_i}$ of the $\{0,1\}$-matrix. It is not hard to check[4] that $a_{\lambda,\mu}$ vanishes unless $\lambda^t \triangleright \mu$. One can also check that $a_{\lambda,\lambda^t} = 1$, so this expansion is again unitriangular with appropriate ordering of rows and columns.

Assume now that $\mathbf{k}$ is a field of characteristic 0. One has $p_\lambda = \sum_\mu b_{\lambda,\mu} m_\mu$ where $b_{\lambda,\mu}$ counts the ways to partition the nonzero parts $\lambda_1, \ldots, \lambda_\ell$ into blocks such that the sums of the blocks give μ; more formally, $b_{\lambda,\mu}$ is the number of maps $\varphi : \{1, 2, \ldots, \ell\} \to \{1, 2, 3, \ldots\}$ having $\mu_j = \sum_{i : \varphi(i) = j} \lambda_i$ for $j = 1, 2, \ldots$. Again it is not hard to check that $b_{\lambda,\mu}$ vanishes unless $\lambda \triangleleft \mu$, and hence this expansion is triangular, for appropriate ordering of rows and columns (but not unitriangular, as $b_{\lambda,\lambda} \neq 1$ in general). The diagonal entries $b_{\lambda,\lambda}$ are positive integers and thus invertible in $\mathbf{k}$, so $\{p_\lambda\}$ is a basis. $\square$

Remark 2.15. When $\mathbf{k}$ is not a field of characteristic 0, the family $\{p_\lambda\}$ is not (in general) a basis of $\Lambda_{\mathbf{k}}$; for instance, $e_2 = \frac{1}{2}(p_{11} - p_2) \in \Lambda_{\mathbb{Q}}$ is not in the $\mathbb{Z}$-span of this family. However, if we define $b_{\lambda,\mu}$ as in the above proof, then the $\mathbb{Z}$-linear span of all p_λ equals the $\mathbb{Z}$-linear span of all $b_{\lambda,\lambda} m_\lambda$. Indeed, if $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$ with $k = \ell(\mu)$, then $b_{\mu,\mu}$ is the size of the subgroup of $\mathfrak{S}_k$ consisting of all permutations $\sigma \in \mathfrak{S}_k$ having each i satisfy $\mu_{\sigma(i)} = \mu_i$. As a consequence, $b_{\mu,\mu} \mid b_{\lambda,\mu}$ for every partition μ of the same size as λ (because this group acts[5] freely on the set which is enumerated by $b_{\lambda,\mu}$), so that by rescaling m_λ with the factor $b_{\lambda,\lambda}$ we obtain a unitriangular integer transition matrix.

2.3. **Comultiplications.** Thinking about comultiplication $\Lambda \xrightarrow{\Delta} \Lambda \otimes \Lambda$ on Schur functions forces us to immediately confront the following.

Definition 2.16. For partitions μ, λ say that $\mu \subseteq \lambda$ if $\mu_i \leq \lambda_i$ for $i = 1, 2, \ldots$, so the Ferrers diagram for μ is a subset of the cells for the Ferrers diagram of λ. In this case, define the *skew (Ferrers) diagram* λ/μ to be their set difference.

Then define the *skew Schur function* $s_{\lambda/\mu}(\mathbf{x})$ to be the sum $s_{\lambda/\mu} := \sum_T \mathbf{x}^{\mathrm{cont}(T)}$, where the sum ranges over all *column-strict tableaux* T of shape λ/μ, that is, assignments of a value in $\{1, 2, 3, \ldots\}$ to each cell of λ/μ, weakly increasing left-to-right in rows, and strictly increasing top-to-bottom in columns.

Example 2.17.

$$T = \begin{array}{ccccc} \cdot & \cdot & \cdot & 2 & 5 \\ \cdot & 1 & 1 & & \\ 2 & 2 & 4 & & \\ 4 & 5 & & & \end{array}$$

is a column-strict tableau of shape $\lambda/\mu = (5, 3, 3, 2)/(3, 1, 0, 0)$ and it has $\mathbf{x}^{\mathrm{cont}(T)} = x_1^2 x_2^3 x_3^0 x_4^2 x_5^2$.

[4]This is the easy implication in the *Gale-Ryser Theorem*.

[5]Specifically, an element σ of the group takes $\varphi : \{1, 2, \ldots, \ell\} \to \{1, 2, 3, \ldots\}$ to $\sigma \circ \varphi$.

Proposition 2.18. *The comultiplication $\Lambda \stackrel{\Delta}{\to} \Lambda \otimes \Lambda$ has the following effect on the symmetric functions discussed so far*[6]*:*

(i) $\Delta p_n = 1 \otimes p_n + p_n \otimes 1$, *that is, the power sums p_n are primitive.*
(ii) $\Delta e_n = \sum_{i+j=n} e_i \otimes e_j$
(iii) $\Delta h_n = \sum_{i+j=n} h_i \otimes h_j$
(iv) $\Delta s_\lambda = \sum_{\mu \subseteq \lambda} s_\mu \otimes s_{\lambda/\mu}$

Proof. Recall that Δ sends $f(\mathbf{x}) \mapsto f(\mathbf{x}, \mathbf{y})$, and one can easily check that

(i) $p_n(\mathbf{x}, \mathbf{y}) = \sum_i x_i^n + \sum_i y_i^n = p_n(\mathbf{x}) \cdot 1 + 1 \cdot p_n(\mathbf{y})$
(ii) $e_n(\mathbf{x}, \mathbf{y}) = \sum_{i+j=n} e_i(\mathbf{x}) e_j(\mathbf{y})$
(iii) $h_n(\mathbf{x}, \mathbf{y}) = \sum_{i+j=n} h_i(\mathbf{x}) h_j(\mathbf{y})$

For assertion (iv), note that by Remark 2.6, one has

$$s_\lambda(\mathbf{x}, \mathbf{y}) = \sum_T (\mathbf{x}, \mathbf{y})^{\mathrm{cont}(T)}$$

where the sum is over column-strict tableaux T of shape λ having entries in the linearly ordered alphabet

$$x_1 < x_2 < \cdots < y_1 < y_2 < \cdots . \tag{2.7}$$

For example,

$$T = \begin{matrix} x_1 & x_1 & x_1 & y_2 & y_5 \\ x_2 & y_1 & y_1 & & \\ y_2 & y_2 & y_4 & & \\ y_4 & y_5 & & & \end{matrix}$$

is such a tableau of shape $\lambda = (5, 3, 3, 2)$. Note that the restriction of T to the alphabet $\mathbf{x}$ gives a column-strict tableau $T_{\mathbf{x}}$ of some shape $\mu \subseteq \lambda$, and the restriction of T to the alphabet $\mathbf{y}$ gives a column-strict tableau $T_{\mathbf{y}}$ of shape λ/μ (e.g. for T in the example above, the tableau $T_{\mathbf{y}}$ appeared in Example 2.17). Consequently, one has

$$\begin{aligned} s_\lambda(\mathbf{x}, \mathbf{y}) &= \sum_T \mathbf{x}^{\mathrm{cont}(T_{\mathbf{x}})} \cdot \mathbf{y}^{\mathrm{cont}(T_{\mathbf{y}})} \\ &= \sum_{\mu \subseteq \lambda} \left(\sum_{T_{\mathbf{x}}} \mathbf{x}^{\mathrm{cont}(T_{\mathbf{x}})} \right) \left(\sum_{T_{\mathbf{y}}} \mathbf{y}^{\mathrm{cont}(T_{\mathbf{y}})} \right) = \sum_{\mu \subseteq \lambda} s_\mu(\mathbf{x}) s_{\lambda/\mu}(\mathbf{y}). \end{aligned}$$

$\square$

2.4. **The antipode, the involution ω, and algebra generators.** Since Λ is a graded connected $\mathbf{k}$-bialgebra, it will have an antipode $\Lambda \stackrel{S}{\to} \Lambda$ making it a Hopf algebra by Proposition 1.30. However, several issues will be resolved by identifying S more explicitly now.

Proposition 2.19. *Each of $\{e_n\}_{n=1,2,\ldots}$, $\{h_n\}_{n=1,2,\ldots}$ are algebraically independent, and generate $\Lambda_{\mathbf{k}}$ as a polynomial algebra for any field $\mathbf{k}$ or $\mathbf{k} = \mathbb{Z}$. The same holds for $\{p_n\}_{n=1,2,\ldots}$ when $\mathbf{k}$ is a field of characteristic zero.*

Furthermore, the antipode S acts as follows:

(i) $S(p_n) = -p_n$
(ii) $S(e_n) = (-1)^n h_n$
(iii) $S(h_n) = (-1)^n e_n$

[6] The abbreviated summation indexing $\sum_{i+j=n} t_{i,j}$ used here is intended to mean

$$\sum_{\substack{(i,j): \\ 0 \le i,j \le n, \\ i+j=n}} t_{i,j}.$$

Proof. The assertions that $\{e_n\}, \{p_n\}$ are algebraically independent and generate Λ are equivalent to Proposition 2.14 asserting $\{e_\lambda\}, \{p_\lambda\}$ give bases for Λ. The assertion $S(p_n) = -p_n$ follows from Proposition 1.31 since p_n is primitive by Proposition 2.18(i).

For the remaining assertions, start with the easy generating function identities

$$H(t) := \prod_{i=1}^{\infty}(1 - x_i t)^{-1} = 1 + h_1(\mathbf{x})t + h_2(\mathbf{x})t^2 + \cdots = \sum_{n \geq 0} h_n(\mathbf{x})t^n \tag{2.8}$$

$$E(t) := \prod_{i=1}^{\infty}(1 + x_i t) = 1 + e_1(\mathbf{x})t + e_2(\mathbf{x})t^2 + \cdots = \sum_{n \geq 0} e_n(\mathbf{x})t^n \tag{2.9}$$

which shows that

$$1 = E(-t)H(t) = \left(\sum_{n \geq 0} e_n(\mathbf{x})(-t)^n\right)\left(\sum_{n \geq 0} h_n(\mathbf{x})t^n\right) \tag{2.10}$$

and hence, equating coefficients of powers of t, that for $n = 0, 1, 2, \ldots$ one has

$$\sum_{i+j=n} (-1)^i e_i h_j = \delta_{0,n}. \tag{2.11}$$

This lets one recursively express the e_n in terms of h_n and vice-versa:

$$\begin{aligned} e_0 &:= 1 =: h_0 \\ e_n &= e_{n-1}h_1 - e_{n-2}h_2 + e_{n-3}h_3 - \cdots \\ h_n &= h_{n-1}e_1 - h_{n-2}e_2 + h_{n-3}e_3 - \cdots \end{aligned} \tag{2.12}$$

for $n = 1, 2, \ldots$ Thus if one uses the algebraic independence of the generators $\{e_n\}$ for Λ to define an algebra endomorphism as follows

$$\begin{array}{ccc} \Lambda & \overset{\omega}{\to} & \Lambda \\ e_n & \longmapsto & h_n, \end{array} \tag{2.13}$$

then the identical form of the two recursions in (2.12) shows that ω also sends $h_n \mapsto e_n$. Therefore ω is an involutive automorphism of Λ, and the $\{h_n\}$ are another algebraically independent generating set for Λ.

For the assertion about the antipode S applied to e_n or h_n, note that the coproduct formulas for e_n, h_n in Proposition 2.18(ii),(iii) show that the defining relations for their antipodes (1.12) will in this case be

$$\sum_{i+j=n} S(e_i)e_j = \delta_{0,n} = \sum_{i+j=n} e_i S(e_j)$$
$$\sum_{i+j=n} S(h_i)h_j = \delta_{0,n} = \sum_{i+j=n} h_i S(h_j)$$

because $u\epsilon(e_n) = u\epsilon(h_n) = \delta_{0,n}$. Comparing these to (2.11), one concludes via induction on n that $S(e_n) = (-1)^n h_n$ and $S(h_n) = (-1)^n e_n$. ∎

Proposition 2.19 shows that the antipode S on Λ is, up to sign, the same as the fundamental involution ω: one has

$$S(f) = (-1)^n \omega(f) \text{ for } f \in \Lambda_n \tag{2.14}$$

since this formula holds for all elements of the generating set $\{e_n\}$ (or $\{h_n\}$).

Remark 2.20. Up to now we have not yet derived how the involution ω and the antipode S act on (skew) Schur functions, which is quite beautiful:

$$\begin{aligned} \omega(s_{\lambda/\mu}) &= s_{\lambda^t/\mu^t} \\ S(s_{\lambda/\mu}) &= (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t} \end{aligned} \tag{2.15}$$

where recall that λ^t is the transpose or conjugate partition to λ, and $|\lambda/\mu|$ is the number of squares in the skew diagram λ/μ, that is, $|\lambda/\mu| = n - k$ if λ, μ lie in $\mathrm{Par}_n, \mathrm{Par}_k$ respectively.

We will deduce this later in two ways, once as an exercise using skewing operators in Exercise 2.49, and for the second time from the action of the antipode in QSym on P-partition enumerators in Corollary 5.24.

However, one could also deduce it immediately from our knowledge of the action of ω and S on e_n, h_n, if we were to prove the following famous *Jacobi-Trudi* and *dual Jacobi-Trudi* formulas.

Theorem 2.21. *Skew Schur functions are the following polynomials in* $\{h_n\}, \{e_n\}$*:*

$$s_{\lambda/\mu} = \det(h_{\lambda_i - \mu_j - i + j})_{i,j=1,2,\ldots,\ell}$$
$$s_{\lambda^t/\mu^t} = \det(e_{\lambda_i - \mu_j - i + j})_{i,j=1,2,\ldots,\ell}$$

if λ *has at most* ℓ *nonzero parts.*

Since we appear not to need these formulas in the sequel, we will omit the proofs. Various proofs are well-explained in [49, §I.5], [62, §4.5], [72, §7.16]. An elegant treatment of Schur polynomials taking the Jacobi-Trudi formula as the *definition* of s_λ is given by Tamvakis [78].

2.5. **Cauchy product, Hall inner product, self-duality.** The Schur functions, although a bit unmotivated right now, have special properties with regard to the Hopf structure. One property is intimately connected with the following *Cauchy identity*.

Theorem 2.22. *One has the following Schur function expansion for the Cauchy product/Cauchy kernel*

$$\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \mathrm{Par}} s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}). \tag{2.16}$$

Remark 2.23. Some readers may be bothered by the ambient ring in which this expansion takes place, which is a certain completion of $R(\mathbf{x}) \otimes R(\mathbf{y})$. One simple way to understand it is to replace each x_i by $x_i t$, and write the equivalent identity in the power series ring $R(\mathbf{x}) \otimes R(\mathbf{y})[[t]]$

$$\prod_{i,j=1}^{\infty} (1 - t x_i y_j)^{-1} = \sum_{\lambda \in \mathrm{Par}} t^{|\lambda|} s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}). \tag{2.17}$$

(Recall that $|\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_\ell$ for any partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$.)

Proof. We follow the standard combinatorial proof (see [62, §4.8],[72, §7.11,7.12]), which rewrites the left and right sides of (2.17), and then compares them with the *Robinson-Schensted-Knuth* (RSK) bijection. On the left side, expanding out each geometric series

$$(1 - t x_i y_j)^{-1} = 1 + t x_i y_j + (t x_i y_j)^2 + (t x_i y_j)^3 + \cdots$$

and thinking of $(x_i y_j)^m$ as m occurrences of a *biletter* $\binom{i}{j}$, one can think of the left side as a sum over multisets of biletters $\binom{i_1}{j_1}, \ldots, \binom{i_\ell}{j_\ell}$. Order the biletters in such a multiset in a lexicographic order $\leq_{lex}$ that first checks if $i_1 \leq i_2$ and then if $i_1 = i_2$ checks if $j_1 \leq j_2$. Defining a *biword* to be an array $\binom{\mathbf{i}}{\mathbf{j}} = \binom{i_1 \cdots i_\ell}{j_1 \cdots j_\ell}$ in which the biletters are ordered $\binom{i_1}{j_1} \leq_{lex} \cdots \leq_{lex} \binom{i_\ell}{j_\ell}$, then the left side of (2.17) is the sum $\sum t^\ell \mathbf{x}^{\mathrm{cont}(\mathbf{i})} \mathbf{y}^{\mathrm{cont}(\mathbf{j})}$ over all biwords $\binom{\mathbf{i}}{\mathbf{j}}$, where ℓ stands for the number of biletters in the biword. On the right side, expanding out the Schur functions as sums of tableaux gives $\sum_{(P,Q)} t^\ell \mathbf{x}^{\mathrm{cont}(Q)} \mathbf{y}^{\mathrm{cont}(P)}$ in which the sum is over all ordered pairs (P, Q) of column-strict tableaux *having the same shape*, with ℓ cells.

The Robinson-Schensted-Knuth insertion algorithm gives us a bijection between the biwords $\binom{\mathbf{i}}{\mathbf{j}}$ and the tableau pairs (P, Q), with the property that

$$\mathrm{cont}(\mathbf{i}) = \mathrm{cont}(Q),$$
$$\mathrm{cont}(\mathbf{j}) = \mathrm{cont}(P).$$

Starting with the pair $(P_0, Q_0) = (\varnothing, \varnothing)$ and $m = 0$, it inserts one at a time the next biletter $\binom{i_{m+1}}{j_{m+1}}$ of the biword into the pair of tableaux (P_m, Q_m) already built; see Example 2.24 below. The bottom letter j_{m+1} tries to insert itself into the first row of P_m by either bumping out the leftmost letter in the first row strictly larger than j_{m+1}, or else placing itself at the right end of the row if no such larger letter exists. If a letter was bumped from the first row, it follows the same rules to insert itself into the second row, and so on. At the end of the bumping, the tableau P_{m+1} created has an extra corner cell not present in P_m, and one creates Q_{m+1} from Q_m by adding the top letter i_{m+1} of $\binom{i_{m+1}}{j_{m+1}}$ to Q_m in this extra corner cell location. After all of the biletters have been inserted, the result is $(P_\ell, Q_\ell) =: (P, Q)$.

Example 2.24. The term in the expansion of the left side of (2.16) corresponding to

$$(x_1y_2)^1(x_1y_4)^1(x_2y_1)^1(x_4y_1)^1(x_4y_3)^2(x_5y_2)^1$$

is the biword $\binom{\mathbf{i}}{\mathbf{j}} = \binom{1124445}{2411332}$, whose RSK insertion goes as follows:

$$\begin{array}{rcl@{\qquad\qquad}rcl}
P_0 & = & \varnothing & Q_0 & = & \varnothing \\[2ex]
P_1 & = & 2 & Q_1 & = & 1 \\[2ex]
P_2 & = & \begin{array}{ll} 2 & 4 \end{array} & Q_2 & = & \begin{array}{ll} 1 & 1 \end{array} \\[2ex]
P_3 & = & \begin{array}{ll} 1 & 4 \\ 2 & \end{array} & Q_3 & = & \begin{array}{ll} 1 & 1 \\ 2 & \end{array} \\[2ex]
P_4 & = & \begin{array}{ll} 1 & 1 \\ 2 & 4 \end{array} & Q_4 & = & \begin{array}{ll} 1 & 1 \\ 2 & 4 \end{array} \\[2ex]
P_5 & = & \begin{array}{lll} 1 & 1 & 3 \\ 2 & 4 & \end{array} & Q_5 & = & \begin{array}{lll} 1 & 1 & 4 \\ 2 & 4 & \end{array} \\[2ex]
P_6 & = & \begin{array}{llll} 1 & 1 & 3 & 3 \\ 2 & 4 & & \end{array} & Q_6 & = & \begin{array}{llll} 1 & 1 & 4 & 4 \\ 2 & 4 & & \end{array} \\[2ex]
P := P_7 & = & \begin{array}{llll} 1 & 1 & 2 & 3 \\ 2 & 3 & & \\ 4 & & & \end{array} & Q := Q_7 & = & \begin{array}{llll} 1 & 1 & 4 & 4 \\ 2 & 4 & & \\ 5 & & & \end{array}
\end{array}$$

It requires some thought, but is not too hard, to see that the bumping rule maintains the property that P_m is a column-strict tableau of some Ferrers shape throughout. It should be clear that (P_m, Q_m) have the same shape at each stage. Also, the construction of Q_m shows that it is at least weakly increasing in rows and weakly increasing in columns throughout. What is perhaps least clear is that Q_m remains strictly increasing down columns. That is, when one has a string of equal letters on top $i_m = i_{m+1} = \cdots = i_{m+r}$, so that on bottom one bumps in $j_m \leq j_{m+1} \leq \cdots \leq j_{m+r}$, one needs to know that the new cells form a *horizontal strip*, that is, no two of them lie in the same column. This follows once one observes that when one bumps in two letters $j \leq j'$, with j bumped in first, the *bumping path* for j' (the cells into which letters bump) stays strictly to the right, within each row, of the bumping path for j. As an example, when $j_{m+1} = 1$ is inserted into the tableau P_m shown below, the result P_{m+1} is shown with bumping path entries underlined:

$$P_m = \begin{array}{lllll} 1 & 1 & 2 & 2 & 3 \\ 2 & 2 & 4 & 4 & \\ 3 & 4 & 5 & & \\ 4 & 6 & 6 & & \end{array} \qquad \overset{\text{insert}}{\underset{j_{m+1}=1}{\longmapsto}} \qquad P_{m+1} = \begin{array}{lllll} 1 & 1 & \underline{1} & 2 & 3 \\ 2 & 2 & \underline{2} & 4 & \\ 3 & 4 & \underline{4} & & \\ 4 & \underline{5} & 6 & & \\ \underline{6} & & & & \end{array}$$

To see that the map is a bijection, we show how to recover $\binom{\mathbf{i}}{\mathbf{j}}$ from (P, Q). This is done by *reverse bumping* from (P_{m+1}, Q_{m+1}) to recover both the biletter $\binom{i_{m+1}}{j_{m+1}}$ and the tableaux (P_m, Q_m), as follows. Firstly, i_{m+1} is the maximum entry of Q_{m+1}, and Q_m is obtained by removing the rightmost occurrence of this letter i_{m+1} from Q_{m+1}. [7] To produce P_m and j_{m+1}, find the position of the rightmost occurrence of i_{m+1} in Q_{m+1}, and start *reverse bumping* in P_{m+1} from the entry in this same position, where reverse bumping an entry means inserting it into one row higher by having it bump out the rightmost entry which is strictly smaller. The entry bumped out of the first row is j_{m+1}, and the resulting tableau is P_m.

Finally, to see that the RSK map is surjective, one needs to show that the reverse bumping procedure can be applied to any pair (P, Q) of column-strict tableaux of the same shape, and will result in a (lexicographically ordered) biword $\binom{\mathbf{i}}{\mathbf{j}}$. We leave this verification to the reader. $\square$

[7] It necessarily has to be the rightmost occurrence, since (by our previous observation on bumping paths) the cell into which i_{m+1} was filled at the step from Q_m to Q_{m+1} lies further right than any existing cell of Q_m containing the letter i_{m+1}.

Corollary 2.25. *In the Schur function basis $\{s_\lambda\}$ for Λ, the structure constants for multiplication and comultiplication are the same, that is, if one defines $c^\lambda_{\mu,\nu}, \hat{c}^\lambda_{\mu,\nu}$ via the unique expansions*

$$\begin{aligned} s_\mu s_\nu &= \sum_\lambda c^\lambda_{\mu,\nu} s_\lambda \\ \Delta(s_\lambda) &= \sum_{\mu,\nu} \hat{c}^\lambda_{\mu,\nu} s_\mu \otimes s_\nu \end{aligned} \tag{2.18}$$

then $c^\lambda_{\mu,\nu} = \hat{c}^\lambda_{\mu,\nu}$.

Proof. The identity (2.16) lets one interpret both $c^\lambda_{\mu,\nu}, \hat{c}^\lambda_{\mu,\nu}$ as the coefficient of $s_\mu(\mathbf{x})s_\nu(\mathbf{y})s_\lambda(\mathbf{z})$ in the product

$$\begin{aligned} \prod_{i,j=1}^{\infty}(1-x_iz_j)^{-1}\prod_{i,j=1}^{\infty}(1-y_iz_j)^{-1} &= \left(\sum_\mu s_\mu(\mathbf{x})s_\mu(\mathbf{z})\right)\left(\sum_\nu s_\nu(\mathbf{y})s_\nu(\mathbf{z})\right) \\ &= \sum_{\mu,\nu} s_\mu(\mathbf{x})s_\nu(\mathbf{y})\cdot s_\mu(\mathbf{z})s_\nu(\mathbf{z}) \\ &= \sum_{\mu,\nu} s_\mu(\mathbf{x})s_\nu(\mathbf{y})\left(\sum_\lambda c^\lambda_{\mu,\nu} s_\lambda(\mathbf{z})\right) \end{aligned}$$

since, regarding $x_1, x_2, \ldots, y_1, y_2, \ldots$ as lying in a single variable set $(\mathbf{x},\mathbf{y})$, separate from the variables $\mathbf{z}$, the Cauchy identity (2.16) expands the same product as

$$\begin{aligned} \prod_{i,j=1}^{\infty}(1-x_iz_j)^{-1}\prod_{i,j=1}^{\infty}(1-y_iz_j)^{-1} &= \sum_\lambda s_\lambda(\mathbf{x},\mathbf{y})s_\lambda(\mathbf{z}) \\ &= \sum_\lambda\left(\sum_{\mu,\nu}\hat{c}^\lambda_{\mu,\nu}s_\mu(\mathbf{x})s_\nu(\mathbf{y})\right)s_\lambda(\mathbf{z}). \end{aligned}$$

□

Definition 2.26. The coefficients $c^\lambda_{\mu,\nu} = \hat{c}^\lambda_{\mu,\nu}$ appearing in the expansions (2.18) are called *Littlewood-Richardson coefficients.*

Remark 2.27*.* Noting on one hand the expansion

$$s_\lambda(\mathbf{x},\mathbf{y}) = \Delta(s_\lambda) = \sum_{\mu,\nu} c^\lambda_{\mu,\nu} s_\mu(\mathbf{x})s_\nu(\mathbf{y})$$

and on the other hand

$$s_\lambda(\mathbf{x},\mathbf{y}) = \sum_{\mu\subseteq\lambda} s_\mu(\mathbf{x})s_{\lambda/\mu}(\mathbf{y})$$

one concludes another standard interpretation for $c^\lambda_{\mu,\nu}$:

$$s_{\lambda/\mu} = \sum_\nu c^\lambda_{\mu,\nu} s_\nu.$$

In particular, $c^\lambda_{\mu,\nu}$ vanishes unless $\mu,\nu \subseteq \lambda$ and $|\mu|+|\nu| = |\lambda|$. Note also that $c^\lambda_{\mu,\nu} = c^\lambda_{\nu,\mu}$. We will interpret $c^\lambda_{\mu,\nu}$ combinatorially in Section 2.6.

Definition 2.28. Define the *Hall inner product* on Λ to be the $\mathbf{k}$-bilinear form $(\cdot,\cdot)$ which makes $\{s_\lambda\}$ an orthonormal basis, that is, $(s_\lambda, s_\nu) = \delta_{\lambda,\nu}$.

Corollary 2.29. *The isomorphism $\Lambda^o \cong \Lambda$ induced by the Hall inner product is an isomorphism of Hopf algebras.*

Proof. We have seen that the orthonormal basis $\{s_\lambda\}$ of Schur functions is **self-dual**, in the sense that its multiplication and comultiplication structure constants are the same. Thus the isomorphism $\Lambda^o \cong \Lambda$ induced by the Hall inner product is an isomorphism of bialgebras, and hence also a Hopf algebra isomorphism by Proposition 1.35(c). □

We next identify two other dual pairs of bases, by expanding the Cauchy product in two other ways.

Proposition 2.30. *One can also expand*

$$\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1}=\sum_{\lambda\in\mathrm{Par}}h_\lambda(\mathbf{x})m_\lambda(\mathbf{y})=\sum_{\lambda\in\mathrm{Par}}z_\lambda^{-1}p_\lambda(\mathbf{x})p_\lambda(\mathbf{y}) \tag{2.19}$$

where $z_\lambda := m_1!\cdot 1^{m_1}\cdot m_2!\cdot 2^{m_2}\cdots$ *if* $\lambda=(1^{m_1},2^{m_2},\ldots)$ *with multiplicity* m_i *for the part* i.

Remark 2.31. It is relevant later (and explains the notation) that z_λ is the size of the $\mathfrak{S}_n$-centralizer subgroup for a permutation having cycle type λ with $|\lambda|=n$.

Proof. For the first expansion, note that (2.8) shows

$$\begin{aligned}\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1}&=\prod_{j=1}^{\infty}\sum_{n\geq 0}h_n(\mathbf{x})y_j^n\\&=\sum_{\substack{\text{weak}\\ \text{compositions}\\ (n_1,n_2,\ldots)}}(h_{n_1}(\mathbf{x})h_{n_2}(\mathbf{x})\cdots)(y_1^{n_1}y_2^{n_2}\cdots)\\&=\sum_{\lambda\in\mathrm{Par}}h_\lambda(\mathbf{x})m_\lambda(\mathbf{y}).\end{aligned}$$

For the second expansion (and for later use in the proof of Theorem 4.35) note that

$$\log H(t)=\log\prod_{i=1}^{\infty}(1-x_it)^{-1}=\sum_{i=1}^{\infty}-\log(1-x_it)=\sum_{i=1}^{\infty}\sum_{m=1}^{\infty}\frac{(x_it)^m}{m}=\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})t^m$$

so that taking $\frac{d}{dt}$ then shows that

$$P(t):=\sum_{m\geq 0}p_{m+1}t^m=\frac{H'(t)}{H(t)}=H'(t)E(-t). \tag{2.20}$$

A similar calculation shows that

$$\log\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1}=\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})p_m(\mathbf{y})$$

and hence

$$\begin{aligned}\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1}&=\exp\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})p_m(\mathbf{y})\right)=\sum_{k=0}^{\infty}\frac{1}{k!}\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})p_m(\mathbf{y})\right)^k\\&=\sum_{k=0}^{\infty}\frac{1}{k!}\sum_{\substack{(m_1,m_2,\ldots)\in\mathbb{N}^\infty:\\ m_1+m_2+\cdots=k}}\binom{k}{m_1,m_2,\ldots}\left(\frac{p_1(\mathbf{x})p_1(\mathbf{y})}{1}\right)^{m_1}\left(\frac{p_2(\mathbf{x})p_2(\mathbf{y})}{2}\right)^{m_2}\cdots\\&=\sum_{\substack{\text{weak}\\ \text{compositions}\\ (m_1,m_2,\ldots)}}\frac{(p_1(\mathbf{x})p_1(\mathbf{y}))^{m_1}}{1^{m_1}m_1!}\cdot\frac{(p_2(\mathbf{x})p_2(\mathbf{y}))^{m_2}}{2^{m_2}m_2!}\cdots=\sum_{\lambda\in\mathrm{Par}}\frac{p_\lambda(\mathbf{x})p_\lambda(\mathbf{y})}{z_\lambda}.\end{aligned}$$

∎

Corollary 2.32. *With respect to the Hall inner product on* Λ, *one also has dual bases* $\{h_\lambda\}$ *and* $\{m_\lambda\}$, *as well as another orthonormal basis* $\{\frac{p_\lambda}{\sqrt{z_\lambda}}\}$.

Proof. Since (2.16) and (2.19) showed

$$\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1}=\sum_{\lambda\in\mathrm{Par}}s_\lambda(\mathbf{x})s_\lambda(\mathbf{y})=\sum_{\lambda\in\mathrm{Par}}h_\lambda(\mathbf{x})m_\lambda(\mathbf{y})=\sum_{\lambda\in\mathrm{Par}}\frac{p_\lambda(\mathbf{x})}{\sqrt{z_\lambda}}\frac{p_\lambda(\mathbf{y})}{\sqrt{z_\lambda}}$$

it suffices to show that any pair of bases $\{u_\lambda\},\{v_\lambda\}$ having

$$\sum_{\lambda\in\mathrm{Par}}s_\lambda(\mathbf{x})s_\lambda(\mathbf{y})=\sum_{\lambda\in\mathrm{Par}}u_\lambda(\mathbf{x})v_\lambda(\mathbf{y})$$

will be dual with respect to $(\cdot,\cdot)$. Write transition matrices $A=(a_{\nu,\lambda}), B=(b_{\nu,\lambda})$ uniquely expressing

$$u_\lambda = \sum_\nu a_{\nu,\lambda} s_\nu,$$
$$v_\lambda = \sum_\nu b_{\nu,\lambda} s_\nu.$$

Then orthonormality of the s_λ gives $(u_\alpha, v_\beta) = \sum_\nu a_{\nu,\alpha} b_{\nu,\beta}$, and hence we want $A^t B = I$, that is, $B^{-1} = A^t$. On the other hand, one has

$$\sum_\lambda s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}) = \sum_\lambda u_\lambda(\mathbf{x}) v_\lambda(\mathbf{y}) = \sum_\lambda \sum_\nu a_{\nu,\lambda} s_\nu(\mathbf{x}) \sum_\rho b_{\rho,\lambda} s_\rho(\mathbf{y}).$$

Comparing coefficients of $s_\nu(\mathbf{x}) s_\rho(\mathbf{y})$ forces $\sum_\lambda a_{\nu,\lambda} b_{\rho,\lambda} = \delta_{\nu,\rho}$, or in other words, $AB^t = I$. Since A and B^t are block-diagonal matrices with each block having finite size (as Λ is graded), this yields $B^t A = I$, and hence $A^t B = I$, as desired. ◻

2.6. **Bialternants, Littlewood-Richardson: Stembridge's concise proof.** There is a more natural way in which Schur functions arise as a $\mathbf{k}$-basis for Λ, coming from consideration of polynomials in a finite variable set, and the relation between those which are symmetric and those which are *alternating*.

For the remainder of this section, fix a positive integer n, and let $\mathbf{x} = (x_1, \ldots, x_n)$ be a finite variable set. This means that $s_{\lambda/\mu} = s_{\lambda/\mu}(\mathbf{x}) = \sum_T \mathbf{x}^{\mathrm{cont}(T)}$ is a generating function for column-strict tableaux T as in Definition 2.16, but with the extra condition that T have entries in $\{1, 2, \ldots, n\}$. We will assume without further mention that all partitions appearing in the section have at most n parts. Lastly, we also take $\mathbf{k} = \mathbb{Z}$ or a field of characteristic not equal to 2, to avoid certain annoyances in the discussion of alternating polynomials in characteristic 2.

Definition 2.33. Say that a polynomial $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ is *alternating* if for every permutation w in $\mathfrak{S}_n$ one has that

$$(wf)(\mathbf{x}) = f(x_{w(1)}, \ldots, x_{w(n)}) = \mathrm{sgn}(w) f(\mathbf{x}).$$

Let $\Lambda^{\mathrm{sgn}} \subset \mathbf{k}[x_1, \ldots, x_n]$ denote the subset of alternating polynomials[8].

As with Λ and its monomial basis $\{m_\lambda\}$, there is an obvious $\mathbf{k}$-basis for Λ^{sgn}, coming from the fact that a polynomial $f = \sum_\alpha c_\alpha \mathbf{x}^\alpha$ is alternating if and only if $c_{w(\alpha)} = \mathrm{sgn}(w) c_\alpha$ for every w in $\mathfrak{S}_n$ and every weak composition α. This means that every alternating f is a $\mathbf{k}$-linear combination of the following elements.

Definition 2.34. For $\alpha = (\alpha_1, \ldots, \alpha_n)$ in $\mathbb{N}^n$, define the *alternant*

$$a_\alpha := \sum_{w \in \mathfrak{S}_n} \mathrm{sgn}(w) w(\mathbf{x}^\alpha) = \det \begin{bmatrix} x_1^{\alpha_1} & \cdots & x_1^{\alpha_n} \\ x_2^{\alpha_1} & \cdots & x_2^{\alpha_n} \\ \vdots & \ddots & \vdots \\ x_n^{\alpha_1} & \cdots & x_n^{\alpha_n} \end{bmatrix}.$$

Example 2.35. One has

$$a_{(1,5,0)} = x_1^1 x_2^5 x_3^0 - x_1^5 x_2^1 x_3^0 - x_1^1 x_2^0 x_3^5 - x_1^0 x_2^5 x_3^1 + x_1^0 x_2^1 x_3^5 + x_1^5 x_2^0 x_3^1 = -a_{(5,1,0)}.$$

Meanwhile, $a_{(5,2,2)} = 0$ since the transposition $t = \binom{123}{132}$ fixes $(5,2,2)$ and hence

$$a_{(5,2,2)} = t(a_{(5,2,2)}) = \mathrm{sgn}(t) a_{(5,2,2)} = -a_{(5,2,2)}.$$

Alternatively, $a_{(5,2,2)} = 0$ as it is a determinant of a matrix with two equal columns.

This example illustrates that, for a $\mathbf{k}$-basis for Λ^{sgn}, one can restrict attention to alternants a_α in which α is a *strict partition*, i.e., in which α satisfies $\alpha_1 > \alpha_2 > \cdots > \alpha_n$. One can therefore uniquely express $\alpha = \lambda + \rho$, where λ is a (weak) partition $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ and where $\rho := (n-1, n-2, \ldots, 2, 1, 0)$ is sometimes called the *staircase partition* due to its Ferrers shape. For example $\alpha = (5,1,0) = (3,0,0) + (2,1,0) = \lambda + \rho$.

[8]When $\mathbf{k}$ has characteristic 2, it is probably best to define the alternating polynomials $\Lambda^{\mathrm{sgn}}_{\mathbf{k}}$ as the subspace $\Lambda^{\mathrm{sgn}} \otimes_{\mathbb{Z}} \mathbf{k} \subset \mathbb{Z}[x_1, \ldots, x_n] \otimes_{\mathbb{Z}} \mathbf{k} \cong \mathbf{k}[x_1, \ldots, x_n]$

Proposition 2.36. *The alternants* $\{a_{\lambda+\rho}\}$ *as* λ *runs through the partitions with at most* n *parts form a* $\mathbf{k}$*-basis for* Λ^{sgn}*. In addition, the* **bialternants** $\{\frac{a_{\lambda+\rho}}{a_\rho}\}$ *as* λ *runs through the same set form a* $\mathbf{k}$*-basis for* $\Lambda(x_1,\ldots,x_n)=\mathbf{k}[x_1,\ldots,x_n]^{\mathfrak{S}_n}$.

Proof. The first assertion should be clear from our previous discussion: the alternants $\{a_{\lambda+\rho}\}$ span Λ^{sgn} by definition, and they are $\mathbf{k}$-linearly independent because they are supported on disjoint sets of monomials $\mathbf{x}^\alpha$.

The second assertion follows from the first, after proving the following **Claim**: $f(\mathbf{x})$ lies in Λ^{sgn} if and only if $f(\mathbf{x})=a_\rho\cdot g(\mathbf{x})$ where $g(\mathbf{x})$ lies in $\mathbf{k}[\mathbf{x}]^{\mathfrak{S}_n}$ and where

$$a_\rho=\det(x_i^{n-j})_{i,j=1,2,\ldots,n}=\prod_{1\le i<j\le n}(x_i-x_j)$$

is the *Vandermonde determinant/product*. In other words

$$\Lambda^{\mathrm{sgn}}=a_\rho\cdot\mathbf{k}[\mathbf{x}]^{\mathfrak{S}_n}$$

is a free $\mathbf{k}[\mathbf{x}]^{\mathfrak{S}_n}$-module of rank one, with a_ρ as its $\mathbf{k}[\mathbf{x}]^{\mathfrak{S}_n}$-basis element.

To see the Claim, first note the inclusion

$$\Lambda^{\mathrm{sgn}}\supset a_\rho\cdot\mathbf{k}[\mathbf{x}]^{\mathfrak{S}_n}$$

since the product of a symmetric polynomial and an alternating polynomial is an alternating polynomial. For the reverse inclusion, note that since an alternating polynomial $f(\mathbf{x})$ changes sign whenever one exchanges two distinct variables x_i,x_j, it must vanish upon setting $x_i=x_j$, and therefore be divisible by x_i-x_j, so divisible by the entire product $\prod_{1\le i<j\le n}(x_i-x_j)=a_\rho$. But then the quotient $g(\mathbf{x})=\frac{f(\mathbf{x})}{a_\rho}$ is symmetric, as it is a quotient of two alternating polynomials. ☐

Our goal is to show that the mysterious bialternant basis $\{\frac{a_{\lambda+\rho}}{a_\rho}\}$ are actually the Schur functions $\{s_\lambda(\mathbf{x})\}$. Stembridge [75] noted that one could give a remarkably concise proof of an even stronger assertion, which simultaneously gives one of the standard combinatorial interpretations for the Littlewood-Richardson coefficients $c^\lambda_{\mu,\nu}$. For the purposes of stating it, we introduce for a tableau T the notation $T|_{\mathrm{cols}\ge j}$ (resp. $T|_{\mathrm{cols}\le j}$) to indicate the subtableau which is the restriction of T to the union of its columns $j,j+1,j+2,\ldots$ (resp. columns $1,2,\ldots,j$).

Theorem 2.37. *For partitions* λ,μ,ν *with* $\mu\subseteq\lambda$*, one has*

$$a_{\nu+\rho}s_{\lambda/\mu}=\sum_T a_{\nu+\mathrm{cont}(T)+\rho}$$

where T *runs through all column-strict tableaux with entries in* $\{1,2,\ldots,n\}$ *of shape* λ/μ *with the property that for all* $j=1,2,\ldots$ *one has* $\nu+\mathrm{cont}(T|_{\mathrm{cols}\ge j})$ *a partition.*

Before proving Theorem 2.37, let us see some of its consequences.

Corollary 2.38.

$$s_\lambda(\mathbf{x})=\frac{a_{\lambda+\rho}}{a_\rho}.$$

Proof. Take $\nu=\mu=\varnothing$ in Theorem 2.37. Note that for any λ, there is only one column-strict tableau T of shape λ having each $\mathrm{cont}(T|_{\mathrm{cols}\ge j})$ a partition, namely the one having every entry in row i equal to i:

$$\begin{matrix}1&1&1&1&1\\2&2&2&&\\3&3&3&&\\4&4&&&\end{matrix}$$

Furthermore, this T has $\mathrm{cont}(T)=\lambda$, so the theorem says $a_\rho s_\lambda=a_{\lambda+\rho}$. ☐

Example 2.39. For $n = 2$, so that $\rho = (1, 0)$, if we take $\lambda = (4, 2)$, then one has

$$\begin{aligned}\frac{a_{\lambda+\rho}}{a_\rho} &= \frac{a_{(4,2)+(1,0)}}{a_{(1,0)}} = \frac{a_{(5,2)}}{a_{(1,0)}} \\ &= \frac{x_1^5x_2^2 - x_1^2x_2^5}{x_1 - x_2} \\ &= x_1^4x_2^2 + x_1^3x_2^3 + x_1^2x_2^4 \\ &= \mathbf{x}^{\mathrm{cont}\begin{pmatrix}1111\\22\end{pmatrix}} + \mathbf{x}^{\mathrm{cont}\begin{pmatrix}1112\\22\end{pmatrix}} + \mathbf{x}^{\mathrm{cont}\begin{pmatrix}1122\\22\end{pmatrix}} \\ &= s_{(4,2)} = s_\lambda.\end{aligned}$$

Next divide through by a_ρ on both sides of the theorem to give the following.

Corollary 2.40. *For partitions λ, μ, ν having at most n parts, one has*

$$s_\nu s_{\lambda/\mu} = \sum_T s_{\nu+\mathrm{cont}(T)}$$

where T runs through the same set as in Theorem 2.37. In particular, taking $\nu = \varnothing$,

$$s_{\lambda/\mu} = \sum_T s_{\mathrm{cont}(T)}$$

where in the sum T runs through all column-strict tableaux of shape λ/μ for which each $\mathrm{cont}(T|_{\mathrm{cols}\geq j})$ is a partition.

By taking the number of variables n sufficiently large, one deduces from this last assertion the following version of the Littlewood-Richardson rule.

Corollary 2.41. *For partitions λ, μ, ν (of any lengths), the Littlewood-Richardson coefficient $c^\lambda_{\mu,\nu}$ counts column-strict tableaux T of shape λ/μ with $\mathrm{cont}(T) = \nu$ having the property that each $\mathrm{cont}(T|_{\mathrm{cols}\geq j})$ is a partition.*

Proof of Theorem 2.37. Start by rewriting the left side of the theorem, and using the fact that $w(s_{\lambda/\mu}) = s_{\lambda/\mu}$ for any w in $\mathfrak{S}_n$:

$$\begin{aligned}a_{\nu+\rho}s_{\lambda/\mu} &= \sum_{w\in\mathfrak{S}_n} \mathrm{sgn}(w)\mathbf{x}^{w(\nu+\rho)}w(s_{\lambda/\mu}) \\ &= \sum_{w\in\mathfrak{S}_n} \mathrm{sgn}(w)\mathbf{x}^{w(\nu+\rho)} \sum_{\substack{\text{column-strict } T\\ \text{of shape } \lambda/\mu}} \mathbf{x}^{w(\mathrm{cont}(T))} \\ &= \sum_{\substack{\text{column-strict } T\\ \text{of shape } \lambda/\mu}} \sum_{w\in\mathfrak{S}_n} \mathrm{sgn}(w)\mathbf{x}^{w(\nu+\mathrm{cont}(T)+\rho)} \\ &= \sum_{\substack{\text{column-strict } T\\ \text{of shape } \lambda/\mu}} a_{\nu+\mathrm{cont}(T)+\rho}\end{aligned}$$

We wish to cancel out all the summands indexed by column-strict tableaux T which fail any of the conditions that $\nu + \mathrm{cont}(T|_{\mathrm{cols}\geq j})$ be a partition. Given such a T, find the maximal j for which it fails this condition, and then find the minimal k for which

$$\nu_k + \mathrm{cont}_k(T|_{\mathrm{cols}\geq j}) < \nu_{k+1} + \mathrm{cont}_{k+1}(T|_{\mathrm{cols}\geq j}).$$

Maximality of j forces

$$\nu_k + \mathrm{cont}_k(T|_{\mathrm{cols}\geq j+1}) \geq \nu_{k+1} + \mathrm{cont}_{k+1}(T|_{\mathrm{cols}\geq j+1}).$$

Since column-strictness implies that column j of T can contain at most one occurrence of k or of $k + 1$ (or neither or both), the previous two inequalities imply that column j must contain an occurrence of $k+1$ and no occurrence of k, so that

$$\nu_k + \mathrm{cont}_k(T|_{\mathrm{cols}\geq j}) + 1 = \nu_{k+1} + \mathrm{cont}_{k+1}(T|_{\mathrm{cols}\geq j}).$$

This implies that the adjacent transposition $t_{k,k+1}$ swapping k and $k+1$ fixes the vector $\nu+\mathrm{cont}(T|_{\mathrm{cols}\geq j})+\rho$.

Now create a new tableau T^* from T by applying the Bender-Knuth involution (from the proof of Proposition 2.9) on letters $k, k+1$, but *only to columns* $1, 2, \ldots, j-1$ of T, leaving columns $j, j+1, j+2, \ldots$ unchanged. One should check that T^* is still column-strict, but this holds because column j of T has no occurrences of letter k. Note that

$$t_{k,k+1} \operatorname{cont}(T|_{\mathrm{cols} \leq j-1}) = \operatorname{cont}(T^*|_{\mathrm{cols} \leq j-1})$$

and hence

$$t_{k,k+1}(\nu + \operatorname{cont}(T) + \rho) = \nu + \operatorname{cont}(T^*) + \rho$$

so that $a_{\nu+\operatorname{cont}(T)+\rho} = -a_{\nu+\operatorname{cont}(T^*)+\rho}$.

Because T, T^* have exactly the same columns $j, j+1, j+2, \ldots$, the tableau T^* is also a violator of at least one of the conditions that $\nu + \operatorname{cont}(T^*|_{\mathrm{cols} \geq j})$ be a partition, and has the same choice of maximal j and minimal k as did T. Hence the map $T \mapsto T^*$ is an involution on the violators that lets one cancel their summands $a_{\nu+\operatorname{cont}(T)+\rho}$ and $a_{\nu+\operatorname{cont}(T^*)+\rho}$ in pairs. $\square$

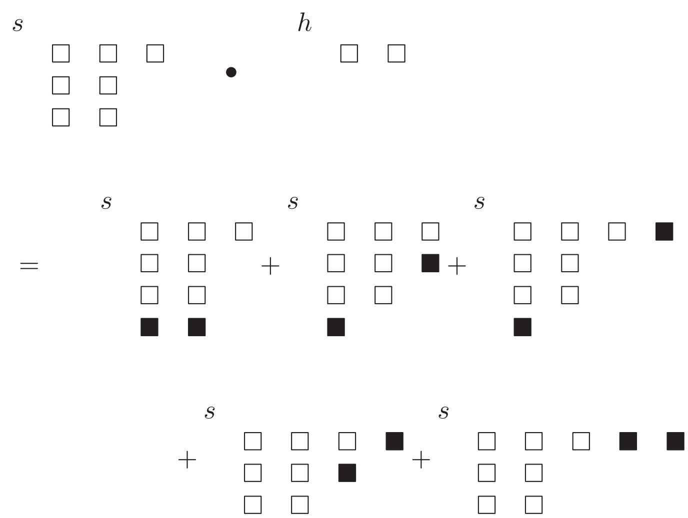
2.7. **The Pieri and Assaf-McNamara skew Pieri rule.** The classical Pieri rule refers to two special cases of the Littlewood-Richardson rule. To state them, recall that a skew shape with n cells is called a *horizontal (resp. vertical) n-strip* if no two of its cells lie in the same column (resp. row).

Theorem 2.42.

$$s_\lambda h_n = \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{\lambda^+}$$

$$s_\lambda e_n = \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^+}$$

Example 2.43.



Proof. For the first Pieri formula involving h_n, as $h_n = s_{(n)}$ one has

$$s_\lambda h_n = \sum_{\lambda^+} c^{\lambda^+}_{\lambda,(n)} s_{\lambda^+}.$$

Corollary 2.41 says $c^{\lambda^+}_{\lambda,(n)}$ counts column-strict tableaux T of shape λ^+/λ having $\operatorname{cont}(T) = (n)$ (i.e. all entries of T are 1's), with an extra condition. Since its entries are all equal, such a T must certainly have shape being a horizontal strip. Conversely for any horizontal strip, there is a unique such filling, and it will trivially satisfy the extra condition that $\operatorname{cont}(T|_{\mathrm{cols} \geq j})$ is a partition for each j. Hence $c^{\lambda^+}_{\lambda,(n)}$ is 1 if λ^+/λ is a horizontal n-strip, and 0 else.

For the second Pieri formula involving e_n, using $e_n = s_{(n)}$ one has

$$s_\lambda e_n = \sum_{\lambda^+} c^{\lambda^+}_{\lambda,(1^n)} s_{\lambda^+}.$$

Corollary 2.41 says $c^{\lambda^+}_{\lambda,(1^n)}$ counts column-strict tableaux T of shape λ^+/λ having $\mathrm{cont}(T) = (1^n)$, so its entries are $1, 2, \ldots, n$ each occurring once, with the extra condition that $1, 2, \ldots, n$ appear from right-to-left. Together with the tableau condition, this forces at most one entry in each row, that is λ^+/λ is a vertical strip, and then there is a unique way to fill it maintaining column-strictness. Thus $c^{\lambda^+}_{\lambda,(1^n)}$ is 1 if λ^+/λ is a vertical n-strip, and 0 else. $\square$

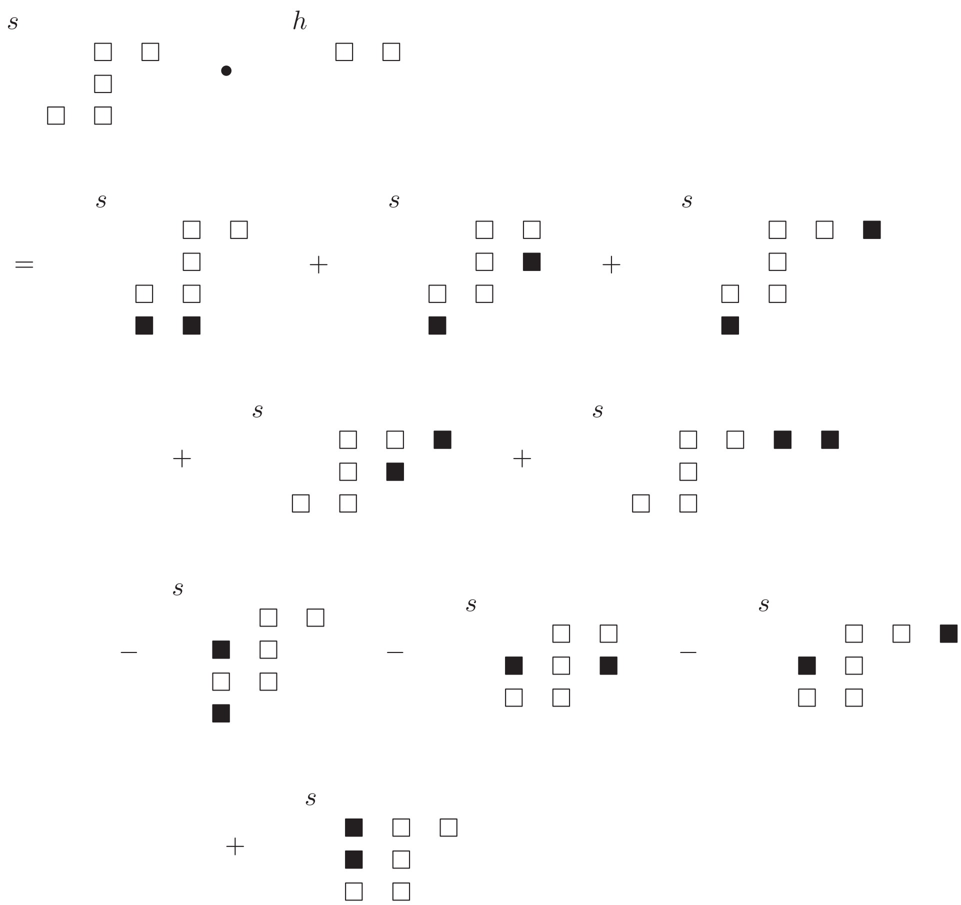
Assaf and McNamara [7] recently proved an elegant generalization.

Theorem 2.44.

$$s_{\lambda/\mu} h_n = \sum_{\substack{\lambda^+,\mu^-: \\ \lambda^+/\lambda \text{ a horizontal strip} \\ \mu/\mu^- \text{ a vertical strip} \\ |\lambda^+/\lambda|+|\mu/\mu^-|=n}} (-1)^{|\mu/\mu^-|} s_{\lambda^+/\mu^-}$$

$$s_{\lambda/\mu} e_n = \sum_{\substack{\lambda^+,\mu^-: \\ \lambda^+/\lambda \text{ a vertical strip} \\ \mu/\mu^- \text{ a horizontal strip} \\ |\lambda^+/\lambda|+|\mu/\mu^-|=n}} (-1)^{|\mu/\mu^-|} s_{\lambda^+/\mu^-}$$

Example 2.45.



Theorem 2.44 is proven in the next section, using an important Hopf algebra tool.

2.8. **Skewing and Lam's proof of the skew Pieri rule.** We codify here the operation $s_\mu^\perp$ of *skewing by* s_μ, acting on Schur functions via

$$s_\mu^\perp(s_\lambda) = s_{\lambda/\mu}$$

where one defines $s_{\lambda/\mu} = 0$ if $\mu \not\subseteq \lambda$. These operations play a crucial role

- in Lam's proof of the skew Pieri rule,
- in Lam, Lauve, and Sottile's proof [42] of a more general skew Littlewood-Richardson rule that had been conjectured by Assaf and McNamara, and
- in Zelevinsky's structure theory of PSH-algebras to be developed in the next chapter.

Definition 2.46. Given a Hopf algebra A of finite type, and its (restricted) dual A^o, let $(\cdot,\cdot) = (\cdot,\cdot)_A$ be the pairing $(f,a) := f(a)$ for f in A^o and a in A. Then define for each f in A^o an operator $A \xrightarrow{f^\perp} A$ as follows[9]: for a in A with $\Delta(a) = \sum a_1 \otimes a_2$, let

$$f^\perp(a) = \sum (f, a_1) a_2.$$

Note that when one takes $A = \Lambda = A^o$, the element $a = s_\lambda$ has $\Delta s_\lambda = \sum_\mu s_\mu \otimes s_{\lambda/\mu}$. Hence if $f = s_\mu$, then one has $f^\perp(a) = s_{\lambda/\mu} = s_\mu^\perp(s_\lambda)$ as desired.

Proposition 2.47. *The $f^\perp$ operators $A \to A$ have the following properties.*

(i) *$f^\perp$ is adjoint to left multiplication $A^o \xrightarrow{f\cdot} A^o$ in the sense that*
$$(g, f^\perp(a)) = (fg, a).$$
(ii) *$(fg)^\perp(a) = g^\perp(f^\perp(a))$, that is, A becomes a right A^o-module via the $f^\perp$ action.*
(iii) *If $\Delta(f) = \sum f_1 \otimes f_2$, then*
$$f^\perp(ab) = \sum f_1^\perp(a) f_2^\perp(b).$$
In particular, if f is primitive in A^o, so that $\Delta(f) = f \otimes 1 + 1 \otimes f$, then $f^\perp$ is a **derivation***:*
$$f^\perp(ab) = f^\perp(a) \cdot b + a \cdot f^\perp(b).$$

Proof. For (i), note that

$$(g, f^\perp(a)) = \sum (f, a_1)(g, a_2) = (f \otimes g, \Delta_A(a)) = (m_{A^o}(f \otimes g), a) = (fg, a).$$

For (ii), using (i) and considering any h in A^o, one has that

$$(h, (fg)^\perp(a)) = (fgh, a) = (gh, f^\perp(a)) = (h, g^\perp(f^\perp(a))).$$

For (iii), noting that

$$\Delta(ab) = \Delta(a)\Delta(b) = \left(\sum_{(a)} a_1 \otimes a_2\right)\left(\sum_{(b)} b_1 \otimes b_2\right) = \sum_{(a),(b)} a_1 b_1 \otimes a_2 b_2,$$

one has that

$$\begin{aligned} f^\perp(ab) &= \sum_{(a),(b)} (f, a_1 b_1)_A \; a_2 b_2 = \sum_{(a),(b)} (\Delta(f), a_1 \otimes b_1)_{A \otimes A} \; a_2 b_2 \\ &= \sum_{(f),(a),(b)} (f_1, a_1)_A (f_2, b_1)_A \; a_2 b_2 \\ &= \sum_{(f)} \left(\sum_{(a)} (f_1, a_1)_A a_2\right)\left(\sum_{(b)} (f_2, b_1)_A b_2\right) = \sum_{(f)} f_1^\perp(a) f_2^\perp(b). \end{aligned}$$

$\square$

The following interaction between multiplication and $h^\perp$ is the key to deducing the skew Pieri formula from the usual Pieri formulas.

Lemma 2.48. *For any f, g in Λ, one has*

$$f \cdot h_n^\perp(g) = \sum_{k=0}^{n} (-1)^k h_{n-k}^\perp (e_k^\perp(f) \cdot g).$$

[9]This $f^\perp(a)$ is called $a \leftharpoonup f$ in Montgomery [56, Example 1.6.5]

Proof. Starting with the right side, first apply Proposition 2.47(iii):

$$\begin{aligned}
&\sum_{k=0}^{n}(-1)^k h_{n-k}^{\perp}(e_k^{\perp}(f)\cdot g)\\
&=\sum_{k=0}^{n}(-1)^k\sum_{j=0}^{n-k} h_j^{\perp}(e_k^{\perp}(f))\cdot h_{n-k-j}^{\perp}(g)\\
&=\sum_{i=0}^{n}(-1)^{n-i}\left(\sum_{j=0}^{n-i}(-1)^j h_j^{\perp}(e_{n-i-j}^{\perp}(f))\right)\cdot h_i^{\perp}(g) \quad (\text{ reindexing } i:=n-k-j\ )\\
&=\sum_{i=0}^{n}(-1)^{n-i}\left(\sum_{j=0}^{n-i}(-1)^j e_{n-i-j}h_j\right)^{\perp}(f)\cdot h_i^{\perp}(g) \quad (\text{ by Proposition 2.47(ii) })\\
&=1^{\perp}(f)\cdot h_n^{\perp}(g)=f\cdot h_n^{\perp}(g)
\end{aligned}$$

where the second-to-last equality used (2.11). □

Proof of Theorem 2.44. We prove the first skew Pieri rule; the second is analogous, swapping $h_i \leftrightarrow e_i$ and swapping the words "vertical" $\leftrightarrow$ "horizontal". Symmetry of $(\cdot,\cdot)_\Lambda$, commutativity of Λ, and Proposition 2.47(i) imply for f in Λ,

$$\big(s_{\lambda/\mu}, f\big) = \big(s_\mu^{\perp}(s_\lambda), f\big) = (s_\lambda, s_\mu f) \tag{2.21}$$

Hence for any g in Λ, one can compute that

$$\begin{aligned}
(h_n s_{\lambda/\mu}\ ,\ g) &\overset{\text{Prop. } 2.47(i)}{=} (s_{\lambda/\mu}\ ,\ h_n^{\perp} g) \overset{(2.21)}{=} (s_\lambda\ ,\ s_\mu\cdot h_n^{\perp} g)\\
&\overset{\text{Lemma } 2.48}{=} \sum_{k=0}^{n}(-1)^k(s_\lambda\ ,\ h_{n-k}^{\perp}(e_k^{\perp}(s_\mu)\cdot g))\\
&\overset{\text{Prop. } 2.47(i)}{=} \sum_{k=0}^{n}(-1)^k(h_{n-k}s_\lambda\ , e_k^{\perp}(s_\mu)\cdot g)
\end{aligned} \tag{2.22}$$

The first Pieri rule in Theorem 2.42 lets one rewrite $h_{n-k}s_\lambda = \sum_{\lambda^+} s_{\lambda^+}$, with the sum running through λ^+ for which λ^+/λ is a horizontal $(n-k)$-strip. The second Pieri rule in Theorem 2.42 lets one rewrite $e_k^{\perp} s_\mu = \sum_{\mu^-} s_{\mu^-}$, with the sum running through μ^- for which μ/μ^- is a vertical k-strip, since $(s_{\mu^-}, e_k^{\perp} s_\mu) = (e_k s_{\mu^-}, s_\mu)$. Thus the last line of (2.22) becomes

$$\sum_{k=0}^{n}(-1)^k\left(\sum_{\lambda^+} s_{\lambda^+}\ ,\ \sum_{\mu^-} s_{\mu^-}\cdot g\right) \overset{(2.21)}{=} \left(\sum_{k=0}^{n}(-1)^k\sum_{(\lambda^+,\mu^-)} s_{\lambda^+/\mu^-}, g\right)$$

where the sum is over the pairs (λ^+,μ^-) for which λ^+/λ is a horizontal $(n-k)$-strip and μ/μ^- is a vertical k-strip. □

Exercise 2.49. The goal of this exercise is to prove (2.15) using the skewing operators that we have developed. Recall the involution $\omega:\Lambda\to\Lambda$ defined in (2.13).

(a) Show that $\omega(p_\lambda) = (-1)^{|\lambda|-\ell(\lambda)} p_\lambda$ for any $\lambda\in\operatorname{Par}$, where $\ell(\lambda)$ denotes the length of the partition λ.
(b) Show that ω is an isometry.
(c) Show that this same map $\omega:\Lambda\to\Lambda$ is a Hopf automorphism.
(d) Prove that $\omega\left(a^{\perp}b\right) = (\omega(a))^{\perp}(\omega(b))$ for every $a\in\Lambda$ and $b\in\Lambda$.
(e) For any partition $\lambda=(\lambda_1,\ldots,\lambda_\ell)$ with length $\ell(\lambda)=\ell$, prove that

$$e_\ell^{\perp} s_\lambda = s_{(\lambda_1-1,\lambda_2-1,\ldots,\lambda_\ell-1)}.$$

(f) For any partition $\lambda=(\lambda_1,\lambda_2,\ldots)$, prove that

$$h_{\lambda_1}^{\perp} s_\lambda = s_{(\lambda_2,\lambda_3,\lambda_4,\ldots)}.$$

(g) Prove (2.15).

3. Zelevinsky's structure theory of positive self-dual Hopf algebras

Section 2 showed that, as a $\mathbb{Z}$-basis for the Hopf algebra $\Lambda = \Lambda_{\mathbb{Z}}$, the Schur functions $\{s_\lambda\}$ have two special properties: they have the *same* structure constants $c^\lambda_{\mu,\nu}$ for their multiplication as for their comultiplication (Corollary 2.25), and these structure constants are all *nonnegative* integers (Corollary 2.41). Zelevinsky [81, §2,3] isolated these two properties as crucial.

Definition 3.1. Say that a graded connected Hopf algebra A over $\mathbf{k} = \mathbb{Z}$ with a distinguished $\mathbb{Z}$-basis $\{\sigma_\lambda\}$ consisting of homogeneous elements is a *positive self-dual Hopf algebra* (or PSH) if it satisfies the two further axioms

- **(self-duality)** The same structure constants $a^\lambda_{\mu,\nu}$ appear for the product $\sigma_\mu \sigma_\nu = \sum_\lambda a^\lambda_{\mu,\nu} \sigma_\lambda$ and the coproduct $\Delta \sigma_\lambda = \sum_{\mu,\nu} a^\lambda_{\mu,\nu} \sigma_\mu \otimes \sigma_\nu$.
- **(positivity)** The $a^\lambda_{\mu,\nu}$ are all nonnegative (integers).

Call $\{\sigma_\lambda\}$ the *PSH-basis* of A.

He then developed a beautiful structure theory for PSH-algebras, explaining how they can be uniquely expressed as tensor products of copies of PSH's each isomorphic to Λ after rescaling their grading. The next few sections explain this, following his exposition closely.

3.1. Self-duality implies polynomiality.

We begin with a property that forces a Hopf algebra to have algebra structure which is a *polynomial* algebra, specifically the symmetric algebra $\mathrm{Sym}(\mathfrak{p})$, where $\mathfrak{p}$ is the subspace of primitive elements.

Recall from Exercise 1.22 that for a graded connected Hopf algebra $A = \bigoplus_{n=0}^\infty A_n$, every x in the two-sided ideal $I := \ker \epsilon = \bigoplus_{n>0} A_n$ has the property that its comultiplication takes the form

$$\Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x)$$

where $\Delta_+(x)$ lies in $I \otimes I$. Recall also that the elements x for which $\Delta_+(x) = 0$ are called the *primitives*. Denote by $\mathfrak{p}$ the $\mathbb{Z}$-submodule of primitive elements inside A.

Given a Hopf algebra A over $\mathbf{k} = \mathbb{Z}$ with a PSH-basis $\{\sigma_\lambda\}$, we identify A^o with A via the bilinear form $(\cdot,\cdot)_A$ on A that makes this basis orthonormal. Similarly, the elements $\{\sigma_\lambda \otimes \sigma_\mu\}$ give an orthonormal basis for a form $(\cdot,\cdot)_{A\otimes A}$ on $A \otimes A$. This is an instance of the following notion of self-duality.

Definition 3.2. Say that a bialgebra A is *self-dual* with respect to a given symmetric bilinear form $(\cdot,\cdot)$ if one has $(a, m(b \otimes c))_A = (\Delta(a), b \otimes c)$ and $(1_A, a) = \epsilon(a)$ for a, b, c in A. If A is a graded Hopf algebra of finite type then this is equivalent to the $\mathbf{k}$-module map $A \to A^o$ induced by $(\cdot,\cdot)_A$ giving a Hopf algebra isomorphism.

Proposition 3.3. *Let A be a Hopf algebra over $\mathbf{k} = \mathbb{Z}$ or $\mathbf{k} = \mathbb{Q}$ which is graded, connected, and self-dual with respect to a positive definite graded*[10] *bilinear form. Then within the ideal I, the subspace of primitives $\mathfrak{p}$ is the orthogonal complement to the subspace I^2. In particular, $\mathfrak{p} \cap I^2 = 0$, and when $\mathbf{k} = \mathbb{Q}$, one has $I = \mathfrak{p} \oplus I^2$.*

Proof. Note that $I^2 = m(I \otimes I)$. Hence an element x in I lies in the perpendicular space to I^2 if and only if one has for all y in $I \otimes I$ that

$$0 = (x, m(y))_A = (\Delta(x), y)_{A\otimes A} = (\Delta_+(x), y)_{A\otimes A}$$

where the second equality uses self-duality, while the third equality uses the fact that y lies in $I \otimes I$ and the form $(\cdot,\cdot)_{A\otimes A}$ makes distinct homogeneous components orthogonal. Since y was arbitrary, this means x is perpendicular to I^2 if and only if $\Delta_+(x) = 0$, that is, x lies in $\mathfrak{p}$. ☐

Remark 3.4. One might wonder why we didn't just say $I = \mathfrak{p} \oplus I^2$ even when $\mathbf{k} = \mathbb{Z}$ in Proposition 3.3. However, this is false even for $A = \Lambda_{\mathbb{Z}}$: the second homogeneous component $(\mathfrak{p} \oplus I^2)_2$ is the index 2 sublattice of Λ_2 which is $\mathbb{Z}$-spanned by $\{p_2, e_1^2\}$, containing $2e_2$, but not containing e_2 itself.

Already the fact that $\mathfrak{p} \cap I^2 = 0$ has a strong implication.

Lemma 3.5. *A graded connected Hopf algebra A over any ring $\mathbf{k}$ having $\mathfrak{p} \cap I^2 = 0$ has algebra structure which is commutative.*

[10] That is, $(A_i, A_j) = 0$ for $i \neq j$.

Proof. The component $A_0 = \mathbf{k}$ commutes with all of A. This forms the base case for an induction on $i+j$ in which one shows that any elements x in A_i and y in A_j with $i,j>0$ will have $[x,y] := xy - yx = 0$. Since $[x,y]$ lies in I^2, it suffices to show that $[x,y]$ also lies in $\mathfrak{p}$:

$$\begin{aligned}
\Delta[x,y] &= [\Delta(x), \Delta(y)] \\
&= [1\otimes x + x\otimes 1 + \Delta_+(x), 1\otimes y + y\otimes 1 + \Delta_+(y)] \\
&= [1\otimes x + x\otimes 1, 1\otimes y + y\otimes 1] \\
&\qquad + [1\otimes x + x\otimes 1, \Delta_+(y)] + [\Delta_+(x), 1\otimes y + y\otimes 1] + [\Delta_+(x), \Delta_+(y)] \\
&= [1\otimes x + x\otimes 1, 1\otimes y + y\otimes 1] \\
&= 1\otimes [x,y] + [x,y]\otimes 1
\end{aligned}$$

showing that $[x,y]$ lies in $\mathfrak{p}$. Here the second-to-last equality used the inductive hypotheses: homogeneity implies that $\Delta_+(x)$ is a sum of homogeneous tensors of the form $z_1 \otimes z_2$ satisfying $\deg(z_1), \deg(z_2) < i$, so that by induction they will commute with $1\otimes y, y\otimes 1$, thus proving that $[\Delta_+(x), 1\otimes y + y\otimes 1] = 0$; a symmetric argument shows $[1\otimes x + x\otimes 1, \Delta_+(y)] = 0$, and, a similar argument shows $[\Delta_+(x), \Delta_+(y)] = 0$. The last equality is an easy calculation, and was done already in (1.9). ☐

Remark 3.6. Zelevinsky actually shows [81, Proof of A.1.3, p. 150] that the self-duality assumption (along with hypotheses of unit, counit, graded, connected, and Δ being a morphism for multiplication) already implies the *associativity* of the multiplication in A ! One shows by induction on $i+j+k$ that any x,y,z in A_i, A_j, A_k with $i,j,k>0$ have vanishing *associator* $\operatorname{assoc}(x,y,z) := x(yz) - (xy)z$. In the inductive step, one first notes that $\operatorname{assoc}(x,y,z)$ lies in I^2, and then checks that $\operatorname{assoc}(x,y,z)$ also lies in $\mathfrak{p}$, by a calculation very similar to the one above, repeatedly using the fact that $\operatorname{assoc}(x,y,z)$ is multilinear in its three arguments.

This leads to a general structure theorem.

Theorem 3.7. *If a graded, connected Hopf algebra A over a field $\mathbf{k}$ of characteristic zero has $I = \mathfrak{p} \oplus I^2$, then the inclusion $\mathfrak{p} \hookrightarrow A$ extends to a Hopf algebra isomorphism from the symmetric algebra $\operatorname{Sym}_{\mathbf{k}}(\mathfrak{p}) \to A$. In particular, A is both commutative and cocommutative.*

Note that these hypotheses are valid, using Proposition 3.3, whenever A is obtained from a PSH (over $\mathbb{Z}$) by tensoring with $\mathbb{Q}$.

Proof. Since Lemma 3.5 implies that A is commutative, the universal property of $\operatorname{Sym}_{\mathbf{k}}(\mathfrak{p})$ as a free commutative algebra on generators $\mathfrak{p}$ shows that the inclusion $\mathfrak{p} \hookrightarrow A$ at least extends to an algebra morphism $\operatorname{Sym}_{\mathbf{k}}(\mathfrak{p}) \xrightarrow{\varphi} A$. Since the Hopf structure on $\operatorname{Sym}_{\mathbf{k}}(\mathfrak{p})$ makes the elements of $\mathfrak{p}$ primitive (see Example 1.18), this φ is actually a coalgebra morphism (since $\Delta \circ \varphi = (\varphi \otimes \varphi) \circ \Delta$ and $\epsilon \circ \varphi = \epsilon$ need only to be checked on algebra generators), hence a bialgebra morphism, hence a Hopf algebra morphism (by Proposition 1.35 (c)). It remains to show that φ is surjective, and injective.

For the surjectivity of φ, note that the hypothesis $I = \mathfrak{p} \oplus I^2$ implies that the composite $\mathfrak{p} \hookrightarrow I \to I/I^2$ gives a $\mathbf{k}$-vector space isomorphism. What follows is a standard argument to deduce that $\mathfrak{p}$ generates A as a commutative graded $\mathbf{k}$-algebra. One shows by induction on n that any homogeneous element a in A_n lies in the $\mathbf{k}$-subalgebra generated by $\mathfrak{p}$. The base case $n=0$ is trivial as a lies in $\mathbf{k}$. In the inductive step where a lies in I, write $a \equiv p \bmod I^2$ for some p in $\mathfrak{p}$. Thus $a = p + \sum_i b_i c_i$, where b_i, c_i lie in I but have strictly smaller degree, so that by induction they lie in the subalgebra generated by $\mathfrak{p}$, and hence so does a.

Note that the surjectivity argument did not use the assumption that $\mathbf{k}$ has characteristic zero, but we will now use it in the injectivity argument for φ, to establish the following

(3.1) **Claim:** Every primitive element of $\operatorname{Sym}(\mathfrak{p})$ lies in $\mathfrak{p} = \operatorname{Sym}^1(\mathfrak{p})$.

Note that this claim fails in positive characteristic, e.g. if $\mathbf{k}$ has characteristic 2 then x^2 lies in $\operatorname{Sym}^2(\mathfrak{p})$, however

$$\Delta(x^2) = 1 \otimes x^2 + 2x \otimes x + x^2 \otimes 1 = 1 \otimes x^2 + x^2 \otimes 1.$$

To see the claim, assume not, so that by gradedness, there must exist some primitive element $y \neq 0$ lying in some $\operatorname{Sym}^n(\mathfrak{p})$ with $n \geq 2$. This would mean that the composite map f that follows the coproduct with a component projection

$$\mathrm{Sym}^n(\mathfrak{p}) \xrightarrow{\Delta} \bigoplus_{i+j=n} \mathrm{Sym}^i(\mathfrak{p}) \otimes \mathrm{Sym}^j(\mathfrak{p}) \longrightarrow \mathrm{Sym}^1(\mathfrak{p}) \otimes \mathrm{Sym}^{n-1}(\mathfrak{p})$$

has $f(y) = 0$. However, one can check on a basis that the multiplication backward $\mathrm{Sym}^1(\mathfrak{p}) \otimes \mathrm{Sym}^{n-1}(\mathfrak{p}) \xrightarrow{m} \mathrm{Sym}^n(\mathfrak{p})$ has the property that $m \circ f = n \cdot 1_{\mathrm{Sym}^n(\mathfrak{p})}$:

$$(m \circ f)(x_1 \cdots x_n) = m\left(\sum_{j=1}^{n} x_j \otimes x_1 \cdots \widehat{x_j} \cdots x_n\right) = n \cdot x_1 \cdots x_n$$

for $x_1, \ldots, x_n$ in $\mathfrak{p}$. Then $n \cdot y = m(f(y)) = m(0) = 0$ leads to the contradiction that $y = 0$, since $\mathbf{k}$ has characteristic zero.

Now one can argue the injectivity of the (graded) map[11] φ by assuming that one has a nonzero homogeneous element u in $\ker(\varphi)$ of minimum degree. In particular, $\deg(u) \geq 1$. Also since $\mathfrak{p} \hookrightarrow A$, one has that u is not in $\mathrm{Sym}^1(\mathfrak{p}) = \mathfrak{p}$, and hence u is not primitive by the previous Claim. Consequently $\Delta_+(u) \neq 0$, and one can find a nonzero component $u^{(i,j)}$ of $\Delta_+(u)$ lying in $\mathrm{Sym}(\mathfrak{p})_i \otimes \mathrm{Sym}(\mathfrak{p})_j$ for some $i, j > 0$. Since this forces $i, j < \deg(u)$, one has that φ maps both $\mathrm{Sym}(\mathfrak{p})_i, \mathrm{Sym}(\mathfrak{p})_j$ injectively into A_i, A_j. Hence the tensor product map

$$\mathrm{Sym}(\mathfrak{p})_i \otimes \mathrm{Sym}(\mathfrak{p})_j \xrightarrow{\varphi \otimes \varphi} A_i \otimes A_j$$

is also injective[12]. This implies $(\varphi \otimes \varphi)(u^{(i,j)}) \neq 0$, giving the contradiction that

$$0 = \Delta_+^A(0) = \Delta_+^A(\varphi(u)) = (\varphi \otimes \varphi)(\Delta_+^{\mathrm{Sym}(\mathfrak{p})}(u))$$

contains the nonzero $A_i \otimes A_j$-component $(\varphi \otimes \varphi)(u^{(i,j)})$. ☐

Before closing this section, we mention one nonobvious corollary of the Claim (3.1), when applied to the ring of symmetric functions $\Lambda_{\mathbb{Q}}$ with $\mathbb{Q}$-coefficients, since Proposition 2.19 says that $\Lambda_{\mathbb{Q}} = \mathbb{Q}[p_1, p_2, \ldots] = \mathrm{Sym}(V)$ where $V = \mathbb{Q}\{p_1, p_2, \ldots\}$.

Corollary 3.8. *The subspace $\mathfrak{p}$ of primitives in $\Lambda_{\mathbb{Q}}$ is one-dimensional in each degree $n = 1, 2, \ldots$, and spanned by $\{p_1, p_2, \ldots\}$.*

3.2. **The decomposition theorem.** Our goal here is Zelevinsky's theorem [81, Theorem 2.2] giving a canonical decomposition of any PSH as a tensor product into PSH's that each have only one primitive element in their PSH-basis. For the sake of stating it, we introduce some notation.

Definition 3.9. Given A a PSH with PSH-basis Σ, let $\mathcal{C} := \Sigma \cap \mathfrak{p}$ be the primitive elements in Σ. For each ρ in $\mathcal{C}$, let $A(\rho) \subset A$ be the $\mathbb{Z}$-span of

$$\Sigma(\rho) := \{\sigma \in \Sigma : \text{ there exists } n \geq 0 \text{ with } (\sigma, \rho^n) \neq 0\}.$$

Theorem 3.10. *Any PSH A has a canonical tensor product decomposition*

$$A = \bigotimes_{\rho \in \mathcal{C}} A(\rho)$$

with $A(\rho)$ a PSH, and ρ the only primitive element in its PSH-basis $\Sigma(\rho)$.

---

[11] The grading on $\mathrm{Sym}(\mathfrak{p})$ is induced from the grading on $\mathfrak{p}$, a homogeneous subspace of $I \subset A$ as it is the kernel of the graded map $I \xrightarrow{\Delta_+} A \otimes A$.

[12] One needs to know that for two injective maps $V_i \xrightarrow{\varphi_i} W_i$ of $\mathbf{k}$-vector spaces V_i, W_i with $i = 1, 2$, the tensor product $\varphi_1 \otimes \varphi_2$ is also injective. Factoring it as $\varphi_1 \otimes \varphi_2 = (1 \otimes \varphi_2) \circ (\varphi_1 \otimes 1)$, one sees that it suffices to show that for an injective map $V \overset{\varphi}{\hookrightarrow} W$ of free $\mathbf{k}$-modules, and any free $\mathbf{k}$-module U, the map $V \otimes U \xrightarrow{\varphi \otimes 1} W \otimes U$ is also injective. Since tensor products commute with direct sums, and U is (isomorphic to) a direct sum of copies of $\mathbf{k}$, this reduces to the easy-to-check case where $U = \mathbf{k}$.

Note that some kind of freeness or flatness hypothesis on U is needed here since, e.g. the injective $\mathbb{Z}$-module maps $\mathbb{Z} \xrightarrow{\varphi_1 = (\cdot \times 2)} \mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z} \xrightarrow{\varphi_2 = 1} \mathbb{Z}/2\mathbb{Z}$ have $\varphi_1 \otimes \varphi_2 = 0$ on $\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \neq 0$.

Although in all the applications, $\mathcal{C}$ will be finite, when $\mathcal{C}$ is infinite one should interpret the tensor product in the theorem as the inductive limit of tensor products over finite subsets of $\mathcal{C}$, that is, linear combinations of basic tensors $\bigotimes_\rho a_\rho$ in which there are only finitely many factors $a_\rho \neq 1$.

The first step toward the theorem uses a certain unique factorization property.

Lemma 3.11. *Given $\mathcal{P}$ a set of pairwise orthogonal primitives in a PSH A,*

$$(\rho_1 \cdots \rho_r, \pi_1 \cdots \pi_s) = 0$$

for ρ_i, π_j in $\mathcal{P}$ unless $r = s$ and one can reindex so that $\rho_i = \pi_i$.

Proof. Induct on $\min(r, s)$. One has

$$\begin{aligned}(\rho_1 \cdots \rho_r, \pi_1 \cdots \pi_s) &= (\rho_2 \cdots \rho_r, \rho_1^\perp(\pi_1 \cdots \pi_s)) \\ &= (\rho_2 \cdots \rho_r, \sum_{j=1}^{s} (\pi_1 \cdots \pi_{j-1} \cdot \rho_1^\perp(\pi_j) \cdot \pi_{j+1} \cdots \pi_s))\end{aligned}$$

from Proposition 2.47(iii) because ρ_1 is primitive. On the other hand, since each π_j is primitive, one has $\rho_1^\perp(\pi_j) = (\rho_1, 1) \cdot \pi_j + (\rho_1, \pi_j) \cdot 1 = (\rho_1, \pi_j)$ which vanishes unless $\rho_1 = \pi_j$. Hence $(\rho_1 \cdots \rho_r, \pi_1 \cdots \pi_s) = 0$ unless $\rho_1 \in \{\pi_1, \ldots, \pi_s\}$, in which case after reindexing so that $\pi_1 = \rho_1$, it equals

$$n \cdot (\rho_1, \rho_1) \cdot (\rho_2 \cdots \rho_r, \pi_2 \cdots \pi_s)$$

if there are exactly n occurrences of ρ_1 among $\pi_1, \ldots, \pi_s$. Now apply induction. $\square$

So far the positivity hypothesis for a PSH has played little role. Now we use it to introduce a certain partial order on the PSH A, and then a semigroup grading.

Definition 3.12. Let $\mathbb{N} := \{0, 1, 2, \ldots\}$, and for a subset S of an abelian group, let $\mathbb{Z}S$ (resp. $\mathbb{N}S$) denote the subgroup of $\mathbb{Z}$-linear combinations (resp. subsemigroup of $\mathbb{N}$-linear combinations) of the elements of S.

In a PSH A with PSH-basis Σ, the subset $\mathbb{N}\Sigma$ forms a subsemigroup, and lets one define a partial order on A via $a \leq b$ if $b - a$ lies in $\mathbb{N}\Sigma$.

We note a few trivial properties of this partial order:

- The positivity hypothesis implies that $\mathbb{N}\Sigma \cdot \mathbb{N}\Sigma \subset \mathbb{N}\Sigma$.
- Hence multiplication by an element $c \geq 0$ (meaning c lies in $\mathbb{N}\Sigma$) preserves the order: $a \leq b$ implies $ac \leq bc$ since $(b - a)c$ lies in $\mathbb{N}\Sigma$.
- Thus $0 \leq c \leq d$ and $0 \leq a \leq b$ implies $ac \leq bc \leq bd$.

This allows one to introduce a semigroup grading on A.

Definition 3.13. Let $\mathbb{N}^{\mathcal{C}}_{\mathrm{fin}}$ denote the additive subsemigroup of $\mathbb{N}^{\mathcal{C}}$ consisting of those $\alpha = (\alpha_\rho)_{\rho \in \mathcal{C}}$ with finite support.

Note that for any α in $\mathbb{N}^{\mathcal{C}}_{\mathrm{fin}}$, one has that the product $\prod_{\rho \in \mathcal{C}} \rho^{\alpha_\rho} \geq 0$. Define

$$\Sigma(\alpha) := \{\sigma \in \Sigma : \sigma \leq \prod_{\rho \in \mathcal{C}} \rho^{\alpha_\rho}\},$$

that is, the subset of Σ on which $\prod_{\rho \in \mathcal{C}} \rho^{\alpha_\rho}$ has support. Also define

$$A_{(\alpha)} := \mathbb{Z}\Sigma(\alpha) \subset A.$$

Proposition 3.14. *The PSH A has an $\mathbb{N}^{\mathcal{C}}_{\mathrm{fin}}$-semigroup-grading: one has an orthogonal direct sum decomposition*

$$A = \bigoplus_{\alpha \in \mathbb{N}^{\mathcal{C}}_{\mathrm{fin}}} A_{(\alpha)}$$

for which

$$A_{(\alpha)} A_{(\beta)} \subset A_{(\alpha+\beta)} \tag{3.2}$$

$$\Delta A_{(\alpha)} \subset \bigoplus_{\alpha = \beta + \gamma} A_{(\beta)} \otimes A_{(\gamma)}. \tag{3.3}$$

Proof. We will make free use of the fact that a PSH A is commutative, since it embeds in $A \otimes_{\mathbb{Z}} \mathbb{Q}$, which is commutative by Theorem 3.7.

Note that the orthogonality $\left(A_{(\alpha)}, A_{(\beta)}\right) = 0$ for $\alpha \neq \beta$ is equivalent to the assertion that $\left(\prod_{\rho\in\mathcal{C}} \rho^{\alpha_\rho}, \prod_{\rho\in\mathcal{C}} \rho^{\beta_\rho}\right) = 0$, which follows from Lemma 3.11.

Next let us deal with the assertion (3.2). It suffices to check that when τ, ω in Σ lie in $A_{(\alpha)}, A_{(\beta)}$, respectively, then $\tau\omega$ lies in $A_{(\alpha+\beta)}$. But note that any σ in Σ having $\sigma \leq \tau\omega$ will then have

$$\sigma \leq \tau\omega \leq \prod_{\rho\in\mathcal{C}} \rho^{\alpha_\rho} \cdot \prod_{\rho\in\mathcal{C}} \rho^{\beta_\rho} = \prod_{\rho\in\mathcal{C}} \rho^{\alpha_\rho+\beta_\rho}$$

so that σ lies in $A_{(\alpha+\beta)}$. This means that $\tau\omega$ lies in $A_{(\alpha+\beta)}$.

This lets us check that $\bigoplus_{\alpha\in\mathbb{N}^{\mathcal{C}}_{\mathrm{fin}}} A_{(\alpha)}$ exhaust A. It suffices to check that any σ in Σ lies in some $A_{(\alpha)}$. Proceed by induction on $\deg(\sigma)$, with the case $\sigma = 1$ being trivial; the element 1 always lies in Σ, and hence lies in $A_{(\alpha)}$ for $\alpha = 0$. For σ lying in I, by Proposition 3.3, one either has $(\sigma, a) \neq 0$ for some a in I^2, or else σ lies in $(I^2)^\perp = \mathfrak{p}$, so that σ is in $\mathcal{C}$ and we are done. If $(\sigma, a) \neq 0$ with a in I^2, then σ appears in the support of some $\mathbb{Z}$-linear combination of elements $\tau\omega$ where τ, ω lie in Σ and have strictly smaller degree. There exists at least one such pair τ, ω for which $(\sigma, \tau\omega) \neq 0$, and therefore $\sigma \leq \tau\omega$. Then by induction τ, ω lie in some $A_{(\alpha)}, A_{(\beta)}$, respectively, so $\tau\omega$ lies in $A_{(\alpha+\beta)}$, and hence σ lies in $A_{(\alpha+\beta)}$ also.

Self-duality shows that (3.2) implies (3.3): if a, b, c lie in $A_{(\alpha)}, A_{(\beta)}, A_{(\gamma)}$, respectively, then $(\Delta a, b \otimes c)_{A\otimes A} = (a, bc)_A = 0$ unless $\alpha = \beta + \gamma$. $\square$

Proposition 3.15. *For α, β in $\mathbb{N}^{\mathcal{C}}_{\mathrm{fin}}$ with disjoint support, one has a bijection*

$$\begin{array}{ccc} \Sigma(\alpha) \times \Sigma(\beta) & \longrightarrow & \Sigma(\alpha+\beta) \\ (\sigma, \tau) & \longmapsto & \sigma\tau. \end{array}$$

Thus, the multiplication map $A_{(\alpha)} \otimes A_{(\beta)} \to A_{(\alpha+\beta)}$ is an isomorphism.

Proof. We first check that for σ_1, σ_2 in $\Sigma(\alpha)$ and τ_1, τ_2 in $\Sigma(\beta)$, one has

$$(\sigma_1\tau_1, \sigma_2\tau_2) = \delta_{(\sigma_1,\tau_1),(\sigma_2,\tau_2)}. \tag{3.4}$$

Note that this is equivalent to showing both

- that $\sigma\tau$ lie in $\Sigma(\alpha+\beta)$ so that the map is well-defined, since it shows $(\sigma\tau, \sigma\tau) = 1$, and
- that the map is injective.

One calculates

$$\begin{aligned} (\sigma_1\tau_1, \sigma_2\tau_2)_A &= (\sigma_1\tau_1, m(\sigma_2 \otimes \tau_2))_A \\ &= (\Delta(\sigma_1\tau_1), \sigma_2 \otimes \tau_2)_{A\otimes A} \\ &= (\Delta(\sigma_1)\Delta(\tau_1), \sigma_2 \otimes \tau_2)_{A\otimes A} \end{aligned}$$

Note that due to (3.3), $\Delta(\sigma_1)\Delta(\tau_1)$ lies in $\sum A_{(\alpha'+\beta')} \otimes A_{(\alpha''+\beta'')}$ where

$$\begin{aligned} \alpha' + \alpha'' &= \alpha \\ \beta' + \beta'' &= \beta. \end{aligned}$$

Since $\sigma_2 \otimes \tau_2$ lies in $A_{(\alpha)} \otimes A_{(\beta)}$, the only nonvanishing terms in the inner product come from those with

$$\begin{aligned} \alpha' + \beta' &= \alpha \\ \alpha'' + \beta'' &= \beta. \end{aligned}$$

As α, β have disjoint support, this can only happen if

$$\alpha' = \alpha, \ \alpha'' = 0, \ \beta' = 0, \ \beta'' = \beta,$$

that is, the only nonvanishing term comes from $(\sigma_1 \otimes 1)(1 \otimes \tau_1) = \sigma_1 \otimes \tau_1$. Hence

$$(\sigma_1\tau_1, \sigma_2\tau_2)_A = (\sigma_1 \otimes \tau_1, \sigma_2 \otimes \tau_2)_{A\otimes A} = \delta_{(\sigma_1,\tau_1),(\sigma_2,\tau_2)}.$$

To see that the map is surjective, express

$$\prod_{\rho\in\mathcal{C}} \rho^{\alpha_\rho} = \sum_i \sigma_i$$
$$\prod_{\rho\in\mathcal{C}} \rho^{\beta_\rho} = \sum_j \tau_j$$

with $\sigma_i \in \Sigma(\alpha)$ and τ_j in $\Sigma(\beta)$. Then each product $\sigma_i\tau_j$ is in $\Sigma(\alpha+\beta)$ by (3.4), and

$$\prod_{\rho\in\mathcal{C}} \rho^{\alpha_\rho+\beta_\rho} = \sum_{i,j} \sigma_i\tau_j$$

shows that $\{\sigma_i\tau_j\}$ exhausts $\Sigma(\alpha+\beta)$. This gives surjectivity. $\square$

Proof of Theorem 3.10. Recall from Definition 3.9 that for each ρ in $\mathcal{C}$, one defines $A(\rho) \subset A$ to be the $\mathbb{Z}$-span of

$$\Sigma(\rho) := \{\sigma \in \Sigma : \text{ there exists } n \geq 0 \text{ with } (\sigma, \rho^n) \neq 0\}.$$

In other words, $A(\rho) := \bigoplus_{n\geq 0} A_{(n\cdot e_\rho)}$ where e_ρ in $\mathbb{N}^{\mathcal{C}}_{\text{fin}}$ is the standard basis element indexed by ρ. Proposition 3.14 then shows that $A(\rho)$ is a subHopf algebra of A. Since every α in $\mathbb{N}^{\mathcal{C}}_{\text{fin}}$ can be expressed as the (finite) sum $\sum_\rho \alpha_\rho e_\rho$, and the e_ρ have disjoint support, iterating Proposition 3.15 shows that $A = \bigotimes_{\rho\in\mathcal{C}} A(\rho)$. Lastly, $\Sigma(\rho)$ is clearly a PSH-basis for $A(\rho)$, and if σ is any primitive element in $\Sigma(\rho)$ then $(\sigma,\rho^n) \neq 0$ lets one conclude via Lemma 3.11 that $\sigma = \rho$ (and $n = 1$). $\square$

3.3. **Λ is the unique indecomposable PSH.** The goal here is to prove the rest of Zelevinsky's structure theory for PSH's. Namely, if A has only one primitive element ρ in its PSH-basis Σ, then A must be isomorphic as a PSH to the ring of symmetric functions Λ, after one rescales the grading of A. Note that every σ in Σ has $\sigma \leq \rho^n$ for some n, and hence has degree divisible by the degree of ρ. Thus one can divide all degrees by that of ρ and assume ρ has degree 1.

The idea is to find within A and Σ a set of elements that play the role of

$$\{h_n = s_{(n)}\}_{n=0,1,2,\ldots}, \{e_n = s_{(1^n)}\}_{n=0,1,2,\ldots}$$

within $A = \Lambda$ and its PSH-basis of Schur functions $\Sigma = \{s_\lambda\}$. Zelevinsky's argument does this by isolating some properties that turn out to characterize these elements:

(a) $h_0 = e_0 = 1$, and $h_1 = e_1 =: \rho$ has ρ^2 a sum of two elements of Σ, namely

$$\rho^2 = h_2 + e_2.$$

(b) For all $n = 0, 1, 2, \ldots$, there exist unique elements h_n, e_n in $A_n \cap \Sigma$ that satisfy

$$h_2^\perp e_n = 0,$$
$$e_2^\perp h_n = 0$$

with h_2, e_2 being the two elements of Σ introduced in (a).

(c) For $k = 0, 1, 2, \ldots, n$ one has

$$h_k^\perp h_n = h_{n-k} \text{ and } \sigma^\perp h_n = 0 \text{ for } \sigma \in \Sigma \setminus \{h_0, h_1, \ldots, h_n\}$$
$$e_k^\perp e_n = e_{n-k} \text{ and } \sigma^\perp e_n = 0 \text{ for } \sigma \in \Sigma \setminus \{e_0, e_1, \ldots, e_n\}.$$

In particular, $e_k^\perp h_n = 0 = h_k^\perp e_n$ for $k \geq 2$.

(d) Their coproducts are

$$\Delta(h_n) = \sum_{i+j=n} h_i \otimes h_j,$$
$$\Delta(e_n) = \sum_{i+j=n} e_i \otimes e_j.$$

We will prove Zelevinsky's result [81, Theorem 3.1] as a combination of the following two theorems.

Theorem 3.16. *Let A be a PSH with PSH-basis Σ containing only one primitive ρ, and assume that the grading has been rescaled so that ρ has degree 1. Then, after renaming $\rho = e_1 = h_1$, one can find unique sequences $\{h_n\}_{n=0,1,2,\ldots}, \{e_n\}_{n=0,1,2,\ldots}$ of elements of Σ having properties (a),(b),(c),(d) listed above.*

The second theorem uses the following notion.

Definition 3.17. A *PSH-morphism* $A \overset{\varphi}{\to} A'$ between two PSH's A, A' having PSH-bases Σ, Σ' is a graded Hopf algebra morphism for which $\varphi(\mathbb{N}\Sigma) \subset \mathbb{N}\Sigma'$. If $A = A'$ and $\Sigma = \Sigma'$ it will be called a *PSH-endomorphism*. If φ is an isomorphism and restricts to a bijection $\Sigma \to \Sigma'$, it will be called a *PSH-isomorphism*; if it is both an isomorphism and an endomorphism, it is a *PSH-automorphism*.

Theorem 3.18. *The elements $\{h_n\}_{n=0,1,2,\ldots}, \{e_n\}_{n=0,1,2,\ldots}$ in Theorem 3.16 also satisfy the following.*

(e) *The elements h_n, e_n in A satisfy the same relation (2.11)*

$$\sum_{i+j=n} (-1)^i e_i h_j = \delta_{0,n}.$$

as their counterparts in Λ, along with the property that

$$A = \mathbb{Z}[h_1, h_2, \ldots] = \mathbb{Z}[e_1, e_2, \ldots].$$

(f) *There is exactly one nontrivial automorphism $A \overset{\omega}{\to} A$ as a PSH, swapping $h_n \leftrightarrow e_n$.*

(g) *There are exactly two PSH-isomorphisms $A \to \Lambda$,*
- *one sending h_n to the complete homogeneous symmetric functions $h_n(\mathbf{x})$, while sending e_n to the elementary symmetric functions $e_n(\mathbf{x})$,*
- *the second one (obtained by composing the first with ω) sending $h_n \mapsto e_n(\mathbf{x})$ and $e_n \mapsto h_n(\mathbf{x})$.*

Before embarking on the proof, we mention one more bit of convenient terminology: say that an element σ in Σ is a *constituent* of a in $\mathbb{N}\Sigma$ when $\sigma \leq a$, that is, σ appears with nonzero coefficient c_σ in the unique expansion $a = \sum_{\tau\in\Sigma} c_\tau \tau$.

Proof of Theorem 3.16. One fact that occurs frequently is this:

$$\text{Every } \sigma \text{ in } \Sigma \cap A_n \text{ is a constituent of } \rho^n. \tag{3.5}$$

This follows from Theorem 3.10, since ρ is the only primitive element of Σ: one has $A = A(\rho)$ and $\Sigma = \Sigma(\rho)$, so that σ is a constituent of some ρ^m, and homogeneity considerations force $m = n$.

Assertion (a). Note that

$$(\rho^2, \rho^2) = (\rho^\perp(\rho^2), \rho) = (2\rho, \rho) = 2$$

using the fact that $\rho^\perp$ is a derivation since ρ is primitive (Proposition 2.47(iii)). On the other hand, expressing $\rho^2 = \sum_{\sigma\in\Sigma} c_\sigma \sigma$ with c_σ in $\mathbb{N}$, one has $(\rho^2, \rho^2) = \sum_\sigma c_\sigma^2$. Hence exactly two of the $c_\sigma = 1$, so ρ^2 has exactly two distinct constituents. Denote them by h_2 and e_2. One concludes that $\Sigma \cap A_2 = \{h_2, e_2\}$ from (3.5).

Note also that the same argument shows $\Sigma \cap A_1 = \{\rho\}$, so that $A_1 = \mathbb{Z}\rho$. Since $\rho^\perp h_2$ lies in $A_1 = \mathbb{Z}\rho$ and $(\rho^\perp h_2, \rho) = (h_2, \rho^2) = 1$, we have $\rho^\perp h_2 = \rho$. Similarly $\rho^\perp e_2 = \rho$.

Assertion (b). We will show via induction on n the following three assertions for $n \geq 1$:

(3.6)
- There exists an element h_n in $\Sigma \cap A_n$ with $e_2^\perp h_n = 0$.
- This element h_n is unique.
- Furthermore $\rho^\perp h_n = h_{n-1}$.

In the base cases $n = 1, 2$, it is not hard to check that our previously labelled elements, h_1, h_2 (namely $h_1 := \rho$, and h_2 as named in part (a)) really *are* the unique elements satisfying these hypotheses.

In the inductive step, it turns out that we will find h_n as a constituent of ρh_{n-1}. Thus we again use the derivation property of $\rho^\perp$ to compute that ρh_{n-1} has exactly two constituents:

$$\begin{aligned}
(\rho h_{n-1}, \rho h_{n-1}) &= (\rho^\perp(\rho h_{n-1}), h_{n-1}) \\
&= (h_{n-1} + \rho \cdot \rho^\perp h_{n-1}, h_{n-1}) \\
&= (h_{n-1} + \rho h_{n-2}, h_{n-1}) \\
&= 1 + (h_{n-2}, \rho^\perp h_{n-1}) \\
&= 1 + (h_{n-2}, h_{n-2}) = 1 + 1 = 2
\end{aligned}$$

where the inductive hypothesis $\rho^\perp h_{n-1} = h_{n-2}$ was used twice. We next show that exactly one of the two constituents of ρh_{n-1} is annihilated by $e_2^\perp$. Note that since e_2 lies in A_2, and A_1 has $\mathbb{Z}$-basis element ρ, there is a constant c in $\mathbb{Z}$ such that

$$\Delta(e_2) = e_2 \otimes 1 + c\rho \otimes \rho + 1 \otimes e_2. \tag{3.7}$$

On the other hand, (a) showed

$$1 = (e_2, \rho^2)_A = (\Delta(e_2), \rho \otimes \rho)_{A \otimes A}$$

so one must have $c = 1$. Therefore by Proposition 2.47(iii) again,

$$\begin{array}{rcccccl} e_2^\perp(\rho h_{n-1}) & = & e_2^\perp(\rho)h_{n-1} & + & \rho^\perp(\rho)\rho^\perp(h_{n-1}) & + & \rho e_2^\perp(h_{n-1}) \\ & = & 0 & + & h_{n-2} & + & 0 \\ & = & h_{n-2} & & & & . \end{array} \tag{3.8}$$

where the first term vanished due to degree considerations and the last term vanished by the inductive hypothesis. Bearing in mind that ρh_{n-1} lies in $\mathbb{N}\Sigma$, and in a PSH with PSH-basis Σ, any skewing operator $\sigma^\perp$ for σ in Σ will preserve $\mathbb{N}\Sigma$, one concludes from (3.8) that

- one of the two distinct constituents of the element ρh_{n-1} must be sent by $e_2^\perp$ to h_{n-2}, and
- the other constituent of ρh_{n-1} must be annihilated by $e_2^\perp$; call this second constituent h_n.

Lastly, to see that this h_n is unique, it suffices to show that any element σ of $\Sigma \cap A_n$ which is killed by $e_2^\perp$ must be a constituent of ρh_{n-1}. This holds for the following reason. We know $\sigma \leq \rho^n$ by (3.5), and hence $0 \neq (\rho^n, \sigma) = (\rho^{n-1}, \rho^\perp \sigma)$, implying that $\rho^\perp \sigma \neq 0$. On the other hand, since $0 = \rho^\perp e_2^\perp \sigma = e_2^\perp \rho^\perp \sigma$, one has that $\rho^\perp \sigma$ is annihilated by $e_2^\perp$, and hence $\rho^\perp \sigma$ must be a (positive) multiple of h_{n-1} by part of our inductive hypothesis. Therefore $(\sigma, \rho h_{n-1}) = (\rho^\perp \sigma, h_{n-1})$ is positive, that is, σ is a constituent of ρh_{n-1}.

The preceding argument, applied to $\sigma = h_n$, shows that $\rho^\perp h_n = c h_{n-1}$ for some c in $\{1, 2, \ldots\}$. Since $(\rho^\perp h_n, h_{n-1}) = (h_n, \rho h_{n-1}) = 1$, this c must be 1, so that $\rho^\perp h_n = h_{n-1}$. This completes the induction step in the proof of (3.6).

One can then argue, swapping the roles of e_n, h_n in the above argument, the existence and uniqueness of a sequence $\{e_n\}_{n=0}^\infty$ in Σ satisfying the properties analogous to (3.6), with $e_0 := 1, e_1 := \rho$.

Assertion (c). Iterating the property from (b) that $\rho^\perp h_n = h_{n-1}$ shows that $(\rho^k)^\perp h_n = h_{n-k}$ for $0 \leq k \leq n$. However one also has an expansion

$$\rho^k = c h_k + \sum_{\substack{\sigma \in \Sigma \cap A_k: \\ \sigma \neq h_k}} c_\sigma \sigma$$

for some integers $c, c_\sigma > 0$, since every σ in $\Sigma \cap A_k$ is a constituent of ρ^k. Hence

$$1 = (h_{n-k}, h_{n-k}) = ((\rho^k)^\perp h_n, (\rho^k)^\perp h_n) \geq c^2 (h_k^\perp h_n, h_k^\perp h_n)$$

using Proposition 2.47(ii). Hence if we knew that $h_k^\perp h_n \neq 0$ this would force

$$h_k^\perp h_n = (\rho^k)^\perp h_n = h_{n-k}$$

as well as $\sigma^\perp h_n = 0$ for all $\sigma \notin \{h_0, h_1, \ldots, h_n\}$. But

$$(\rho^{n-k})^\perp h_k^\perp h_n = h_k^\perp (\rho^{n-k})^\perp h_n = h_k^\perp h_k = 1 \neq 0$$

so $h_k^\perp h_n \neq 0$, as desired. The argument for $e_k^\perp e_n = e_{n-k}$ is symmetric.

The last assertion in (c) follows if one checks that $e_n \neq h_n$ for each $n \geq 2$, but this holds since $e_2^\perp(h_n) = 0$ but $e_2^\perp(e_n) = e_{n-2}$.

Assertion (d). Part (c) implies that

$$(\Delta h_n, \sigma \otimes \tau)_{A \otimes A} = (h_n, \sigma\tau)_A = (\sigma^\perp h_n, \tau)_A = 0$$

unless $\sigma = h_k$ for some $k = 0, 1, 2, \ldots, n$ and $\tau = h_{n-k}$. Also one can compute

$$(\Delta h_n, h_k \otimes h_{n-k}) = (h_n, h_k h_{n-k}) = (h_k^\perp h_n, h_{n-k}) \overset{(c)}{=} (h_{n-k}, h_{n-k}) = 1.$$

This is equivalent to the assertion for Δh_n in (d). The argument for Δe_n is symmetric. $\square$

Before proving Theorem 3.18, we note some consequences of Theorem 3.16. Define for each partition $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell)$ the elements of A

$$h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots ,$$
$$e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots .$$

Also, define the *lexicographic order* on Par_n by saying $\lambda <_{\mathrm{lex}} \mu$ if $\lambda \neq \mu$ and the smallest index i for which $\lambda_i \neq \mu_i$ has $\lambda_i < \mu_i$. Recall also that λ^t denotes the *conjugate* or *transpose* partition to λ, obtained by swapping rows and columns in the Ferrers diagram.

The following unitriangularity lemma will play a role in the proof of Theorem 3.18(e).

Lemma 3.19. *Under the hypotheses of Theorem 3.16, for λ, μ in Par_n, one has*

$$e_\mu^\perp h_\lambda = \begin{cases} 1 & \text{if } \mu = \lambda^t \\ 0 & \text{if } \mu >_{\mathrm{lex}} \lambda^t. \end{cases}$$

Consequently

$$\det \left[(e_{\mu^t}, h_\lambda) \right]_{\lambda, \mu \in \mathrm{Par}_n} = \pm 1.$$

Proof. Induct on the length of μ. If λ has length ℓ, so that $\lambda_1^t = \ell$, then

$$\begin{aligned} e_\mu^\perp h_\lambda &= e_{(\mu_2,\mu_3,\ldots)}^\perp \left(e_{\mu_1}^\perp (h_{\lambda_1} \cdots h_{\lambda_\ell}) \right) \\ &= e_{(\mu_2,\mu_3,\ldots)}^\perp \sum_{i_1+\ldots+i_\ell=\mu_1} e_{i_1}^\perp (h_{\lambda_1}) \cdots e_{i_\ell}^\perp (h_{\lambda_\ell}) \\ &= \begin{cases} 0 & \text{if } \mu_1 > \ell = \lambda_1^t \\ e_{(\mu_2,\mu_3,\ldots)}^\perp h_{(\lambda_1-1,\ldots,\lambda_\ell-1)} & \text{if } \mu_1 = \ell = \lambda_1^t \end{cases} \end{aligned}$$

where the second-to-last equality used Proposition 2.47(iii) along with (the iterates of) the coproduct formula for $\Delta(e_n)$ in Theorem 3.16(d), and the last equality used

$$e_k^\perp (h_n) = \begin{cases} h_{n-1} & \text{if } k = 1, \\ 0 & \text{if } k \geq 2 \end{cases}$$

Now apply induction, since $(\lambda_1 - 1, \ldots, \lambda_\ell - 1)^t = (\lambda_2^t, \lambda_3^t, \ldots)$.

For the last assertion, note $(e_{\mu^t}, h_\lambda) = (e_{\mu^t}^\perp (h_\lambda), 1) = e_{\mu^t}^\perp (h_\lambda)$ for λ, μ in Par_n. ∎

The following proposition will be the crux of the proof of Theorem 3.18(f) and (g), and turns out to be closely related to Kerov's *asymptotic theory of characters of the symmetric groups* [39].

Proposition 3.20. *Given a PSH A with PSH-basis Σ containing only one primitive ρ, the two maps $A \to \mathbb{Z}$ defined on $A = \bigoplus_{n \geq 0} A_n$ via*

$$\delta_h = \bigoplus_n h_n^\perp ,$$
$$\delta_e = \bigoplus_n e_n^\perp$$

are characterized as the only two $\mathbb{Z}$-linear maps $A \xrightarrow{\delta} \mathbb{Z}$ with the three properties of being

- **positive**: $\delta(\mathbb{N}\Sigma) \subset \mathbb{N}$,
- **multiplicative**: $\delta(a_1 a_2) = \delta(a_1)\delta(a_2)$, *and*
- **normalized**: $\delta(\rho) = 1$.

Proof. It should be clear from their definitions that δ_h, δ_e are $\mathbb{Z}$-linear, positive and normalized. To see that δ_h is multiplicative, by $\mathbb{Z}$-linearity, it suffices to check that for a_1, a_2 in A_{n_1}, A_{n_2} with $n_1 + n_2 = n$, one has

$$\delta_h(a_1 a_2) = h_n^\perp (a_1 a_2) = \sum_{i_1+i_2=n} h_{i_1}^\perp (a_1) h_{i_2}^\perp (a_2) = h_{n_1}^\perp (a_1) h_{n_2}^\perp (a_2) = \delta_h(a_1)\delta_h(a_2)$$

in which the second equality used Proposition 2.47(iii) and Theorem 3.16(d). The argument for δ_e is symmetric.

Conversely, given $A \xrightarrow{\delta} \mathbb{Z}$ which is $\mathbb{Z}$-linear, positive, multiplicative, and normalized, note that

$$\delta(h_2) + \delta(e_2) = \delta(h_2 + e_2) = \delta(\rho^2) = \delta(\rho)^2 = 1^2 = 1$$

and hence positivity implies that either $\delta(h_2) = 0$ or $\delta(e_2) = 0$. Assume the latter holds, and we will show that $\delta = \delta_h$.

Given any σ in $\Sigma \cap A_n \setminus \{h_n\}$, note that $e_2^\perp \sigma \neq 0$ by Theorem 3.16(b), and hence $0 \neq (e_2^\perp \sigma, \rho^{n-2}) = (\sigma, e_2 \rho^{n-2})$. Thus σ is a constituent of $e_2 \rho^{n-2}$, so positivity implies

$$0 \leq \delta(\sigma) \leq \delta(e_2 \rho^{n-2}) = \delta(e_2)\delta(\rho^{n-2}) = 0.$$

Thus $\delta(\sigma) = 0$ for σ in $\Sigma \cap A_n \setminus \{h_n\}$. Since $\delta(\rho^n) = \delta(\rho)^n = 1^n = 1$, this forces $\delta(h_n) = 1$, for each $n \geq 0$ (including $n = 0$, as $1 = \delta(\rho) = \delta(\rho \cdot 1) = \delta(\rho)\delta(1) = 1 \cdot \delta(1) = \delta(1)$.) Thus $\delta = \delta_h$. The argument when $\delta(h_2) = 0$ showing $\delta = \delta_e$ is symmetric. □

Proof of Theorem 3.18. Many of the assertions of parts (e) and (f) will come from constructing the unique nontrivial PSH-automorphism ω of A from the antipode S: for homogeneous a in A_n, define $\omega(a) := (-1)^n S(a)$. We now study some of the properties of S and ω.

Since A is a PSH, it is commutative by Theorem 3.7. This implies both that S, ω are actually algebra endomorphisms by Proposition 1.26, and that $S^2 = 1_A = \omega^2$ by Corollary 1.28.

Since A is self-dual and the defining diagram (1.11) satisfied by the antipode S is sent to itself when one replaces A by A^o and all maps by their adjoints, one concludes that $S = S^t$, i.e., S is self-adjoint. Since S is an algebra endomorphism, and $S = S^t$, in fact S is also a coalgebra endomorphism, a bialgebra endomorphism, and a Hopf endomorphism (by Proposition 1.35(c)) The same properties are shared by ω.

Since $1_A = S^2 = SS^t$, one concludes that S is an isometry, and hence so is ω.

Since ρ is primitive, one has $S(\rho) = -\rho$ and $\omega(\rho) = \rho$. Therefore $\omega(\rho^n) = \rho^n$ for $n = 1, 2, \ldots$. Use this as follows to check that ω is a PSH-automorphism, which amounts to checking that every σ in Σ has $\omega(\sigma)$ in Σ:

$$(\omega(\sigma), \omega(\sigma)) = (\sigma, \sigma) = 1$$

so that $\pm\omega(\sigma)$ lies in Σ, but also if σ lies in A_n, then

$$(\omega(\sigma), \rho^n) = (\sigma, \omega(\rho^n)) = (\sigma, \rho^n) > 0.$$

In summary, ω is a PSH-automorphism of A, an isometry, and an involution.

Let us try to determine the action of ω on the $\{h_n\}$. By similar reasoning as in (3.7), one has

$$\Delta(h_2) = h_2 \otimes 1 + \rho \otimes \rho + 1 \otimes h_2.$$

Thus $0 = S(h_2) + S(\rho)\rho + h_2$, and combining this with $S(\rho) = -\rho$, one has $S(h_2) = e_2$. Thus also $\omega(h_2) = (-1)^2 S(h_2) = e_2$.

We claim that this forces $\omega(h_n) = e_n$, because $h_2^\perp \omega(h_n) = 0$ via the following calculation: for any a in A one has

$$\begin{aligned}
(h_2^\perp \omega(h_n), a) &= (\omega(h_n), h_2 a) \\
&= (h_n, \omega(h_2 a)) \\
&= (h_n, e_2 \omega(a)) \\
&= (e_2^\perp h_n, \omega(a)) = (0, \omega(a)) = 0.
\end{aligned}$$

Consequently the involution ω swaps h_n and e_n, while the antipode S has $S(h_n) = (-1)^n e_n$ and $S(e_n) = (-1)^n h_n$. Thus the coproduct formulas in (d) and definition of the antipode S imply the relation (2.11) between $\{h_n\}$ and $\{e_n\}$.

This relation (2.11) also lets one recursively express the h_n as polynomials with integer coefficients in the $\{e_n\}$, and vice-versa, so that $\{h_n\}$ and $\{e_n\}$ each generate the same $\mathbb{Z}$-subalgebra A' of A. We wish to show that A' exhausts A.

We argue that Lemma 3.19 implies the *Gram matrix* $[(h_\mu, h_\lambda)]_{\mu,\lambda \in \mathrm{Par}_n}$ has determinant ± 1 as follows. Since $\{h_n\}$ and $\{e_n\}$ both generate A', there exists a $\mathbb{Z}$-matrix $(a_{\mu,\lambda})$ expressing $e_{\mu^t} = \sum_\lambda a_{\mu,\lambda} h_\lambda$, and one has

$$[(e_{\mu^t}, h_\lambda)] = [a_{\mu,\lambda}] \cdot [(h_\mu, h_\lambda)].$$

Taking determinants of these three $\mathbb{Z}$-matrices, and using the fact that the determinant on the left is ± 1, both determinants on the right must also be ± 1.

Now we will show that every $\sigma \in \Sigma \cap A_n$ lies in A'_n. Uniquely express $\sigma = \sigma' + \sigma''$ in which σ' lies in the $\mathbb{R}$-span $\mathbb{R}A'_n$ and σ'' lies in the real perpendicular space $(\mathbb{R}A'_n)^\perp$ inside $\mathbb{R} \otimes_{\mathbb{Z}} A_n$. One can compute $\mathbb{R}$-coefficients $(c_\mu)_{\mu \in \mathrm{Par}_n}$ that express $\sigma' = \sum_\mu c_\mu h_\mu$ by solving the system

$$\left(\sum_\mu c_\mu h_\mu, h_\lambda \right) = (\sigma, h_\lambda) \text{ for } \lambda \in \mathrm{Par}_n .$$

This linear system is governed by the Gram matrix $[(h_\mu, h_\lambda)]_{\mu,\lambda \in \mathrm{Par}_n}$ with determinant ± 1, and its right side has $\mathbb{Z}$-entries since σ, h_λ lie in A. Hence the solution $(c_\mu)_{\mu \in \mathrm{Par}_n}$ will have $\mathbb{Z}$-entries, so σ' lies in A'. Furthermore, $\sigma'' = \sigma - \sigma'$ will lie in A, and hence by the orthogonality of σ', σ'',

$$1 = (\sigma, \sigma) = (\sigma', \sigma') + (\sigma'', \sigma'').$$

One concludes that either $\sigma'' = 0$, or $\sigma' = 0$. The latter cannot occur since it would mean that $\sigma = \sigma''$ is perpendicular to all of A'. But $\rho^n = h_1^n$ lies in A', and $(\sigma, \rho^n) \neq 0$. Thus $\sigma'' = 0$, meaning $\sigma = \sigma'$ lies in A'. This completes the proof of assertion (e). Note that in the process, having shown $\det(h_\mu, h_\lambda)_{\lambda,\mu \in \mathrm{Par}_n} = \pm 1$, one also knows that $\{h_\lambda\}_{\lambda \in \mathrm{Par}_n}$ are $\mathbb{Z}$-linearly independent, so that $\{h_1, h_2, \ldots\}$ are algebraically independent, and $A = \mathbb{Z}[h_1, h_2, \ldots]$ is the polynomial algebra generated by $\{h_1, h_2, \ldots\}$.

For assertion (f), we have seen that ω gives such a PSH-automorphism $A \to A$, swapping $h_n \leftrightarrow e_n$. Conversely, given a PSH-automorphism $A \xrightarrow{\varphi} A$, consider the positive, multiplicative, normalized $\mathbb{Z}$-linear map $\delta := \delta_h \circ \varphi : A \to \mathbb{Z}$. Proposition 3.20 shows that either

- $\delta = \delta_h$, which then forces $\varphi(h_n) = h_n$ for all n, so $\varphi = 1_A$, or
- $\delta = \delta_e$, which then forces $\varphi(e_n) = h_n$ for all n, so $\varphi = \omega$.

For assertion (g), given a PSH A with PSH-basis Σ having exactly one primitive ρ, since we have seen $A = \mathbb{Z}[h_1, h_2, \ldots]$, where h_n in A is as defined in Theorem 3.16, one can uniquely define an algebra morphism $A \xrightarrow{\varphi} \Lambda$ that sends the element h_n to the complete homogeneous symmetric function $h_n(\mathbf{x})$. Assertions (d) and (e) show that φ is a bialgebra isomorphism, and hence it is a Hopf isomorphism. To show that it is a PSH-isomorphism, we first note that it is an isometry because one can iterate Proposition 2.47(iii) together with assertions (c) and (d) to compute all inner products

$$(h_\mu, h_\lambda)_A = (1, h_\mu^\perp h_\lambda)_A = (1, h_{\mu_1}^\perp h_{\mu_2}^\perp \cdots (h_{\lambda_1} h_{\lambda_2} \cdots))_A$$

for μ, λ in Par_n. Hence

$$(h_\mu, h_\lambda)_A = (h_\mu(\mathbf{x}), h_\lambda(\mathbf{x}))_\Lambda = (\varphi(h_\mu), \varphi(h_\lambda))_\Lambda$$

Once one knows φ is an isometry, then elements ω in $\Sigma \cap A_n$ are characterized in terms of the form $(\cdot, \cdot)$ by $(\omega, \omega) = 1$ and $(\omega, \rho^n) > 0$. Hence φ sends each σ in Σ to a Schur function s_λ, and is a PSH-isomorphism. $\square$

4. Complex representations for $\mathfrak{S}_n$, wreath products, $GL_n(\mathbb{F}_q)$

After reviewing the basics that we will need from representation and character theory of finite groups, we give Zelevinsky's three main examples of PSH's arising as spaces of virtual characters for three towers of finite groups:

- *symmetric* groups,
- their *wreath products* with any finite group, and
- the finite *general linear* groups.

4.1. **Review of complex character theory.**

A good source for this material, including the crucial Mackey formula, is Serre [66, Chaps. 1-7].

4.1.1. *Basic definitions, Maschke, Schur.* For a group G, a *representation* is a homomorphism $G \xrightarrow{\varphi} GL(V)$ for some vector space V over a field. We will take the field to be $\mathbb{C}$ from now on, and we will also assume that V is finite-dimensional over $\mathbb{C}$. Thus a representation of G is the same as a finite-dimensional (left-)$\mathbb{C}G$-module V.

We also assume that G is finite, so that Maschke's Theorem[13] says that $\mathbb{C}G$ is semisimple, meaning that every $\mathbb{C}G$-module $U \subset V$ has a $\mathbb{C}G$-module complement U' with $V = U \oplus U'$. Equivalently, *indecomposable* $\mathbb{C}G$-modules are the same thing as *simple* (=*irreducible*) $\mathbb{C}G$-modules.

Schur's Lemma implies that for two simple $\mathbb{C}G$-modules V_1, V_2, one has

$$\operatorname{Hom}_{\mathbb{C}G}(V_1, V_2) \cong \begin{cases} \mathbb{C} & \text{if } V_1 \cong V_2, \\ 0 & \text{if } V_1 \not\cong V_2. \end{cases}$$

4.1.2. *Characters and Hom spaces.* A $\mathbb{C}G$-module V is completely determined up to isomorphism by its *character*

$$\begin{array}{ccc} G & \xrightarrow{\chi_V} & \mathbb{C} \\ g & \longmapsto & \chi_V(g) := \operatorname{trace}(g : V \to V). \end{array}$$

This character χ_V is a *class function*, meaning it is constant on G-conjugacy classes. The space $R_{\mathbb{C}}(G)$ of class functions $G \to \mathbb{C}$ has a Hermitian, positive definite form

$$(f_1, f_2)_G := \frac{1}{|G|} \sum_{g \in G} f_1(g) \overline{f_2(g)}.$$

For any two $\mathbb{C}G$-modules V_1, V_2,

$$(\chi_{V_1}, \chi_{V_2})_G = \dim_{\mathbb{C}} \operatorname{Hom}_{\mathbb{C}G}(V_1, V_2). \tag{4.1}$$

The set of all *irreducible characters*

$$\operatorname{Irr}(G) = \{\chi_V : V \text{ is a simple } \mathbb{C}G\text{-module}\}$$

forms an orthonormal basis of $R_{\mathbb{C}}(G)$ with respect to this form, and spans a $\mathbb{Z}$-sublattice

$$R(G) := \mathbb{Z} \operatorname{Irr}(G) \subset R_{\mathbb{C}}(G)$$

sometimes called the *virtual characters* of G.

Instead of working with the Hermitian form $(\cdot, \cdot)_G$ on G, we could also (and some authors do) define a $\mathbb{C}$-bilinear form $\langle \cdot, \cdot \rangle_G$ on $R_{\mathbb{C}}(G)$ by

$$\langle f_1, f_2 \rangle_G := \frac{1}{|G|} \sum_{g \in G} f_1(g) f_2(g^{-1}).$$

This form is not identical with $(\cdot, \cdot)_G$ (indeed, $\langle \cdot, \cdot \rangle_G$ is bilinear while $(\cdot, \cdot)_G$ is Hermitian), but it still satisfies (4.1), and thus is identical with $(\cdot, \cdot)_G$ on $R(G) \times R(G)$. Hence, for all we are going to do, we could just as well use the form $\langle \cdot, \cdot \rangle_G$ instead of $(\cdot, \cdot)_G$.

[13]... which has a beautiful generalization to finite-dimensional Hopf algebras due to Larson and Sweedler; see Montgomery [56, §2.2].

4.1.3. *Tensor products.* Given two groups G_1, G_2 and $\mathbb{C}G_i$-modules V_i for $i = 1, 2$, their tensor product $V_1 \otimes_{\mathbb{C}} V_2$ becomes a $\mathbb{C}[G_1 \times G_2]$-module via $(g_1, g_2)(v_1 \otimes v_2) = g_1(v_1) \otimes g_2(v_2)$. When V_1, V_2 are both simple, then so is $V_1 \otimes V_2$, and all simple $\mathbb{C}[G_1 \times G_2]$-modules arise this way. Thus one has identifications and isomorphisms

$$\begin{aligned} \operatorname{Irr}(G_1 \times G_2) &= \operatorname{Irr}(G_1) \times \operatorname{Irr}(G_2) \\ R(G_1 \times G_2) &\cong R(G_1) \otimes_{\mathbb{Z}} R(G_2); \end{aligned}$$

here, $\chi_{V_1} \otimes \chi_{V_2} \in R(G_1) \otimes_{\mathbb{Z}} R(G_2)$ is being identified with $\chi_{V_1 \otimes V_2} \in R(G_1 \times G_2)$ for all $\mathbb{C}G_1$-modules V_1 and all $\mathbb{C}G_2$-modules V_2. Given two $\mathbb{C}G_1$-modules V_1 and W_1 and two $\mathbb{C}G_2$-modules V_2 and W_2, we have

$$(\chi_{V_1 \otimes V_2}, \chi_{W_1 \otimes W_2})_{G_1 \times G_2} = (\chi_{V_1}, \chi_{W_1})_{G_1} (\chi_{V_2}, \chi_{W_2})_{G_2}. \tag{4.2}$$

Exercise 4.1. Let G_1 and G_2 be two groups. Let V_i and W_i be $\mathbb{C}G_i$-modules for every $i \in \{1, 2\}$. Prove that the map

$$\operatorname{Hom}_{\mathbb{C}G_1}(V_1, W_1) \otimes \operatorname{Hom}_{\mathbb{C}G_2}(V_2, W_2) \to \operatorname{Hom}_{\mathbb{C}[G_1 \times G_2]}(V_1 \otimes V_2, W_1 \otimes W_2)$$

sending each tensor $f \otimes g$ to the tensor product $f \otimes g$ of homomorphisms is a vector space isomorphism. Conclude that (4.2) holds.

4.1.4. *Induction and restriction.* Given a subgroup $H < G$ and $\mathbb{C}H$-module U, one can use the fact that $\mathbb{C}G$ is a $(\mathbb{C}G, \mathbb{C}H)$-bimodule to form the *induced $\mathbb{C}G$-module*

$$\operatorname{Ind}_H^G U := \mathbb{C}G \otimes_{\mathbb{C}H} U.$$

The fact that $\mathbb{C}G$ is free as a (right-)$\mathbb{C}H$-module[14] on basis elements $\{t_g\}_{gH \in G/H}$ makes this tensor product easy to analyze. For example one can compute its character

$$\chi_{\operatorname{Ind}_H^G U}(g) = \operatorname{Ind}_H^G(\chi_U)(g) := \frac{1}{|H|} \sum_{\substack{k \in G: \\ kgk^{-1} \in H}} \chi_U(kgk^{-1}). \tag{4.3}$$

One can also recognize when a $\mathbb{C}G$-module V is isomorphic to $\operatorname{Ind}_H^G U$ for some $\mathbb{C}H$-module U: this happens if and only if there is an H-stable subspace $U \subset V$ having the property that $V = \bigoplus_{gH \in G/H} gU$.

It is fairly easy to show that

$$\operatorname{Ind}_{H_1 \times H_2}^{G_1 \times G_2}(U_1 \otimes U_2) \cong \left(\operatorname{Ind}_{H_1}^{G_1} U_1\right) \otimes \left(\operatorname{Ind}_{H_2}^{G_2} U_2\right) \tag{4.4}$$

whenever G_1 and G_2 are two groups, $H_1 < G_1$ and $H_2 < G_2$ are two subgroups, and each U_i is a $\mathbb{C}H_i$-module.

The *restriction* operation $V \mapsto \operatorname{Res}_H^G V$ restricts a $\mathbb{C}G$-module V to a $\mathbb{C}H$-module. *Frobenius reciprocity* asserts the adjointness between Ind_H^G and Res_H^G

$$\operatorname{Hom}_{\mathbb{C}G}(\operatorname{Ind}_H^G U, V) \cong \operatorname{Hom}_{\mathbb{C}H}(U, \operatorname{Res}_H^G V), \tag{4.5}$$

as a special case ($S = A = \mathbb{C}G, R = \mathbb{C}H, B = U, C = V$) of the general *adjoint associativity*

$$\operatorname{Hom}_S(A \otimes_R B, C) \cong \operatorname{Hom}_R(B, \operatorname{Hom}_S(A, C))$$

for S, R two rings, A an (S, R)-bimodule, B a left R-module, C a left S-module. Taking characters in (4.5), we obtain

$$(\operatorname{Ind}_H^G \chi_U, \chi_V)_G = (\chi_U, \operatorname{Res}_H^G \chi_V)_H, \tag{4.6}$$

where restriction $\operatorname{Res}_H^G f$ of an $f \in R_{\mathbb{C}}(G)$ is defined just by restricting the map $f : G \to \mathbb{C}$ to H.

[14]... which also has a beautiful generalization to finite-dimensional Hopf algebras due to Nichols and Zoeller; see [56, §3.1].

4.1.5. *Mackey's formula.* Mackey gave an alternate description of a module which has been induced and then restricted. To state it, for a subgroup $H < G$ and g in G, let $H^g := g^{-1}Hg$ and ${}^gH := gHg^{-1}$. Given a $\mathbb{C}H$-module U, say defined by a homomorphism $H \xrightarrow{\varphi} GL(U)$, let U^g denote the $\mathbb{C}[gHg^{-1}]$-module on the same $\mathbb{C}$-vector space U defined by the composite homomorphism

$$\begin{array}{ccccc} {}^gH & \longrightarrow & H & \xrightarrow{\varphi} & GL(U). \\ h & \longmapsto & g^{-1}hg & & \end{array}$$

Theorem 4.2. *(Mackey's formula) Consider subgroups $H, K < G$, and any $\mathbb{C}H$-module U. If $\{g_1, \ldots, g_t\}$ are double coset representatives for $K\backslash G/H$, then*

$$\operatorname{Res}^G_K \operatorname{Ind}^G_H U \cong \bigoplus_{i=1}^t \operatorname{Ind}^K_{{}^{g_i}H\cap K} \left(\left(\operatorname{Res}^H_{H\cap K^{g_i}} U \right)^{g_i} \right)$$

Proof. In this proof, all tensor product symbols $\otimes$ should be interpreted as $\otimes_{\mathbb{C}H}$. Recall $\mathbb{C}G$ has $\mathbb{C}$-basis $\{t_g\}_{g\in G}$. For subsets $S \subset G$, let $\mathbb{C}[S]$ denote the $\mathbb{C}$-span of $\{t_g\}_{g\in S}$ in $\mathbb{C}G$.

Note that each double coset KgH gives rise to a sub-(K,H)-bimodule $\mathbb{C}[KgH]$ within $\mathbb{C}G$, and one has a $\mathbb{C}K$-module direct sum decomposition

$$\operatorname{Ind}^G_H U = \mathbb{C}G \otimes U = \bigoplus_{i=1}^t \mathbb{C}[Kg_iH] \otimes U.$$

Hence it suffices to check for any element g in G that

$$\mathbb{C}[KgH] \otimes U \cong \operatorname{Ind}^K_{{}^gH\cap K} \left(\left(\operatorname{Res}^H_{H\cap K^g} U \right)^g \right).$$

Note that ${}^gH \cap K$ is the subgroup of K consisting of the elements k in K for which $kgH = gH$. Hence by picking $\{k_1, \ldots, k_s\}$ to be coset representatives for $K/({}^gH \cap K)$, one disjointly decomposes the double coset

$$KgH = \bigsqcup_{j=1}^s k_j({}^gH \cap K)gH,$$

giving a $\mathbb{C}$-vector space direct sum decomposition

$$\begin{aligned} \mathbb{C}[KgH] \otimes U &= \bigoplus_{j=1}^s \mathbb{C}\left[k_j\left({}^gH \cap K\right)gH\right] \otimes U \\ &\cong \operatorname{Ind}^K_{{}^gH\cap K} \left(\mathbb{C}[\left({}^gH \cap K\right)gH] \otimes U\right). \end{aligned}$$

So it remains to check that one has a $\mathbb{C}[{}^gH \cap K]$-module isomorphism

$$\mathbb{C}[\left({}^gH \cap K\right)gH] \otimes U \cong \left(\operatorname{Res}^H_{H\cap K^g} U \right)^g.$$

Bearing in mind that, for each k in ${}^gH \cap K$ and h in H, one has $g^{-1}kg$ in H and hence

$$t_{kgh} \otimes u = t_g \cdot t_{g^{-1}kg\cdot h} \otimes u = t_g \otimes g^{-1}kgh \cdot u$$

one sees that this isomorphism can be defined by mapping

$$t_{kgh} \otimes u \longmapsto g^{-1}kgh \cdot u.$$

$\square$

4.1.6. *Inflation and fixed points.* There are two (adjoint) constructions on representations that apply when one has a normal subgroup $K \triangleleft G$. Given a $\mathbb{C}[G/K]$-module U, say defined by the homomorphism $G/K \xrightarrow{\varphi} GL(U)$, the *inflation* of U to a $\mathbb{C}G$-module $\operatorname{Infl}^G_{G/K} U$ has the same underlying space U, and is defined by the composite homomorphism $G \to G/K \xrightarrow{\varphi} GL(U)$. We will later use the easily-checked fact that when $H < G$ is any other subgroup, one has

$$\operatorname{Res}^G_H \operatorname{Infl}^G_{G/K} U = \operatorname{Infl}^H_{H/H\cap K} \operatorname{Res}^{G/K}_{H/H\cap K} U. \tag{4.7}$$

Inflation turns out to be adjoint to the *K-fixed space construction* sending a $\mathbb{C}G$-module V to the $\mathbb{C}[G/K]$-module

$$V^K := \{v \in V : k(v) = v \text{ for } k \in K\}$$

Note that V^K is indeed a G-stable subspace: for any v in V^K and g in G, one has that $g(v)$ lies in V^K since an element k in K satisfies $kg(v) = g \cdot g^{-1}kg(v) = g(v)$ as $g^{-1}kg$ lies in K. One has this adjointness

$$\mathrm{Hom}_{\mathbb{C}G}(\mathrm{Infl}_{G/K}^{G} U, V) = \mathrm{Hom}_{\mathbb{C}[G/K]}(U, V^K) \tag{4.8}$$

because any $\mathbb{C}G$-module homomorphism φ on the left must have the property that $k\varphi(u) = \varphi(k(u)) = \varphi(u)$ for all k in K, so that φ actually lies on the right. Taking characters in (4.8), we obtain

$$(\mathrm{Infl}_{G/K}^{G} \chi_U, \chi_V)_G = (\chi_U, \chi_{V^K})_{G/K}, \tag{4.9}$$

where inflation $\mathrm{Infl}_{G/K}^{G} f$ of an $f \in R_{\mathbb{C}}(G/K)$ is defined as the composition $G \twoheadrightarrow G/K \xrightarrow{f} \mathbb{C}$.

We will also need the following formula for the character χ_{V^K} in terms of the character χ_V:

$$\chi_{V^K}(gK) = \frac{1}{|K|} \sum_{k \in K} \chi_V(gk). \tag{4.10}$$

To see this, note that when one has a $\mathbb{C}$-linear endomorphism φ on a space V that preserves some $\mathbb{C}$-subspace $W \subset V$, if $V \xrightarrow{\pi} W$ is any idempotent projection onto W, then the trace of the restriction $\varphi|_W$ equals the trace of $\varphi \circ \pi$ on V. Applying this to $W = V^K$ and $\varphi = g$, with $\pi = \frac{1}{|K|} \sum_{k \in K} k$, gives (4.10).

Another way to restate (4.10) is:

$$\chi_{V^K}(gK) = \frac{1}{|K|} \sum_{h \in gK} \chi_V(h). \tag{4.11}$$

There is also an analogue of (4.4):

Lemma 4.3. *Let G_1 and G_2 be two groups, and $K_1 < G_1$ and $K_2 < G_2$ be two respective subgroups. Let U_i be a $\mathbb{C}G_i$-module for each $i \in \{1, 2\}$. Then,*

$$(U_1 \otimes U_2)^{K_1 \times K_2} = U_1^{K_1} \otimes U_2^{K_2} \tag{4.12}$$

(as subspaces of $U_1 \otimes U_2$).

Proof. The subgroup $K_1 = K_1 \times 1$ of $G_1 \times G_2$ acts on $U_1 \otimes U_2$, and its fixed points are $(U_1 \otimes U_2)^{K_1} = U_1^{K_1} \otimes U_2$ (because for a $\mathbb{C}K_1$-module, tensoring with U_2 is the same as taking a direct power, which clearly commutes with taking fixed points). Similarly, $(U_1 \otimes U_2)^{K_2} = U_1 \otimes U_2^{K_2}$. Now,

$$(U_1 \otimes U_2)^{K_1 \times K_2} = (U_1 \otimes U_2)^{K_1} \cap (U_1 \otimes U_2)^{K_2} = \left(U_1^{K_1} \otimes U_2\right) \cap \left(U_1 \otimes U_2^{K_2}\right) = U_1^{K_1} \otimes U_2^{K_2}$$

according to the known linear-algebraic fact stating that if P and Q are subspaces of two vector spaces U and V, respectively, then $(P \otimes V) \cap (U \otimes Q) = P \otimes Q$. $\square$

4.1.7. *Semidirect products.* Recall that a *semidirect product* is a group $G \ltimes K$ having two subgroups G, K with

- $K \triangleleft (G \ltimes K)$ is a normal subgroup,
- $G \ltimes K = GK = KG$, and
- $G \cap K = \{e\}$.

In this setting one has two interesting adjoint constructions, applied in Section 4.5.

Proposition 4.4. *Fix a $\mathbb{C}[G \ltimes K]$-module V.*

(i) *For any $\mathbb{C}G$-module U, one has $\mathbb{C}[G \ltimes K]$-module structure*

$$\Phi(U) := U \otimes V,$$

determined via

$$k(u \otimes v) = u \otimes k(v),$$
$$g(u \otimes v) = g(u) \otimes g(v).$$

(ii) *For any $\mathbb{C}[G \ltimes K]$-module W, one has $\mathbb{C}G$-module structure*

$$\Psi(W) := \mathrm{Hom}_{\mathbb{C}K}(\mathrm{Res}_K^{G \ltimes K} V, \mathrm{Res}_K^{G \ltimes K} W),$$

determined via $g(\varphi) = g \circ \varphi \circ g^{-1}$.

(iii) *The maps*

$$\mathbb{C}G-\text{mods} \underset{\Psi}{\overset{\Phi}{\rightleftharpoons}} \mathbb{C}[G \ltimes K]-\text{mods}$$

are adjoint in the sense that one has an isomorphism

$$\begin{array}{ccc} \operatorname{Hom}_{\mathbb{C}G}(U, \Psi(W)) & \longrightarrow & \operatorname{Hom}_{\mathbb{C}[G\ltimes K]}(\Phi(U), W) \\ \| & & \| \\ \operatorname{Hom}_{\mathbb{C}G}(U, \operatorname{Hom}_{\mathbb{C}K}(\operatorname{Res}^{G\ltimes K}_K V, \operatorname{Res}^{G\ltimes K}_K W)) & & \operatorname{Hom}_{\mathbb{C}[G\ltimes K]}(U\otimes V, W) \\ & & \\ \varphi & \longmapsto & \overline{\varphi}(u\otimes v) := \varphi(u)(v) \end{array}$$

(iv) *One has a $\mathbb{C}G$-module isomorphism*

$$(\Psi\circ\Phi)(U) \cong U \otimes \operatorname{End}_{\mathbb{C}K}(\operatorname{Res}^{G\ltimes K}_K V).$$

In particular, if $\operatorname{Res}^{G\ltimes K}_K V$ is a simple $\mathbb{C}K$-module, then $(\Psi\circ\Phi)(U)\cong U$.

Proof. These are mostly straightforward exercises in the definitions. To check assertion (iv), for example, note that K acts only in the right tensor factor in $\operatorname{Res}^{G\ltimes K}_K(U\otimes V)$, and hence as $\mathbb{C}G$-modules one has

$$\begin{aligned} (\Psi\circ\Phi)(U) &= \operatorname{Hom}_{\mathbb{C}K}(\operatorname{Res}^{G\ltimes K}_K V,\ \operatorname{Res}^{G\ltimes K}_K(U\otimes V)) \\ &= \operatorname{Hom}_{\mathbb{C}K}(\operatorname{Res}^{G\ltimes K}_K V,\ U\otimes\operatorname{Res}^{G\ltimes K}_K V) \\ &= U\otimes\operatorname{Hom}_{\mathbb{C}K}(\operatorname{Res}^{G\ltimes K}_K V,\ \operatorname{Res}^{G\ltimes K}_K V) \\ &= U\otimes\operatorname{End}_{\mathbb{C}K}(\operatorname{Res}^{G\ltimes K}_K V) \end{aligned}$$

□

4.2. **Three towers of groups.** Here we consider three towers of groups

$$G_* = (G_1 < G_2 < G_3 < \cdots)$$

where either

- $G_n = \mathfrak{S}_n$, the *symmetric group*, or
- $G_n = \mathfrak{S}_n[\Gamma]$, the *wreath product* of the symmetric group with some arbitrary finite group Γ, or
- $G_n = GL_n(\mathbb{F}_q)$, the finite general linear group.

Here the wreath product $\mathfrak{S}_n[\Gamma]$ can be thought of informally as the group of *monomial* $n\times n$ matrices whose nonzero entries lie in Γ, that is, $n\times n$ matrices having exactly one nonzero entry in each row and column, and that entry is an element of Γ. E.g.

$$\begin{bmatrix} 0 & g_2 & 0 \\ g_1 & 0 & 0 \\ 0 & 0 & g_3 \end{bmatrix} \begin{bmatrix} 0 & 0 & g_6 \\ 0 & g_5 & 0 \\ g_4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & g_2g_5 & 0 \\ 0 & 0 & g_1g_6 \\ g_3g_4 & 0 & 0 \end{bmatrix}.$$

More formally, $\mathfrak{S}_n[\Gamma]$ is the semidirect product $\mathfrak{S}_n \ltimes \Gamma^n$ in which $\mathfrak{S}_n$ acts on Γ^n via $\sigma(\gamma_1,\ldots,\gamma_n) = (\gamma_{\sigma^{-1}(1)},\ldots,\gamma_{\sigma^{-1}(n)})$.

For each of the three towers G_*, there are embeddings $G_i\times G_j \hookrightarrow G_{i+j}$ and we introduce maps $\operatorname{ind}^{i+j}_{i,j}$ taking $\mathbb{C}[G_i\times G_j]$-modules to $\mathbb{C}G_{i+j}$-modules, as well as maps $\operatorname{res}^{i+j}_{i,j}$ carrying modules in the reverse direction which are adjoint:

$$\operatorname{Hom}_{\mathbb{C}G_{i+j}}(\operatorname{ind}^{i+j}_{i,j} U, V) = \operatorname{Hom}_{\mathbb{C}[G_i\times G_j]}(U, \operatorname{res}^{i+j}_{i,j} V). \tag{4.13}$$

Definition 4.5. For $G_n = \mathfrak{S}_n$, one embeds $\mathfrak{S}_i\times\mathfrak{S}_j$ into $\mathfrak{S}_{i+j}$ as the permutations that permute $\{1,2,\ldots,i\}$ and $\{i+1,i+2,\ldots,i+j\}$ separately. Here one defines

$$\begin{aligned} \operatorname{ind}^{i+j}_{i,j} &:= \operatorname{Ind}^{\mathfrak{S}_{i+j}}_{\mathfrak{S}_i\times\mathfrak{S}_j}, \\ \operatorname{res}^{i+j}_{i,j} &:= \operatorname{Res}^{\mathfrak{S}_{i+j}}_{\mathfrak{S}_i\times\mathfrak{S}_j}. \end{aligned}$$

For $G_n = \mathfrak{S}_n[\Gamma]$, similarly embed $\mathfrak{S}_i[\Gamma] \times \mathfrak{S}_j[\Gamma]$ into $\mathfrak{S}_{i+j}[\Gamma]$ as block monomial matrices whose two diagonal blocks have sizes i, j respectively, and define

$$\begin{aligned} \mathrm{ind}_{i,j}^{i+j} &:= \mathrm{Ind}_{\mathfrak{S}_i[\Gamma]\times\mathfrak{S}_j[\Gamma]}^{\mathfrak{S}_{i+j}[\Gamma]}, \\ \mathrm{res}_{i,j}^{i+j} &:= \mathrm{Res}_{\mathfrak{S}_i[\Gamma]\times\mathfrak{S}_j[\Gamma]}^{\mathfrak{S}_{i+j}[\Gamma]}. \end{aligned}$$

For $G_n = GL_n(\mathbb{F}_q)$, which we will denote just GL_n, similarly embed $GL_i \times GL_j$ into GL_{i+j} as block diagonal matrices whose two diagonal blocks have sizes i, j respectively. However, one also introduces as an intermediate the *parabolic subgroup* $P_{i,j}$ consisting of the block upper-triangular matrices of the form

$$\begin{bmatrix} g_i & \ell \\ 0 & g_j \end{bmatrix}$$

where g_i, g_j lie in GL_i, GL_j, respectively, and ℓ in $\mathbb{F}_q^{i\times j}$ is arbitrary. One has a quotient map $P_{i,j} \to GL_i \times GL_j$ whose kernel $K_{i,j}$ is the set of matrices of the form

$$\begin{bmatrix} I_i & \ell \\ 0 & I_j \end{bmatrix}$$

with ℓ again arbitrary. Here one defines

$$\begin{aligned} \mathrm{ind}_{i,j}^{i+j} &:= \mathrm{Ind}_{P_{i,j}}^{GL_{i+j}} \mathrm{Infl}_{GL_i\times GL_j}^{P_{i,j}}, \\ \mathrm{res}_{i,j}^{i+j} &:= \left(\mathrm{Res}_{P_{i,j}}^{GL_{i+j}}(-)\right)^{K_{i,j}}. \end{aligned}$$

The operation $\mathrm{ind}_{i,j}^{i+j}$ is sometimes called *parabolic induction* or *Harish-Chandra induction*. The operation $\mathrm{res}_{i,j}^{i+j}$ is essentially just the $K_{i,j}$-fixed point construction $V \mapsto V^{K_{i,j}}$. However writing it as the above two-step composite makes it more obvious, (via (4.5) and (4.8)) that $\mathrm{res}_{i,j}^{i+j}$ is again adjoint to $\mathrm{ind}_{i,j}^{i+j}$.

Definition 4.6. For each of the three towers G_*, define a graded $\mathbb{Z}$-module

$$A := A(G_*) = \bigoplus_{n\geq 0} R(G_n)$$

with a bilinear form $(\cdot,\cdot)_A$ whose restriction to $A_n := R(G_n)$ is the usual form $(\cdot,\cdot)_{G_n}$, and such that $\Sigma = \bigsqcup_{n\geq 0} \mathrm{Irr}(G_n)$ gives an orthonormal $\mathbb{Z}$-basis. Here we adopt the convention that $A_0 = \mathbb{Z}$ has its basis element 1 equal to the unique irreducible character of the trivial group G_0.

Bearing in mind that $A_n = R(G_n)$ and

$$A_i \otimes A_j = R(G_i) \otimes R(G_j) \cong R(G_i \times G_j)$$

one then has candidates for product and coproduct defined by

$$\begin{array}{rcrcl} m := \mathrm{ind}_{i,j}^{i+j} & : & A_i \otimes A_j & \longrightarrow & A_{i+j} \\ \Delta := \bigoplus_{i+j=n} \mathrm{res}_{i,j}^{i+j} & : & A_n & \longrightarrow & \bigoplus_{i+j=n} A_i \otimes A_j. \end{array}$$

The coassociativity of Δ is an easy consequence of transitivity of the constructions of restriction and fixed points. We could derive the associativity of m from the transitivity of induction and inflation, but this would be more complicated; we will instead prove it differently.

We first show that the maps m and Δ are adjoint with respect to the forms $(\cdot,\cdot)_A$ and $(\cdot,\cdot)_{A\otimes A}$. In fact, if U, V, W are modules over $\mathbb{C}G_i$, $\mathbb{C}G_j$, $\mathbb{C}G_{i+j}$, respectively, then we can write the $\mathbb{C}[G_i \times G_j]$-module $\mathrm{res}_{i,j}^{i+j} W$ as a direct sum $\bigoplus_k X_k \otimes Y_k$ with X_k being $\mathbb{C}G_i$-modules and Y_k being $\mathbb{C}G_j$-modules; we then have

$$\mathrm{res}_{i,j}^{i+j} \chi_W = \sum_k \chi_{X_k} \otimes \chi_{Y_k} \tag{4.14}$$

and

$$\begin{aligned}(m(\chi_U \otimes \chi_V), \chi_W)_A &= \left(\mathrm{ind}_{i,j}^{i+j}(\chi_{U\otimes V}), \chi_W\right)_A = \left(\mathrm{ind}_{i,j}^{i+j}(\chi_{U\otimes V}), \chi_W\right)_{G_{i+j}} \\ &= \left(\chi_{U\otimes V}, \mathrm{res}_{i,j}^{i+j}\chi_W\right)_{G_i\times G_j} = \left(\chi_{U\otimes V}, \sum_k \chi_{X_k}\otimes\chi_{Y_k}\right)_{G_i\times G_j} \\ &= \sum_k (\chi_{U\otimes V}, \chi_{X_k\otimes Y_k})_{G_i\times G_j} = \sum_k (\chi_U, \chi_{X_k})_{G_i}(\chi_V, \chi_{Y_k})_{G_j}\end{aligned}$$

(the third equality sign follows by taking dimensions in (4.13) and recalling (4.1); the fourth equality sign follows from (4.14); the sixth one follows from (4.2)) and

$$\begin{aligned}(\chi_U\otimes\chi_V, \Delta(\chi_W))_{A\otimes A} &= \left(\chi_U\otimes\chi_V, \mathrm{res}_{i,j}^{i+j}\chi_W\right)_{A\otimes A} = \left(\chi_U\otimes\chi_V, \sum_k\chi_{X_k}\otimes\chi_{Y_k}\right)_{A\otimes A} \\ &= \sum_k(\chi_U,\chi_{X_k})_A(\chi_V,\chi_{Y_k})_A = \sum_k(\chi_U,\chi_{X_k})_{G_i}(\chi_V,\chi_{Y_k})_{G_j}.\end{aligned}$$

(the first equality sign follows by removing all terms in $\Delta(\chi_W)$ whose scalar product with $\chi_U\otimes\chi_V$ vanishes for reasons of gradedness; the second equality sign follows from (4.14)), which in comparison yield $(m(\chi_U\otimes\chi_V),\chi_W)_A = (\chi_U\otimes\chi_V,\Delta(\chi_W))_{A\otimes A}$, thus showing that m and Δ are adjoint maps. Therefore, m is associative (since Δ is coassociative).

Endowing $A = \bigoplus_{n\geq 0} R(G_n)$ with the obvious unit and counit maps, it thus becomes a graded, finite-type $\mathbb{Z}$-algebra and $\mathbb{Z}$-coalgebra.

The next section address the issue of why they form a bialgebra. However, assuming this for the moment, it should be clear that each of these algebras A is a PSH having $\Sigma = \bigsqcup_{n\geq 0}\mathrm{Irr}(G_n)$ as its PSH-basis. Σ is self-dual because m,Δ are defined by adjoint maps, and it is positive because m,Δ take irreducible representations to genuine representations not just virtual ones, and hence have characters which are nonnegative sums of irreducible characters.

4.3. **Bialgebra and double cosets.** To show that the algebra and coalgebras $A = A(G_*)$ are bialgebras, the central issue is checking the pentagonal diagram in (1.8), that is, as maps $A\otimes A\to A\otimes A$, one has

$$\Delta\circ m = (m\otimes m)\circ(1\otimes T\otimes 1)\circ(\Delta\otimes\Delta). \tag{4.15}$$

In checking this, it is convenient to have a lighter notation for various subgroups of the groups G_n corresponding to compositions α.

Definition 4.7. A *composition* is a (finite) tuple $\alpha = (\alpha_1,\alpha_2,\ldots,\alpha_\ell)$ of positive integers. Its *length* is defined to be ℓ and denoted by $\ell(\alpha)$; its *size* is defined to be $\alpha_1+\alpha_2+\cdots+\alpha_\ell$ and denoted by $|\alpha|$; its *parts* are its entries $\alpha_1,\alpha_2,\ldots,\alpha_\ell$. The compositions of size n are called the *compositions of* n. In particular, any partition of n (written without trailing zeroes) is a composition of n. We write $\emptyset$ (and sometimes, sloppily, (0)) for the empty composition $()$.

Definition 4.8. Given a composition $\alpha = (\alpha_1,\ldots,\alpha_\ell)$ of n, define a subgroup

$$G_\alpha \cong G_{\alpha_1}\times\cdots\times G_{\alpha_\ell} < G_n$$

via the block-diagonal embedding with diagonal blocks of sizes $(\alpha_1,\ldots,\alpha_\ell)$. This G_α is called a *Young subgroup* $\mathfrak{S}_\alpha$ when $G_n = \mathfrak{S}_n$, and a *Levi subgroup* when $G_n = GL_n$. In the case when $G_n = \mathfrak{S}_n[\Gamma]$, we also denote G_α by $\mathfrak{S}_\alpha[\Gamma]$. In the case where $G_n = GL_n$, also define the *parabolic subgroup* P_α to be the subgroup of G_n consisting of block-upper triangular matrices whose diagonal blocks have sizes $(\alpha_1,\ldots,\alpha_\ell)$, and let K_α be the kernel of the obvious surjection $P_\alpha\to G_\alpha$ which sends a block upper-triangular matrix to the tuple of its diagonal blocks whose sizes are $\alpha_1,\alpha_2,\ldots,\alpha_\ell$. Notice that $P_{(i,j)} = P_{i,j}$ for any i and j with $i+j=n$; we will also abbreviate $G_{(i,j)} = G_i\times G_j$ by $G_{i,j}$. Also define operators $\mathrm{ind}_\alpha^n, \mathrm{res}_\alpha^n$ for the three towers analogous to those defined in Definition 4.5 when $\alpha = (i,j)$ has only two parts.

Definition 4.9. Let K and H be two groups, $\tau: K\to H$ a group homomorphism, and U a $\mathbb{C}H$-module. Then, U^τ is defined as the $\mathbb{C}K$-module with ground space U and action given by $k\cdot u = \tau(k)\cdot u$ for all $k\in K$ and $u\in U$. This very simple construction generalizes the definition of U^g for an element $g\in G$, where G

is a group containing H as a subgroup; in fact, in this situation we have $U^g = U^\tau$, where $K = {}^gH$ and $\tau : K \to H$ is the map $k \mapsto g^{-1}kg$.

Using homogeneity, checking the bialgebra condition (4.15) in the homogeneous component $(A \otimes A)_n$ amounts to the following: for each pair of representations U_1, U_2 of G_{r_1}, G_{r_2} with $r_1 + r_2 = n$, and for each (c_1, c_2) with $c_1 + c_2 = n$, one must verify that

$$\begin{aligned} &\operatorname{res}^n_{c_1,c_2}\left(\operatorname{ind}^n_{r_1,r_2}(U_1 \otimes U_2)\right) \\ &\cong \bigoplus_A \left(\operatorname{ind}^{c_1}_{a_{11},a_{21}} \otimes \operatorname{ind}^{c_2}_{a_{12},a_{22}}\right)\left(\left(\operatorname{res}^{r_1}_{a_{11},a_{12}} U_1 \otimes \operatorname{res}^{r_2}_{a_{21},a_{22}} U_2\right)^{\tau_A^{-1}}\right) \end{aligned} \tag{4.16}$$

where the direct sum is over all matrices $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ in $\mathbb{N}^{2\times 2}$ with row sums (r_1, r_2) and column sums (c_1, c_2), and where τ_A is the obvious isomorphism between the subgroups

$$\begin{aligned} &G_{a_{11},a_{12},a_{21},a_{22}} \quad (< G_{r_1,r_2}) \\ &G_{a_{11},a_{21},a_{12},a_{22}} \quad (< G_{c_1,c_2}) \end{aligned} \tag{4.17}$$

(we are using the inverse τ_A^{-1} of this isomorphism τ_A to identify modules for the first subgroup with modules for the second subgroup, according to Definition 4.9).

As one might guess, (4.16) comes from the Mackey formula (Theorem 4.2), once one identifies the appropriate double coset representatives. This is just as easy to do in a slightly more general setting.

Definition 4.10. Given compositions α, β of n having lengths ℓ, m and a matrix A in $\mathbb{N}^{\ell\times m}$ with row sums α and column sums β, define a permutation w_A in $\mathfrak{S}_n$ as follows. Disjointly decompose $[n] = \{1, 2, \ldots, n\}$ into consecutive intervals of numbers

$$\begin{aligned} [n] &= I_1 \sqcup \cdots \sqcup I_\ell \\ [n] &= J_1 \sqcup \cdots \sqcup J_m \end{aligned}$$

such that $|I_i| = \alpha_i, |J_j| = \beta_j$. For every $j \in [m]$, disjointly decompose J_j into consecutive intervals of numbers $J_j = J_{j,1} \sqcup J_{j,2} \sqcup \cdots \sqcup J_{j,\ell}$ such that every $i \in [\ell]$ satisfies $|J_{j,i}| = a_{ij}$. For every $i \in [\ell]$, disjointly decompose I_i into consecutive intervals of numbers $I_i = I_{i,1} \sqcup I_{i,2} \sqcup \cdots \sqcup I_{i,m}$ such that every $j \in [m]$ satisfies $|I_{i,j}| = a_{ij}$. Now, for every $i \in [\ell]$ and $j \in [m]$, let $\pi_{i,j}$ be the increasing bijection from $J_{j,i}$ to $I_{i,j}$ (this is well-defined since these two sets both have cardinality a_{ij}). The disjoint union of these bijections $\pi_{i,j}$ over all i and j is a bijection $[n] \to [n]$ (since the disjoint union of the sets $J_{j,i}$ over all i and j is $[n]$, and so is the disjoint union of the sets $I_{i,j}$), that is, a permutation of $[n]$; this permutation is what we call w_A.

Example 4.11. Taking $n = 9$ and $\alpha = (4, 5), \beta = (3, 4, 2)$, one has

$$\begin{aligned} &I_1 = \{1, 2, 3, 4\}, \quad I_2 = \{5, 6, 7, 8, 9\} \\ &J_1 = \{1, 2, 3\}, \quad J_2 = \{4, 5, 6, 7\}, \quad J_3 = \{8, 9\}. \end{aligned}$$

Then one possible matrix A having row and column sums α, β is $A = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 2 & 2 \end{bmatrix}$, and its associated permutation w_A written in two-line notation is

$$\begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ \underline{1} & \underline{2} & \underline{\underline{5}} & | & \underline{3} & \underline{4} & \underline{\underline{6}} & \underline{\underline{7}} & | & \underline{\underline{8}} & \underline{\underline{9}} \end{pmatrix}$$

with vertical lines dividing the sets J_j on top, and with elements of I_i underlined i times on the bottom.

Remark 4.12. Given compositions α and β of n having lengths ℓ and m, and a permutation $w \in \mathfrak{S}_n$. It is easy to see that there exists a matrix $A \in \mathbb{N}^{\ell\times m}$ satisfying $w_A = w$ if and only if the restriction of w to each J_j and the restriction of w^{-1} to each I_i are increasing. In this case, the matrix A is determined by $a_{ij} = |w(J_j) \cap I_i|$.

Among our three towers G_* of groups, the symmetric group tower ($G_n = \mathfrak{S}_n$) is the simplest one. We will now see that it also embeds into the two others, in the sense that $\mathfrak{S}_n$ embeds into $\mathfrak{S}_n[\Gamma]$ for every Γ and into $GL_n(\mathbb{F}_q)$ for every q.

First, for every $n \in \mathbb{N}$ and any group Γ, we embed the group $\mathfrak{S}_n$ into $\mathfrak{S}_n[\Gamma]$ by means of the canonical embedding $\mathfrak{S}_n \to \mathfrak{S}_n \ltimes \Gamma^n = \mathfrak{S}_n[\Gamma]$. If we regard elements of $\mathfrak{S}_n[\Gamma]$ as $n\times n$ monomial matrices with nonzero entries in Γ, then this boils down to identifying every $\pi \in \mathfrak{S}_n$ with the permutation matrix of π (in which the 1's are read as the neutral element of Γ). If α is a composition of n, then this embedding $\mathfrak{S}_n \to \mathfrak{S}_n [\Gamma]$ makes the subgroup $\mathfrak{S}_\alpha$ of $\mathfrak{S}_n$ become a subgroup of $\mathfrak{S}_n [\Gamma]$ and actually into a subgroup of $\mathfrak{S}_\alpha [\Gamma] < \mathfrak{S}_n [\Gamma]$.

For every $n \in \mathbb{N}$ and every q, we embed the group $\mathfrak{S}_n$ into $GL_n (\mathbb{F}_q)$ by identifying every permutation $\pi \in \mathfrak{S}_n$ with its permutation matrix in $GL_n (\mathbb{F}_q)$. If α is a composition of n, then this embedding makes the subgroup $\mathfrak{S}_\alpha$ of $\mathfrak{S}_n$ become a subgroup of $GL_n (\mathbb{F}_q)$. If we let $G_n = GL_n (\mathbb{F}_q)$, then $\mathfrak{S}_\alpha < G_\alpha < P_\alpha$.

The embeddings we have just defined commute with the group embeddings $G_n < G_{n+1}$ on both sides.

Proposition 4.13. *The permutations $\{w_A\}$ as A runs over all matrices in $\mathbb{N}^{\ell \times m}$ having row, column sums α, β give a system of double coset representatives for*

$$\mathfrak{S}_\alpha \backslash \mathfrak{S}_n / \mathfrak{S}_\beta$$
$$\mathfrak{S}_\alpha[\Gamma] \backslash \mathfrak{S}_n[\Gamma] / \mathfrak{S}_\beta[\Gamma]$$
$$P_\alpha \backslash GL_n / P_\beta$$

Proof. First note that double coset representatives for $\mathfrak{S}_\alpha \backslash \mathfrak{S}_n / \mathfrak{S}_\beta$ should also provide double coset representatives for $\mathfrak{S}_\alpha[\Gamma] \backslash \mathfrak{S}_n[\Gamma] / \mathfrak{S}_\beta[\Gamma]$, since

$$\mathfrak{S}_\alpha[\Gamma] = \mathfrak{S}_\alpha \Gamma^n = \Gamma^n \mathfrak{S}_\alpha.$$

We give an algorithm to show that every double coset $\mathfrak{S}_\alpha w \mathfrak{S}_\beta$ contains some w_A. Start by altering w within its coset $w\mathfrak{S}_\beta$, that is, by permuting the *positions* within each set J_j, to obtain a representative w' for $w\mathfrak{S}_\beta$ in which each set $w'(J_j)$ appears in increasing order in the second line of the two-line notation for w'. Then alter w' within its coset $\mathfrak{S}_\alpha w'$, that is, by permuting the *values* within each set I_i, to obtain a representative w_A having the elements of each set I_i appearing in increasing order in the second line; because the values within each set I_i are consecutive, this alteration will not ruin the property that one had each set $w'(J_j)$ appearing in increasing order. For example, one might have

$$w = \begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ \underline{4} & \underline{\underline{8}} & \underline{2} & | & \underline{\underline{5}} & \underline{\underline{3}} & \underline{\underline{\underline{9}}} & \underline{1} & | & \underline{\underline{\underline{7}}} & \underline{\underline{\underline{6}}} \end{pmatrix}$$
$$w' = \begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ \underline{2} & \underline{4} & \underline{\underline{8}} & | & \underline{1} & \underline{\underline{3}} & \underline{\underline{5}} & \underline{\underline{\underline{9}}} & | & \underline{\underline{\underline{6}}} & \underline{\underline{\underline{7}}} \end{pmatrix} \in w\mathfrak{S}_\beta$$
$$w_A = \begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ \underline{1} & \underline{2} & \underline{\underline{5}} & | & \underline{3} & \underline{4} & \underline{\underline{6}} & \underline{\underline{7}} & | & \underline{\underline{\underline{8}}} & \underline{\underline{\underline{9}}} \end{pmatrix} \in \mathfrak{S}_\alpha w' \subset \mathfrak{S}_\alpha w' \mathfrak{S}_\beta = \mathfrak{S}_\alpha w \mathfrak{S}_\beta$$

Next note that $\mathfrak{S}_\alpha w_A \mathfrak{S}_\beta = \mathfrak{S}_\alpha w_B \mathfrak{S}_\beta$ implies $A = B$, since the quantities

$$a_{i,j}(w) := |w(J_j) \cap I_i|$$

are easily seen to be constant on double cosets $\mathfrak{S}_\alpha w \mathfrak{S}_\beta$.

A similar argument shows that $P_\alpha w_A P_\beta = P_\alpha w_B P_\beta$ implies $A = B$: for g in GL_n, the rank $r_{ij}(g)$ of the matrix obtained by restricting g to rows $I_i \sqcup I_{i+1} \sqcup \cdots \sqcup I_\ell$ and columns $J_1 \sqcup J_2 \sqcup \cdots \sqcup J_j$ is constant on double cosets $P_\alpha g P_\beta$, and for a permutation matrix w one can recover $a_{i,j}(w)$ from the formula

$$a_{i,j}(w) = r_{i,j}(w) - r_{i,j-1}(w) - r_{i+1,j}(w) + r_{i+1,j-1}(w).$$

Thus it only remains to show that every double coset $P_\alpha g P_\beta$ contains some w_A. Since $\mathfrak{S}_\alpha < P_\alpha$, and we have seen already that every double coset $\mathfrak{S}_\alpha w \mathfrak{S}_\beta$ contains some w_A, it suffices to show that every double coset $P_\alpha g P_\beta$ contains some permutation w. However, we claim that this is already true for the smaller double cosets BgB where $B = P_{1^n}$ is the *Borel subgroup* of upper triangular invertible matrices, that is, one has the usual *Bruhat decomposition*

$$GL_n = \bigsqcup_{w \in \mathfrak{S}_n} BwB.$$

To prove this decomposition, we show how to find a permutation w in each double coset BgB. The freedom to alter g within its coset gB allows one to scale columns and add scalar multiples of earlier columns to later columns. We claim that using such column operations, one can always find a representative g' for coset gB in which

- the bottommost nonzero entry of each column has been scaled to 1 (call this a *pivot*),
- the entries to right of each pivot within its row are all 0, and

- there is one pivot in each row and each column, so that they lie in the positions of some permutation matrix w.

In fact, we will see below that $BgB = BwB$ in this case. The algorithm which produces g' from g is simple: starting with the leftmost column, find its bottommost nonzero entry, and scale the column to make this entry a 1, creating the pivot in this column. Now use this pivot to clear out all entries in its row to its right, using column operations that subtract multiples of this column from later columns. Having done this, move on to the next column to the right, and repeat, scaling to create a pivot, and using it to eliminate entries to its right.

For example, the typical matrix g lying in the double coset BwB where

$$w = \begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ \underline{4} & \underline{8} & \underline{2} & | & \underline{5} & \underline{3} & \underline{9} & \underline{1} & | & \underline{7} & \underline{6} \end{pmatrix}$$

from before is one that can be altered within its coset gB to look like this:

$$g' = \begin{bmatrix} * & * & * & * & * & * & 1 & 0 & 0 \\ * & * & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & 0 & * & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & * & 0 & * & 1 \\ 0 & * & 0 & 0 & 0 & * & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \in gB.$$

Having found this g' in gB, a similar algorithm using left multiplication by B shows that w lies in $Bg' \subset Bg'B = BgB$. This time no scalings are required to create the pivot entries: starting with the bottom row, one uses its pivot to eliminate all the entries above it in the same column (shown by stars $*$ above) by adding multiples of the bottom row to higher rows. Then do the same using the pivot in the next-to-bottom row, etc. The result is the permutation matrix for w. ☐

Corollary 4.14. *For each of the three towers of groups G_*, the product and coproduct structures on $A = A(G_*)$ endow it with a bialgebra structure, and hence they form PSH-algebras.*

Proof. The first two towers $G_n = \mathfrak{S}_n$ and $G_n = \mathfrak{S}_n[\Gamma]$ have product, coproduct defined by induction, restriction along embeddings $G_i \times G_j < G_{i+j}$. Hence the desired bialgebra equality (4.16) follows from Mackey's Theorem 4.2, taking $G = G_n, H = G_{(r_1,r_2)}, K = G_{(c_1,c_2)}, U = U_1 \otimes U_2$ with double coset representatives[15]

$$\{g_1, \ldots, g_t\} = \left\{ w_{A^t} \mid A \in \mathbb{N}^{2\times 2},\ A \text{ has row sums } (r_1, r_2) \text{ and column sums } (c_1, c_2) \right\}$$

and checking for a given double coset

$$KgH = (G_{c_1,c_2}) w_{A^t} (G_{r_1,r_2})$$

indexed by a matrix A in $\mathbb{N}^{2\times 2}$ with row sums (r_1, r_2) and column sums (c_1, c_2), that the two subgroups appearing on the left in (4.17) are exactly

$$H \cap K^{w_{A^t}} = G_{r_1,r_2} \cap (G_{c_1,c_2})^{w_{A^t}},$$
$${}^{w_{A^t}} H \cap K = {}^{w_{A^t}}(G_{r_1,r_2}) \cap G_{c_1,c_2},$$

respectively. One should also apply (4.4) and check that the isomorphism τ_A between the two subgroups in (4.17) is the conjugation isomorphism by w_{A^t} (that is, $\tau_A(g) = w_{A^t} g w_{A^t}^{-1}$ for every $g \in H \cap K^{w_{A^t}}$). We leave all of these bookkeeping details to the reader to check. [16]

[15]Proposition 4.13 gives as a system of double coset representatives for $G_{(c_1,c_2)} \backslash G_n / G_{(r_1,r_2)}$ the elements

$$\left\{ w_A \mid A \in \mathbb{N}^{2\times 2},\ A \text{ has row sums } (c_1, c_2) \text{ and column sums } (r_1, r_2) \right\}$$
$$= \left\{ w_{A^t} \mid A \in \mathbb{N}^{2\times 2},\ A \text{ has row sums } (r_1, r_2) \text{ and column sums } (c_1, c_2) \right\}$$

where A^t denotes the transpose matrix of A.

[16]It helps to recognize w_{A^t} as the permutation written in two-line notation as

$$\begin{pmatrix} 1 & 2 & \ldots & a_{11} & | & a_{11}+1 & a_{11}+2 & \ldots & r_1 & | & r_1+1 & r_1+2 & \ldots & a'_{22} & | & a'_{22}+1 & a'_{22}+2 & \ldots & n \\ 1 & 2 & \ldots & a_{11} & | & c_1+1 & c_1+2 & \ldots & a'_{22} & | & a_{11}+1 & a_{11}+2 & \ldots & c_1 & | & a'_{22}+1 & a'_{22}+2 & \ldots & n \end{pmatrix},$$

For the tower with $G_n = GL_n$, there is slightly more work to be done to check the equality (4.16). Via Mackey's Theorem 4.2 and Proposition 4.13, the left side is

$$\begin{aligned}
&\operatorname{res}^n_{c_1,c_2}\left(\operatorname{ind}^n_{r_1,r_2}(U_1 \otimes U_2)\right)\\
&= \left(\operatorname{Res}^{G_n}_{P_{c_1,c_2}} \operatorname{Ind}^{G_n}_{P_{r_1,r_2}} \operatorname{Infl}^{P_{r_1,r_2}}_{G_{r_1,r_2}}(U_1 \otimes U_2)\right)^{K_{c_1,c_2}}\\
&= \bigoplus_A \left(\operatorname{Ind}^{P_{c_1,c_2}}_{{}^{w_{A^t}}P_{r_1,r_2}\cap P_{c_1,c_2}}\left(\left(\operatorname{Res}^{P_{r_1,r_2}}_{P_{r_1,r_2}\cap P^{w_{A^t}}_{c_1,c_2}} \operatorname{Infl}^{P_{r_1,r_2}}_{G_{r_1,r_2}}(U_1 \otimes U_2)\right)^{\tau_A^{-1}}\right)\right)^{K_{c_1,c_2}}
\end{aligned} \tag{4.18}$$

where A runs over the usual 2×2 matrices. The right side is a direct sum over this same set of matrices A:

$$\begin{aligned}
&\bigoplus_A \left(\operatorname{ind}^{c_1}_{a_{11},a_{21}} \otimes \operatorname{ind}^{c_2}_{a_{12},a_{22}}\right)\left(\left(\operatorname{res}^{r_1}_{a_{11},a_{12}} U_1 \otimes \operatorname{res}^{r_2}_{a_{21},a_{22}} U_2\right)^{\tau_A^{-1}}\right)\\
&= \bigoplus_A \left(\operatorname{Ind}^{G_{c_1}}_{P_{a_{11},a_{21}}} \otimes \operatorname{Ind}^{G_{c_2}}_{P_{a_{12},a_{22}}}\right) \circ \left(\operatorname{Infl}^{P_{a_{11},a_{21}}}_{G_{a_{11},a_{21}}} \otimes \operatorname{Infl}^{P_{a_{12},a_{22}}}_{G_{a_{12},a_{22}}}\right)\\
&\qquad \left(\left(\left(\operatorname{Res}^{G_{r_1}}_{P_{a_{11},a_{12}}} U_1\right)^{K_{a_{11},a_{12}}} \otimes \left(\operatorname{Res}^{G_{r_2}}_{P_{a_{21},a_{22}}} U_2\right)^{K_{a_{21},a_{22}}}\right)^{\tau_A^{-1}}\right)\\
&= \bigoplus_A \operatorname{Ind}^{G_{c_1,c_2}}_{P_{a_{11},a_{21}}\times P_{a_{12},a_{22}}}\\
&\qquad \operatorname{Infl}^{P_{a_{11},a_{21}}\times P_{a_{12},a_{22}}}_{G_{a_{11},a_{21},a_{12},a_{22}}}\left(\left(\left(\operatorname{Res}^{G_{r_1,r_2}}_{P_{a_{11},a_{12}}\times P_{a_{21},a_{22}}}(U_1 \otimes U_2)\right)^{K_{a_{11},a_{12}}\times K_{a_{21},a_{22}}}\right)^{\tau_A^{-1}}\right)
\end{aligned} \tag{4.19}$$

(by (4.4), (4.12) and their obvious analogues for restriction and inflation). Thus it suffices to check for each 2×2 matrix A that any $\mathbb{C}G_{c_1,c_2}$-module of the form $V_1 \otimes V_2$ has the same inner product with the A-summands of (4.18) and (4.19). Abbreviate $w := w_{A^t}$ and $\tau := \tau_A^{-1}$.

Notice that ${}^wP_{r_1,r_2}$ is the group of all matrices having the block form

$$\begin{bmatrix} g_{11} & h & i & j \\ 0 & g_{21} & 0 & k \\ d & e & g_{12} & \ell \\ 0 & f & 0 & g_{22} \end{bmatrix} \tag{4.20}$$

in which the diagonal blocks g_{ij} for $i, j = 1, 2$ are invertible of size $a_{ij} \times a_{ij}$, while the blocks $h, i, j, k, \ell, d, e, f$ are all arbitrary matrices[17] of the appropriate (rectangular) block sizes. Hence, ${}^wP_{r_1,r_2} \cap P_{c_1,c_2}$ is the group of all matrices having the block form

$$\begin{bmatrix} g_{11} & h & i & j \\ 0 & g_{21} & 0 & k \\ 0 & 0 & g_{12} & \ell \\ 0 & 0 & 0 & g_{22} \end{bmatrix} \tag{4.21}$$

in which the diagonal blocks g_{ij} for $i, j = 1, 2$ are invertible of size $a_{ij} \times a_{ij}$, while the blocks h, i, j, k, ℓ are all arbitrary matrices of the appropriate (rectangular) block sizes; then ${}^wP_{r_1,r_2} \cap G_{c_1,c_2}$ is the subgroup where the blocks i, j, k all vanish. The canonical projection ${}^wP_{r_1,r_2} \cap P_{c_1,c_2} \to {}^wP_{r_1,r_2} \cap G_{c_1,c_2}$ (obtained by restricting the projection $P_{c_1,c_2} \to G_{c_1,c_2}$) has kernel ${}^wP_{r_1,r_2} \cap P_{c_1,c_2} \cap K_{c_1,c_2}$. Consequently,

$$({}^wP_{r_1,r_2} \cap P_{c_1,c_2}) \,/\, ({}^wP_{r_1,r_2} \cap P_{c_1,c_2} \cap K_{c_1,c_2}) = {}^wP_{r_1,r_2} \cap G_{c_1,c_2}. \tag{4.22}$$

Similarly,

$$\left(P_{r_1,r_2} \cap P^w_{c_1,c_2}\right) / \left(P_{r_1,r_2} \cap P^w_{c_1,c_2} \cap K_{r_1,r_2}\right) = G_{r_1,r_2} \cap P^w_{c_1,c_2}. \tag{4.23}$$

---

where $a'_{22} = r_1 + a_{21} = c_1 + a_{12} = n - a_{22}$. In matrix form, w_{A^t} is the block matrix $\begin{bmatrix} I_{a_{11}} & 0 & 0 & 0 \\ 0 & 0 & I_{a_{21}} & 0 \\ 0 & I_{a_{12}} & 0 & 0 \\ 0 & 0 & 0 & I_{a_{22}} \end{bmatrix}$.

[17]The blocks i and j have nothing to do with the indices i, j in g_{ij}.

Computing first the inner product of $V_1 \otimes V_2$ with the A-summand of (4.18), and using adjointness properties, one gets

$$\begin{aligned}&\left(\left(\mathrm{Res}^{P_{r_1,r_2}}_{P_{r_1,r_2}\cap P^w_{c_1,c_2}} \mathrm{Infl}^{P_{r_1,r_2}}_{G_{r_1,r_2}} (U_1 \otimes U_2)\right)^{\tau},\right.\\ &\qquad\left.\mathrm{Res}^{P_{c_1,c_2}}_{{}^wP_{r_1,r_2}\cap P_{c_1,c_2}} \mathrm{Infl}^{P_{c_1,c_2}}_{G_{c_1,c_2}} (V_1 \otimes V_2)\right)_{{}^wP_{r_1,r_2}\cap P_{c_1,c_2}}\\ &\overset{(4.7)}{=} \left(\left(\mathrm{Infl}^{P_{r_1,r_2}\cap P^w_{c_1,c_2}}_{G_{r_1,r_2}\cap P^w_{c_1,c_2}} \mathrm{Res}^{G_{r_1,r_2}}_{G_{r_1,r_2}\cap P^w_{c_1,c_2}} (U_1 \otimes U_2)\right)^{\tau},\right.\\ &\qquad\left.\mathrm{Infl}^{{}^wP_{r_1,r_2}\cap P_{c_1,c_2}}_{{}^wP_{r_1,r_2}\cap G_{c_1,c_2}} \mathrm{Res}^{G_{c_1,c_2}}_{{}^wP_{r_1,r_2}\cap G_{c_1,c_2}} (V_1 \otimes V_2)\right)_{{}^wP_{r_1,r_2}\cap P_{c_1,c_2}}\end{aligned}$$

(by (4.23) and (4.22)). One can compute this inner product by first recalling that ${}^wP_{r_1,r_2} \cap P_{c_1,c_2}$ is the group of matrices having the block form (4.21) in which the diagonal blocks g_{ij} for $i,j = 1,2$ are invertible of size $a_{ij} \times a_{ij}$, while the blocks h, i, j, k, ℓ are all arbitrary matrices of the appropriate (rectangular) block sizes; then ${}^wP_{r_1,r_2} \cap G_{c_1,c_2}$ is the subgroup where the blocks i, j, k all vanish. The inner product above then becomes

$$\begin{aligned}&\frac{1}{|{}^wP_{r_1,r_2} \cap P_{c_1,c_2}|} \sum_{\substack{(g_{ij})\\(h,i,j,k,\ell)}} \chi_{U_1}\begin{pmatrix} g_{11} & i \\ 0 & g_{12}\end{pmatrix} \chi_{U_2}\begin{pmatrix} g_{21} & k \\ 0 & g_{22}\end{pmatrix}\\ &\qquad\qquad \overline{\chi}_{V_1}\begin{pmatrix} g_{11} & h \\ 0 & g_{21}\end{pmatrix} \overline{\chi}_{V_2}\begin{pmatrix} g_{12} & \ell \\ 0 & g_{22}\end{pmatrix}.\end{aligned} \tag{4.24}$$

If one instead computes the inner product of $V_1 \otimes V_2$ with the A-summand of (4.19), using adjointness properties and (4.11) one gets

$$\begin{aligned}&\left(\left(\left(\mathrm{Res}^{G_{r_1,r_2}}_{P_{a_{11},a_{12}}\times P_{a_{21},a_{22}}} (U_1 \otimes U_2)\right)^{K_{a_{11},a_{12}}\times K_{a_{21},a_{22}}}\right)^{\tau},\right.\\ &\qquad\left.\left(\mathrm{Res}^{G_{c_1,c_2}}_{P_{a_{11},a_{21}}\times P_{a_{12},a_{22}}} (V_1 \otimes V_2)\right)^{K_{a_{11},a_{21}}\times K_{a_{12},a_{22}}}\right)_{G_{a_{11},a_{21},a_{12},a_{22}}}\\ &= \frac{1}{|G_{a_{11},a_{21},a_{12},a_{22}}|}\sum_{(g_{ij})} \frac{1}{|K_{a_{11},a_{12}} \times K_{a_{21},a_{22}}|} \sum_{(i,k)} \chi_{U_1}\begin{pmatrix} g_{11} & i \\ 0 & g_{12}\end{pmatrix} \chi_{U_2}\begin{pmatrix} g_{21} & k \\ 0 & g_{22}\end{pmatrix}\\ &\qquad\qquad \frac{1}{|K_{a_{11},a_{21}} \times K_{a_{12},a_{22}}|} \sum_{(h,\ell)} \overline{\chi}_{V_1}\begin{pmatrix} g_{11} & h \\ 0 & g_{21}\end{pmatrix} \overline{\chi}_{V_2}\begin{pmatrix} g_{12} & \ell \\ 0 & g_{22}\end{pmatrix}.\end{aligned}$$

But this right hand side can be seen to equal (4.24), after one notes that

$$|{}^wP_{r_1,r_2} \cap P_{c_1,c_2}| = |G_{a_{11},a_{21},a_{12},a_{22}}| \cdot |K_{a_{11},a_{12}} \times K_{a_{21},a_{22}}| \cdot |K_{a_{11},a_{21}} \times K_{a_{12},a_{22}}| \cdot \#\{j \in \mathbb{F}_q^{a_{11}\times a_{22}}\}$$

and that the summands in (4.24) are independent of the matrix j in the summation. □

4.4. **Symmetric groups.** Finally, some payoff. Consider the tower of symmetric groups $G_n = \mathfrak{S}_n$, and $A = A(G_*) = A(\mathfrak{S})$. Denote by $1_{\mathfrak{S}_n}, \mathrm{sgn}_{\mathfrak{S}_n}$ the *trivial* and *sign* characters on $\mathfrak{S}_n$. For a partition λ of n, denote by $1_{\mathfrak{S}_\lambda}, \mathrm{sgn}_{\mathfrak{S}_\lambda}$ the trivial and sign characters restricted to the Young subgroup $\mathfrak{S}_\lambda = \mathfrak{S}_{\lambda_1} \times \mathfrak{S}_{\lambda_2} \times \cdots$, and denote by 1_λ the class function which is the characteristic function for the $\mathfrak{S}_n$-conjugacy class of permutations of cycle type λ.

Theorem 4.15. *Irreducible complex characters $\{\chi^\lambda\}$ of $\mathfrak{S}_n$ are indexed by partitions λ in* Par_n*, and one has a PSH-isomorphism, the* **Frobenius characteristic map***,*

$$A = A(\mathfrak{S}) \xrightarrow{\mathrm{ch}} \Lambda$$

that sends

$$\begin{array}{rcl} 1_{\mathfrak{S}_n} & \longmapsto & h_n \\ \mathrm{sgn}_{\mathfrak{S}_n} & \longmapsto & e_n \\ 1_{(n)} & \longmapsto & \frac{p_n}{n} \\ \chi^\lambda & \longmapsto & s_\lambda \\ \mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} 1_{\mathfrak{S}_\lambda} & \longmapsto & h_\lambda \\ \mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \mathrm{sgn}_{\mathfrak{S}_\lambda} & \longmapsto & e_\lambda \\ 1_\lambda & \longmapsto & \frac{p_\lambda}{z_\lambda} \end{array}$$

and where the involution ω on Λ corresponds under ch^{-1} to the map on each virtual character space $R(\mathfrak{S}_n)$ given by tensoring with the sign character $\mathrm{sgn}_{\mathfrak{S}_n}$. Here, z_λ is defined as in Proposition 2.30.

Proof. Corollary 4.14 implies that the set $\Sigma = \bigsqcup_{n\geq 0} \mathrm{Irr}(\mathfrak{S}_n)$ gives a PSH-basis for A. Since a character χ of $\mathfrak{S}_n$ has

$$\Delta(\chi) = \bigoplus_{i+j=n} \mathrm{Res}_{\mathfrak{S}_i\times\mathfrak{S}_j}^{\mathfrak{S}_n} \chi, \tag{4.25}$$

such an element χ is never primitive for $n \geq 2$. Hence the unique irreducible character $\rho = 1_{\mathfrak{S}_1}$ of $\mathfrak{S}_1$ is the only element of $\mathcal{C} = \Sigma \cap \mathfrak{p}$.

Thus Theorem 3.18 tells us that there are two isomorphisms $A \to \Lambda$, each of which sends Σ to the PSH-basis of Schur functions $\{s_\lambda\}$ for Λ. It also tells us that we can pin down one of the two isomorphisms to call ch, by insisting that it map the two characters $1_{\mathfrak{S}_2}, \mathrm{sgn}_{\mathfrak{S}_2}$ in $\mathrm{Irr}(\mathfrak{S}_2)$ to h_2, e_2 (and not e_2, h_2).

Bearing in mind the coproduct formula (4.25), and the fact that $1_{\mathfrak{S}_n}, \mathrm{sgn}_{\mathfrak{S}_n}$ restrict, respectively, to trivial and sign characters of $\mathfrak{S}_i \times \mathfrak{S}_j$ for $i + j = n$, one finds that for $n \geq 2$ one has $\mathrm{sgn}_{\mathfrak{S}_2}^\perp$ annihilating $1_{\mathfrak{S}_n}$, and $1_{\mathfrak{S}_2}^\perp$ annihilating $\mathrm{sgn}_{\mathfrak{S}_n}$. Therefore Theorem 3.16(b) implies $1_{\mathfrak{S}_n}, \mathrm{sgn}_{\mathfrak{S}_n}$ are sent under ch to h_n, e_n. Then the fact that $\mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} 1_{\mathfrak{S}_\lambda}, \mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \mathrm{sgn}_{\mathfrak{S}_\lambda}$ are sent to h_λ, e_λ follows via induction products.

For the assertion about $1_{(n)}$, note that it is primitive in A for $n \geq 1$, because as a class function, the indicator function of n-cycles vanishes upon restriction to $\mathfrak{S}_i \times \mathfrak{S}_j$ for $i + j = n$ if both $i, j \geq 1$; these subgroups contain no n-cycles. Hence Corollary 3.8 implies that $\mathrm{ch}(1_{(n)})$ is a scalar multiple of p_n. To pin down the scalar, note $p_n = m_{(n)}$ so $(h_n, p_n)_\Lambda = (h_n, m_n)_\Lambda = 1$, while $\mathrm{ch}^{-1}(h_n) = 1_{\mathfrak{S}_n}$ has

$$(1_{\mathfrak{S}_n}, 1_{(n)}) = \frac{1}{n!} \cdot (n-1)! = \frac{1}{n}.$$

Thus $\mathrm{ch}(1_{(n)}) = \frac{p_n}{n}$. The fact that $\mathrm{ch}(1_\lambda) = \frac{p_\lambda}{z_\lambda}$ then follows via induction product calculations, e.g. using (4.3).

The fact that ω corresponds on $A_n = R(\mathfrak{S}_n)$ to the operation $\chi \longmapsto \chi \otimes \mathrm{sgn}_{\mathfrak{S}_n}$, comes from Theorem 3.18(f), once one notes that the latter operation induces a (nontrivial) PSH-algebra automorphism of A. $\square$

Remark 4.16. The paper of Liulevicius [45] gives a very elegant alternate approach to the Frobenius map as a Hopf isomorphism $A(\mathfrak{S}) \xrightarrow{\mathrm{ch}} \Lambda$, inspired by equivariant K-theory and vector bundles over spaces which are finite sets of points!

4.5. **Wreath products.** Next consider the tower of groups $G_n = \mathfrak{S}_n[\Gamma]$ for a finite group Γ, and the Hopf algebra $A = A(G_*) =: A(\mathfrak{S}[\Gamma])$. Recall from the previous section that irreducible complex representations χ^λ of $\mathfrak{S}_n$ are indexed by partitions λ in Par_n. Index the irreducible complex representations of Γ as $\mathrm{Irr}(\Gamma) = \{\rho_1, \ldots, \rho_d\}$.

Definition 4.17. Define for a partition λ in Par_n and ρ in $\mathrm{Irr}(\Gamma)$ a representation $\chi^{\lambda,\rho}$ of $\mathfrak{S}_n[\Gamma]$ in which σ in $\mathfrak{S}_n$ and $\gamma = (\gamma_1, \ldots, \gamma_n)$ in Γ^n act on the space $\chi^\lambda \otimes (\rho^{\otimes n})$ as follows

$$\begin{aligned} \sigma(u \otimes (v_1 \otimes \cdots \otimes v_n)) &= \sigma(u) \otimes (v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)}) \\ \gamma(u \otimes (v_1 \otimes \cdots \otimes v_n)) &= u \otimes (\gamma_1 v_1 \otimes \cdots \otimes \gamma_n v_n) \end{aligned} \tag{4.26}$$

Theorem 4.18. *The irreducible $\mathbb{C}\mathfrak{S}_n[\Gamma]$-modules are the induced characters*

$$\chi^{\underline{\lambda}} := \mathrm{Ind}_{\mathfrak{S}_\lambda[\Gamma]}^{\mathfrak{S}_n[\Gamma]} \left(\chi^{\lambda^{(1)},\rho_1} \otimes \cdots \otimes \chi^{\lambda^{(d)},\rho_d} \right)$$

as $\underline{\lambda}$ runs through all functions

$$\begin{array}{ccc} \mathrm{Irr}(\Gamma) & \xrightarrow{\underline{\lambda}} & \mathrm{Par} \\ \rho_i & \longmapsto & \lambda^{(i)} \end{array}$$

with the property that $\sum_{i=1}^{d} |\lambda^{(i)}| = n$. Furthermore, one has a PSH-isomorphism

$$\begin{array}{ccc} A(\mathfrak{S}[\Gamma]) & \longrightarrow & \Lambda^{\otimes n} \\ \chi^{\underline{\lambda}} & \longmapsto & s_{\lambda^{(1)}} \otimes \cdots \otimes s_{\lambda^{(d)}}. \end{array}$$

Proof. We know from Corollary 4.14 that $A(\mathfrak{S}[\Gamma])$ is a PSH, with PSH-basis Σ given by the union of all irreducible characters of all groups $\mathfrak{S}_n[\Gamma]$. Therefore Theorem 3.10 tells us that $A(\mathfrak{S}[\Gamma]) \cong \bigotimes_{\rho \in \mathcal{C}} A(\mathfrak{S}[\Gamma])(\rho)$ where $\mathcal{C}$ is the set of irreducible characters which are also primitive. Just as in the case of $\mathfrak{S}_n$, it is clear from the definition of the coproduct that an irreducible character ρ of $\mathfrak{S}_n[\Gamma]$ is primitive if and only if $n = 1$, that $\mathfrak{S}_n[\Gamma] = \Gamma$, and ρ lies in $\mathrm{Irr}(\Gamma) = \{\rho_1, \ldots, \rho_d\}$.

The remaining assertions of the theorem will then follow from the definition of the induction product algebra structure on $A(\mathfrak{S}[\Gamma])$, once we have shown that there is a PSH-isomorphism sending

$$\begin{array}{ccc} A(\mathfrak{S}) & \longrightarrow & A(\mathfrak{S}[\Gamma])(\rho) \\ \chi^{\lambda} & \longmapsto & \chi^{\lambda,\rho}. \end{array} \tag{4.27}$$

Such an isomorphism comes from applying Proposition 4.4 to the semidirect product $\mathfrak{S}_n[\Gamma] = \mathfrak{S}_n \ltimes \Gamma^n$, so that $K = \Gamma^n, G = \mathfrak{S}_n$, and fixing $V = \rho^{\otimes n}$ as $\mathbb{C}\mathfrak{S}_n[\Gamma]$-module with structure as defined in (4.26). One obtains for each n, maps

$$A(\mathfrak{S}_n) \underset{\Psi}{\overset{\Phi}{\rightleftharpoons}} A(\mathfrak{S}_n[\Gamma])$$

where

$$\begin{array}{ccc} \chi & \overset{\Phi}{\longmapsto} & \chi \otimes (\rho^{\otimes n}) \\ \alpha & \overset{\Psi}{\longmapsto} & \mathrm{Hom}_{\mathbb{C}\Gamma^n}(\rho^{\otimes n}, \alpha). \end{array}$$

Taking the direct sum of these maps for all n gives maps $A(\mathfrak{S}) \underset{\Psi}{\overset{\Phi}{\rightleftharpoons}} A(\mathfrak{S}[\Gamma])$.

These maps are coalgebra morphisms because of their interaction with restriction to $\mathfrak{S}_i \times \mathfrak{S}_j$. Since Proposition 4.4(iii) gives the adjointness property that

$$(\chi, \Psi(\alpha))_{A(\mathfrak{S})} = (\Phi(\chi), \alpha)_{A(\mathfrak{S}[\Gamma])},$$

one concludes from the self-duality of $A(\mathfrak{S}), A(\mathfrak{S}[\Gamma])$ that Φ, Ψ are also algebra morphisms. Since they take genuine characters to genuine characters, they are PSH-morphisms. Since ρ being a simple $\mathbb{C}\Gamma$-module implies that $V = \rho^{\otimes n}$ is a simple $\mathbb{C}\Gamma^n$-module, Proposition 4.4(iv) shows that

$$(\Psi \circ \Phi)(\chi) = \chi \tag{4.28}$$

for all $\mathfrak{S}_n$-characters χ. Hence Φ is an injective PSH-morphism. Using adjointness, (4.28) also shows that Φ sends $\mathbb{C}\mathfrak{S}_n$-simples χ to $\mathbb{C}[\mathfrak{S}_n[\Gamma]]$-simples $\Phi(\chi)$:

$$(\Phi(\chi), \Phi(\chi))_{A(\mathfrak{S}[\Gamma])} = ((\Psi \circ \Phi)(\chi), \chi)_{A(\mathfrak{S})} = (\chi, \chi)_{A(\mathfrak{S})} = 1.$$

Since $\Phi(\chi) = \chi \otimes (\rho^{\otimes n})$ has $V = \rho^{\otimes n}$ as a constituent upon restriction to Γ^n, Frobenius Reciprocity shows that the irreducible character $\Phi(\chi)$ is a constituent of $\mathrm{Ind}_{\Gamma^n}^{\mathfrak{S}_n[\Gamma]} \rho^{\otimes n} = \rho^n$. Hence the entire image of Φ lies in $A(\mathfrak{S}[\Gamma])(\rho)$, and so Φ must restrict to an isomorphism as desired in (4.27). $\square$

One of Zelevinsky's sample applications of the theorem is this branching rule.

Corollary 4.19. *Given $\underline{\lambda} = (\lambda^{(1)}, \ldots, \lambda^{(d)})$ with $\sum_{i=1}^{d} |\lambda^{(i)}| = n$, one has*

$$\mathrm{Res}^{\mathfrak{S}_n[\Gamma]}_{\mathfrak{S}_{n-1}[\Gamma] \times \Gamma} \left(\chi^{\underline{\lambda}}\right) = \sum_{i=1}^{d} \sum_{\substack{\lambda^{(i)}_- \subseteq \lambda^{(i)}: \\ |\lambda^{(i)}/\lambda^{(i)}_-| = 1}} \chi^{(\lambda^{(1)}, \ldots, \lambda^{(i)}_-, \ldots, \lambda^{(d)})} \otimes \rho_i.$$

Example 4.20. For Γ a two-element group, so $\mathrm{Irr}(\Gamma) = \{\rho_1, \rho_2\}$ and $d = 2$, then

$$\mathrm{Res}^{\mathfrak{S}_6[\Gamma]}_{\mathfrak{S}_5[\Gamma] \times \Gamma} \left(\chi^{((3,1),(1,1))}\right) = \chi^{((3),(1,1))} \otimes \rho_1 + \chi^{((2,1),(1,1))} \otimes \rho_1 + \chi^{((3,1),(1))} \otimes \rho_2.$$

Proof of Corollary 4.19. By Theorem 4.18, this is equivalent to computing in the Hopf algebra $A := \Lambda^{\otimes d}$ the component of the coproduct of $s_{\lambda^{(1)}} \otimes \cdots \otimes s_{\lambda^{(d)}}$ that lies in $A_{n-1} \otimes A_1$. Working within each tensor factor Λ, the Pieri formula implies that the $\Lambda_{n-1} \otimes \Lambda_1$-component of $\Delta(s_\lambda)$ is

$$\sum_{\substack{\lambda_- \subseteq \lambda: \\ |\lambda/\lambda_-|=1}} s_{\lambda_-} \otimes \rho.$$

One must apply this in each of the d tensor factors of $A = \Lambda^{\otimes d}$, then sum on i. □

4.6. **General linear groups.** We now consider the tower of finite general linear groups $G_n = GL_n = GL_n(\mathbb{F}_q)$ and $A = A(G_*) =: A(GL)$. Corollary 4.14 tells us that $A(GL)$ is a PSH, with PSH-basis Σ given by the union of all irreducible characters of all groups GL_n. Therefore Theorem 3.10 tells us that

$$A(GL) \cong \bigotimes_{\rho \in \mathcal{C}} A(GL)(\rho) \tag{4.29}$$

where $\mathcal{C} = \Sigma \cap \mathfrak{p}$ is the set of primitive irreducible characters.

Definition 4.21. Call an irreducible representation ρ of GL_n *cuspidal* for $n \geq 1$ if it lies in $\mathcal{C}$, that is, its restriction to proper parabolic subgroups $P_{i,j}$ with $i+j = n$ and $i, j > 0$ contain no nonzero vectors which are $K_{i,j}$-invariant. Given an irreducible character σ of GL_n, say that $d(\sigma) = n$, and let $\mathcal{C}_n := \{\rho \in \mathcal{C} : d(\rho) = n\}$ for $n \geq 1$ denote the subset of cuspidal characters of GL_n.

Just as was the case for $\mathfrak{S}_1$ and $\mathfrak{S}_1[\Gamma] = \Gamma$, *every* irreducible character ρ of $GL_1(\mathbb{F}_q) = \mathbb{F}_q^\times$ is cuspidal. However, this does not exhaust the cuspidal characters. In fact, one can predict the number of cuspidal characters in $\mathcal{C}_n$, using knowledge of the number of conjugacy classes in GL_n. Let $\mathcal{F}$ denote the set of all nonconstant monic irreducible polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$. Let $\mathcal{F}_n := \{f \in \mathcal{F} : \deg(f) = n\}$ for $n \geq 1$.

Proposition 4.22. *The number $|\mathcal{C}_n|$ of cuspidal characters of $GL_n(\mathbb{F}_q)$ is the number of $|\mathcal{F}_n|$ of irreducible monic degree n polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$ with nonzero constant term.*

Proof. We show $|\mathcal{C}_n| = |\mathcal{F}_n|$ for $n \geq 1$ by induction on n. For the base case $n = 1$, just as with the families $G_n = \mathfrak{S}_n$ and $G_n = \mathfrak{S}_n[\Gamma]$, when $n = 1$ any irreducible character χ of $G_1 = GL_1(\mathbb{F}_q)$ gives a primitive element of $A = A(GL)$, and hence is cuspidal. Since $GL_1(\mathbb{F}_q) = \mathbb{F}_q^\times$ is abelian, there are $|\mathbb{F}_q^\times| = q - 1$ such cuspidal characters in $\mathcal{C}_1$, which agrees with the fact that there are $q - 1$ monic (irreducible) linear polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$, namely $\mathcal{F}_1 := \{f(x) = x - c : c \in \mathbb{F}_q^\times\}$.

In the inductive step, use the fact that the number $|\Sigma_n|$ of irreducible complex characters χ of $GL_n(\mathbb{F}_q)$ equals its number of conjugacy classes. These conjugacy classes are uniquely represented by *rational canonical forms*, which are parametrized by functions $\underline{\lambda} : \mathcal{F} \to \mathrm{Par}$ with the property that $\sum_{f \in \mathcal{F}} \deg(f)|\lambda(f)| = n$. On the other hand, (4.29) tells us that $|\Sigma_n|$ is similarly parametrized by the functions $\underline{\lambda} : \mathcal{C} \to \mathrm{Par}$ having the property that $\sum_{\rho \in \mathcal{C}} d(\rho)|\lambda(f)| = n$. Thus we have parallel disjoint decompositions

$$\begin{aligned} \mathcal{C} &= \bigsqcup_{n \geq 1} \mathcal{C}_n \qquad \text{where } \mathcal{C}_n = \{\rho \in \mathcal{C} : d(\rho) = n\} \\ \mathcal{F} &= \bigsqcup_{n \geq 1} \mathcal{F}_n \qquad \text{where } \mathcal{F}_n = \{f \in \mathcal{F} : \deg(f) = n\} \end{aligned}$$

and hence an equality for all $n \geq 1$

$$\left| \left\{ \mathcal{C} \xrightarrow{\underline{\lambda}} \mathrm{Par} : \quad \sum_{\rho \in \mathcal{C}} d(\rho)|\lambda(f)| = n \right\} \right| = |\Sigma_n| = \left| \left\{ \mathcal{F} \xrightarrow{\underline{\lambda}} \mathrm{Par} : \quad \sum_{f \in \mathcal{F}} \deg(f)|\lambda(f)| = n \right\} \right|.$$

Since there is only one partition $\lambda = (1)$ of having $|\lambda| = 1$, this leads to parallel recursions

$$|\mathcal{C}_n| = |\Sigma_n| - \left| \left\{ \bigsqcup_{i=1}^{n-1} \mathcal{C}_i \xrightarrow{\underline{\lambda}} \mathrm{Par} : \quad \sum_{\rho} d(\rho)|\lambda(\rho)| = n \right\} \right|$$

$$|\mathcal{F}_n| = |\Sigma_n| - \left| \left\{ \bigsqcup_{i=1}^{n-1} \mathcal{F}_i \xrightarrow{\underline{\lambda}} \mathrm{Par} : \quad \sum_{f} \deg(f)|\lambda(f)| = n \right\} \right|$$

and induction implies that $|\mathcal{C}_n| = |\mathcal{F}_n|$. □

Example 4.23. Taking $q = 2$, let us list the sets $\mathcal{F}_n$ of monic irreducible polynomials $f(x) \neq x$ in $\mathbb{F}_2[x]$ of degree n for $n \leq 3$, so that we know how many cuspidal characters of $GL_n(\mathbb{F}_q)$ in $\mathcal{C}_n$ to expect:

$$\begin{aligned} \mathcal{F}_1 &= \{x+1\} \\ \mathcal{F}_2 &= \{x^2+x+1\} \\ \mathcal{F}_3 &= \{x^3+x+1, x^3+x^2+1\} \end{aligned}$$

Thus we expect

- one cuspidal character of $GL_1(\mathbb{F}_2)$, namely $\rho_1 (= 1_{GL_1(\mathbb{F}_2)})$,
- one cuspidal character ρ_2 of $GL_2(\mathbb{F}_2)$, and
- two cuspidal characters ρ_3, ρ_3' of $GL_3(\mathbb{F}_2)$.

We will say more about ρ_2, ρ_3, ρ_3' in the next section.

Exercise 4.24. (a) Show that for $n \geq 2$,

$$|\mathcal{C}_n| (= |\mathcal{F}_n|) = \frac{1}{n} \sum_{d \text{ dividing } n} \mu\left(\frac{n}{d}\right) q^d \tag{4.30}$$

where $\mu(m)$ is the *number-theoretic Möbius function* of m, that is $\mu(m) = (-1)^d$ if $m = p_1 \cdots p_d$ for d distinct primes, and $\mu(m) = 0$ if m is not squarefree.
(b) Show that (4.30) also counts the *necklaces* with n beads of q colors (=equivalence classes under the $\mathbb{Z}/n\mathbb{Z}$-action of cyclic rotation on sequences $(a_1, \ldots, a_n)$ in $\mathbb{F}_q^n$) which are *primitive* in the sense that no nontrivial rotation fixes any of the sequences within the equivalence class. For example, when $q = 2$, here are representatives of these primitive necklaces for $n = 2, 3, 4$:

$$\begin{aligned} n = 2 :& \{(0,1)\} \\ n = 3 :& \{(0,0,1), (0,1,1)\} \\ n = 4 :& \{(0,0,0,1), (0,0,1,1), (0,1,1,1)\}. \end{aligned}$$

4.7. **Steinberg's unipotent characters.** Not surprisingly, the (cuspidal) character $\iota := 1_{GL_1}$ of $GL_1(\mathbb{F}_q)$ plays a distinguished role. The parabolic subgroup $P_{(1^n)}$ is the Borel subgroup B of upper triangular matrices, and the subalgebra $A(GL)(\iota)$ of $A(GL)$ is the $\mathbb{Z}$-span of the irreducible characters σ that appear as constituents of $\iota^n = \operatorname{Ind}_B^{GL_n} 1_B = \mathbb{C}[GL_n/B]$ for some n.

Definition 4.25. An irreducible character σ of GL_n appearing as a constituent of $\operatorname{Ind}_B^{GL_n} 1_B = \mathbb{C}[GL_n/B]$ is called a *unipotent character*. Equivalently, by Frobenius reciprocity, σ is unipotent if it contains a nonzero B-invariant vector.

In particular, 1_{GL_n} is a unipotent character of GL_n for each n.

Proposition 4.26. *One can choose $\Lambda \cong A(GL)(\iota)$ in Theorem 3.18(g) so that $h_n \longmapsto 1_{GL_n}$.*

Proof. Theorem 3.16(a) tells us $\iota^2 = \operatorname{Ind}_B^{GL_2} 1_B$ must have exactly two irreducible constituents, one of which is 1_{GL_2}; call the other one St_2. Choose the isomorphism so as to send $h_2 \longmapsto 1_{GL_2}$. Then $h_n \mapsto 1_{GL_n}$ follows from the claim that $\mathrm{St}_2^\perp(1_{GL_n}) = 0$ for $n \geq 2$: one has

$$\Delta(1_{GL_n}) = \sum_{i+j=n} \left(\operatorname{Res}_{P_{i,j}}^{G_n} 1_{GL_n}\right)^{K_{i,j}} = \sum_{i+j=n} 1_{GL_i} \otimes 1_{GL_j}$$

so that $\mathrm{St}_2^\perp(1_{GL_n}) = (\mathrm{St}_2, 1_{GL_2}) 1_{GL_{n-2}} = 0$ since $\mathrm{St}_2 \neq 1_{GL_2}$.

∎

This subalgebra $A(GL)(\iota)$, and the unipotent characters χ_q^λ corresponding under this isomorphism to the Schur functions s_λ, were introduced by Steinberg [74]. He wrote down χ_q^λ as a virtual sum of induced characters $\operatorname{Ind}_{P_\alpha}^{GL_n} 1_{P_\alpha} (= 1_{G_{\alpha_1}} \cdots 1_{G_{\alpha_\ell}})$, modelled on the Jacobi-Trudi determinantal expression for $s_\lambda = \det(h_{\lambda_i - i + j})$. Note that $\operatorname{Ind}_{P_\alpha}^{GL_n} 1_{P_\alpha}$ is the transitive permutation representation $\mathbb{C}[G/P_\alpha]$ for GL_n permuting the *finite partial flag variety* G/P_α, that is, the set of *α-flags of subspaces*

$$\{0\} \subset V_{\alpha_1} \subset V_{\alpha_1+\alpha_2} \subset \cdots \subset V_{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1}} \subset \mathbb{F}_q^n$$

where $\dim_{\mathbb{F}_q} V_d = d$ in each case. This character has dimension equal to $|G/P_\alpha|$, with formula given by the *q-multinomial coefficient* (see e.g. Stanley [72, §1.7]):

$$\begin{bmatrix} n \\ \alpha \end{bmatrix}_q = \frac{[n]!_q}{[\alpha_1]!_q \cdots [\alpha_\ell]!_q}$$

where $[n]!_q := [n]_q[n-1]_q \cdots [2]_q[1]_q$ and $[n]_q := 1 + q + \cdots + q^{n-1} = \frac{q^n-1}{q-1}$.

Our terminology St_2 is motivated by the $n = 2$ special case of the *Steinberg character* St_n, which is the unipotent character corresponding under the isomorphism in Proposition 4.26 to $e_n = s_{(1^n)}$. It can be defined by the virtual sum

$$\mathrm{St}_n := \chi_q^{(1^n)} = \sum_\alpha (-1)^{n-\ell(\alpha)} \mathrm{Ind}_{P_\alpha}^{GL_n} 1_{P_\alpha}$$

in which the sum runs through all compositions α of n. This turns out to be the genuine character for $GL_n(\mathbb{F}_q)$ acting on the top homology group of its *Tits building*: the simplicial complex whose vertices are nonzero proper subspaces V of $\mathbb{F}_q^n$, and whose simplices correspond to flags of nested subspaces. One needs to know that this Tits building has only top homology, so that one can deduce the above character formula from the Hopf trace formula; see Björner [12].

4.8. **Examples: $GL_2(\mathbb{F}_2)$ and $GL_3(\mathbb{F}_2)$.** Let's get our hands dirty.

Example 4.27. For $n = 2$, there are two unipotent characters, $\chi_q^{(2)} = 1_{GL_2}$ and

$$\mathrm{St}_2 := \chi_q^{(1,1)} = 1_{GL_1}^2 - 1_{GL_2} = \mathrm{Ind}_B^{GL_2} 1_B - 1_{GL_2} \tag{4.31}$$

since the Jacobi-Trudi formula gives $s_{(1,1)} = \det \begin{bmatrix} h_1 & h_2 \\ 1 & h_1 \end{bmatrix} = h_1^2 - h_2$. The description (4.31) for this Steinberg character St_2 shows that it has dimension

$$|GL_2/B| - 1 = (q+1) - 1 = q$$

and that one can think of it as follows: consider the permutation action of GL_2 on the $q+1$ lines $\{\ell_0, \ell_1, \ldots, \ell_q\}$ in the projective space $\mathbb{P}^1_{\mathbb{F}_q} = GL_2(\mathbb{F}_q)/B$, and take the invariant subspace perpendicular to the sum of basis elements $e_{\ell_0} + \cdots + e_{\ell_q}$.

Example 4.28. Continuing the previous example, but taking $q = 2$, we find that we have constructed two unipotent characters: $1_{GL_2} = \chi_{q=2}^{(2)}$ of dimension 1, and $\mathrm{St}_2 = \chi_{q=2}^{(1,1)}$ of dimension $q = 2$. This lets us identify the unique cuspidal character ρ_2 of $GL_2(\mathbb{F}_2)$, using knowledge of the character table of $GL_2(\mathbb{F}_2) \cong \mathfrak{S}_3$:

		$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$
$1_{GL_2} = \chi_{q=2}^{(2)}$	unipotent	1	1	1
$\mathrm{St}_2 = \chi_{q=2}^{(1,1)}$	unipotent	2	0	-1
ρ_2	cuspidal	1	-1	1

In other words, the cuspidal character ρ_2 of $GL_2(\mathbb{F}_2)$ corresponds under the isomorphism $GL_2(\mathbb{F}_2) \cong \mathfrak{S}_3$ to the sign character $\mathrm{sgn}_{\mathfrak{S}_3}$.

Example 4.29. Continuing the previous example to $q = 2$ and $n = 3$ lets us analyze the irreducible characters of $GL_3(\mathbb{F}_2)$. Recalling our labelling $\rho_1, \rho_2, \rho_3, \rho_3'$ from Example 4.23 of the cuspidal characters of $GL_n(\mathbb{F}_2)$ for $n = 1, 2, 3$, Zelevinsky's Theorem 3.10 tells us that the $GL_3(\mathbb{F}_2)$-irreducible characters should be labelled by function $\{\rho_1, \rho_2, \rho_3, \rho_3'\} \xrightarrow{\underline{\lambda}} \mathrm{Par}$ for which

$$1 \cdot |\lambda(\rho_1)| + 2 \cdot |\lambda(\rho_2)| + 3 \cdot |\lambda(\rho_3)| + 3 \cdot |\lambda(\rho_3')| = 3$$

We will label such an irreducible character $\chi^{\underline{\lambda}} = \chi^{(\lambda(\rho_1), \lambda(\rho_2), \lambda(\rho_3), \lambda(\rho_3'))}$.

Three of these irreducibles will be the unipotent characters, mapping under the isomorphism from Proposition 4.26 as follows:

- $s_{(3)} = h_3 \longmapsto \chi^{((3), \varnothing, \varnothing, \varnothing)} = 1_{GL_3}$ of dimension 1.

- $$s_{(2,1)} = \det \begin{bmatrix} h_2 & h_3 \\ 1 & h_1 \end{bmatrix} = h_2 h_1 - h_3 \longmapsto \chi^{((2,1),\varnothing,\varnothing,\varnothing)} = \mathrm{Ind}_{P_{2,1}}^{GL_3} 1_{P_{2,1}} - 1_{GL_3},$$

  of dimension $\begin{bmatrix} 3 \\ 2,1 \end{bmatrix}_q - \begin{bmatrix} 3 \\ 3 \end{bmatrix}_q = [3]_q - 1 = q^2 + q \overset{q=2}{\rightsquigarrow} 6.$
- Lastly,

$$s_{(1,1,1)} = \det \begin{bmatrix} h_1 & h_2 & h_3 \\ 1 & h_1 & h_2 \\ 0 & 1 & h_1 \end{bmatrix} = h_1^3 - h_2 h_1 - h_1 h_2 + h_3$$
$$\longmapsto \mathrm{St}_3 = \chi^{((1,1,1),\varnothing,\varnothing,\varnothing)} = \mathrm{Ind}_B^{GL_3} 1_B - \mathrm{Ind}_{P_{1,2}}^{GL_3} 1_{P_{1,2}} - \mathrm{Ind}_{P_{2,1}}^{GL_3} 1_{P_{2,1}} + 1_{GL_3}.$$

  of dimension

$$\begin{bmatrix} 3 \\ 1,1,1 \end{bmatrix}_q - \begin{bmatrix} 3 \\ 2,1 \end{bmatrix}_q - \begin{bmatrix} 3 \\ 1,2 \end{bmatrix}_q - \begin{bmatrix} 3 \\ 3 \end{bmatrix}_q$$
$$= [3]!_q - [3]_q - [3]_q + 1 = q^3 \overset{q=2}{\rightsquigarrow} 8.$$

There should also be one non-unipotent, non-cuspidal character, namely

$$\chi^{((1),(1),\varnothing,\varnothing)} = \rho_1 \rho_2 = \mathrm{Ind}_{P_{1,2}}^{GL_3} \mathrm{Infl}_{GL_1 \times GL_2}^{P_{1,2}} (1_{GL_1} \otimes \rho_2)$$

having dimension $\begin{bmatrix} 3 \\ 1,2 \end{bmatrix}_q \cdot 1 \cdot 1 = [3]_q \overset{q=2}{\rightsquigarrow} 7.$

Finally, we expect cuspidal characters $\rho_3 = \chi^{(\varnothing,\varnothing,(1),\varnothing)}$, $\rho_3' = \chi^{(\varnothing,\varnothing,\varnothing,(1))}$, whose dimensions d_3, d_3' can be deduced from the equation

$$1^2 + 6^2 + 8^2 + 7^2 + d_3^2 + (d_3')^2 = |GL_3(\mathbb{F}_2)| = \left[(q^3 - q^0)(q^3 - q^1)(q^3 - q^2)\right]_{q=2} = 168.$$

This forces $d_3^2 + (d_3')^2 = 18$, whose only solution in positive integers is $d_3 = d_3' = 3$.

We can check our predictions of the dimensions for the various $GL_3(\mathbb{F}_2)$-irreducible characters since $GL_3(\mathbb{F}_2)$ is the finite simple group of order 168 (also isomorphic to $PSL_2(\mathbb{F}_7)$), with known character table (see James and Liebeck [38, p. 318]):

	centralizer order	168	8	4	3	7	7
	unipotent?/cuspidal?						
$1_{GL_3} = \chi^{((3),\varnothing,\varnothing,\varnothing)}$	unipotent	1	1	1	1	1	1
$\chi^{((2,1),\varnothing,\varnothing,\varnothing)}$	unipotent	6	2	0	0	-1	-1
$\mathrm{St}_3 = \chi^{((1,1,1),\varnothing,\varnothing,\varnothing)}$	unipotent	8	0	0	-1	1	1
$\chi^{((1),(1),\varnothing,\varnothing)}$		7	-1	-1	1	0	0
$\rho_3 = \chi^{(\varnothing,\varnothing,\varnothing,(1))}$	cuspidal	3	-1	1	0	α	$\overline{\alpha}$
$\rho_3' = \chi^{(\varnothing,\varnothing,(1),\varnothing)}$	cuspidal	3	-1	1	0	$\overline{\alpha}$	α

Here $\alpha := -1/2 + i\sqrt{7}/2$.

Remark 4.30. It is known (see e.g. Bump [13, Cor. 7.4]) that, for $n \geq 2$, the dimension of any cuspidal irreducible character ρ of $GL_n(\mathbb{F}_q)$ is

$$(q^{n-1} - 1)(q^{n-2} - 1) \cdots (q^2 - 1)(q - 1).$$

Note that when $q = 2$,

- for $n = 2$ this gives $2^1 - 1 = 1$ for the dimension of ρ_2, and
- for $n = 3$ it gives $(2^2 - 1)(2 - 1) = 3$ for the dimensions of ρ_3, ρ_3',

agreeing with our calculations above. Much more is known about the character table of $GL_n(\mathbb{F}_q)$; see Remark 4.42 below, Zelevinsky [81, Chap. 11], and Macdonald [49, Chap. IV]

4.9. **The Hall algebra.** There is another interesting Hopf subalgebra (and quotient Hopf algebra) of $A(GL)$, related to unipotent conjugacy classes in $GL_n(\mathbb{F}_q)$.

Definition 4.31. Say that an element g in $GL_n(\mathbb{F}_q)$ is *unipotent* if its eigenvalues are all equal to 1. Denote by $\mathcal{H}_n$ the $\mathbb{C}$-subspace of $R_{\mathbb{C}}(GL_n)$ consisting of those class functions which are supported only on unipotent conjugacy classes, and let $\mathcal{H} = \bigoplus_{n\geq 0} \mathcal{H}_n$ as a $\mathbb{C}$-subspace of $A_{\mathbb{C}}(GL) = \bigoplus_{n\geq 0} R_{\mathbb{C}}(GL_n)$.

Proposition 4.32. *The subspace $\mathcal{H}$ is a Hopf-subalgebra of $A_{\mathbb{C}}(GL)$, which is graded, connected, and of finite type, and self-dual with respect to the inner product on class functions inherited from $A_{\mathbb{C}}(GL)$. It is also a quotient Hopf algebra of $A_{\mathbb{C}}(GL)$, as the $\mathbb{C}$-linear surjection $A_{\mathbb{C}}(GL) \twoheadrightarrow \mathcal{H}$ restricting class functions to unipotent classes has kernel $\mathcal{H}^\perp$ which is both an ideal and a two-sided coideal.*

Proof. Given two class functions χ_i, χ_j on GL_i, GL_j and g in GL_{i+j}, one has

$$(\chi_i \cdot \chi_j)(g) = \frac{1}{|P_{i,j}|} \sum_{\substack{h \in GL_{i+j}: \\ h^{-1}gh = \begin{bmatrix} g_i & * \\ 0 & g_j \end{bmatrix} \in P_{i,j}}} \chi_i(g_i)\chi_j(g_j). \tag{4.32}$$

Since g is unipotent if and only if $h^{-1}gh$ is unipotent if and only if both g_i, g_j are unipotent, the formula (4.32) shows both that $\mathcal{H}$ is a subalgebra and that $\mathcal{H}^\perp$ is a two-sided ideal: χ_i and χ_j are both supported only on unipotent classes if and only if the same holds for $\chi_i \cdot \chi_j$. Similarly, for class functions χ on GL_n and (g_i, g_j) in $GL_{i,j} = GL_i \times GL_j$, one has

$$\Delta(\chi)(g_i, g_j) = \frac{1}{q^{ij}} \sum_{k \in \mathbb{F}_q^{i\times j}} \chi \begin{bmatrix} g_i & k \\ 0 & g_j \end{bmatrix}$$

using (4.10). This shows both that $\mathcal{H}$ is a sub-coalgebra of $A = A_{\mathbb{C}}(GL)$

$$\Delta\mathcal{H} \subset \mathcal{H} \otimes \mathcal{H}$$

and that $\mathcal{H}^\perp$ is a two-sided coideal

$$\Delta(\mathcal{H}^\perp) \subset \mathcal{H}^\perp \otimes A + A \otimes \mathcal{H}^\perp$$

since it shows that χ is supported only on unipotent classes if and only if $\Delta(\chi)$ vanishes on (g_1, g_2) that have either g_1 or g_2 non-unipotent. The rest follows. ◻

The subspace $\mathcal{H}$ is called the *Hall algebra*. It has an obvious orthogonal $\mathbb{C}$-basis, with interesting structure constants.

Definition 4.33. Given a partition λ of n, let J_λ denote the GL_n-conjugacy class of unipotent matrices whose *Jordan type* (= Jordan block sizes) is given by λ, and let $z_\lambda(q)$ denote the size of this conjugacy class J_λ.

The indicator class functions $\{1_{J_\lambda}\}_{\lambda\in\mathrm{Par}}$ form a $\mathbb{C}$-basis for $\mathcal{H}$ whose multiplicative structure constants are called the *Hall coefficients* $g^\lambda_{\mu,\nu}(q)$:

$$1_{J_\mu} 1_{J_\nu} = \sum_\lambda g^\lambda_{\mu,\nu}(q)\ 1_{J_\lambda}.$$

Because the dual basis to $\{1_{J_\lambda}\}$ is $\{z_\lambda(q)^{-1} 1_{J_\lambda}\}$, self-duality of $\mathcal{H}$ shows that the Hall coefficients are (essentially) also structure constants for the comultiplication:

$$\Delta 1_{J_\lambda} = \sum_{\mu,\nu} g^\lambda_{\mu,\nu}(q) \frac{z_\mu(q) z_\nu(q)}{z_\lambda(q)} \cdot 1_{J_\mu} \otimes 1_{J_\nu}.$$

The Hall coefficient $g^\lambda_{\mu,\nu}(q)$ has the following interpretation.

Proposition 4.34. *Fix any g in $GL_n(\mathbb{F}_q)$ acting unipotently on $\mathbb{F}_q^n$ with Jordan type λ. Then $g^\lambda_{\mu,\nu}(q)$ counts the g-stable $\mathbb{F}_q$-subspaces $V \subset \mathbb{F}_q^n$ for which the restriction $g|V$ acts with Jordan type μ, and the induced map $\bar{g}$ on the quotient space $\mathbb{F}_q^n/V$ has Jordan type ν.*

Proof. Given μ, ν partitions of i, j with $i+j=n$, taking χ_i, χ_j equal to $1_{J_\mu}, 1_{J_\nu}$ in (4.32) shows that for any g in GL_n, the value of $\left(1_{J_\mu} \cdot 1_{J_\nu}\right)(g)$ is given by

$$\frac{1}{|P_{i,j}|}\left|\left\{h \in GL_n : h^{-1}gh = \begin{bmatrix} g_i & * \\ 0 & g_j \end{bmatrix} \text{ with } g_i \in J_\mu, g_j \in J_\nu\right\}\right| \tag{4.33}$$

Let S denote the set appearing in (4.33), and let $\mathbb{F}_q^i$ denote the i-dimensional subspace of $\mathbb{F}_q^n$ spanned by the first i standard basis vectors. Note that the condition on an element h in S saying that $h^{-1}gh$ is in block upper-triangular form can be re-expressed by saying that the subspace $V := h(\mathbb{F}_q^i)$ is g-stable. One then sees that the map $h \overset{\varphi}{\longmapsto} V = h(\mathbb{F}_q^i)$ surjects S onto the set of i-dimensional g-stable subspaces V of $\mathbb{F}_q^n$ for which $g|V$ and $\bar{g}$ are unipotent of types μ, ν, respectively. Furthermore, for any particular such V, its fiber $\varphi^{-1}(V)$ in S is the stabilizer within GL_n of V, which is conjugate to $P_{i,j}$, and hence has cardinality $|\varphi^{-1}(V)| = |P_{i,j}|$. This proves the assertion of the proposition. $\square$

The Hall algebra $\mathcal{H}$ will turn out to be isomorphic to the ring $\Lambda_\mathbb{C}$ of symmetric functions with $\mathbb{C}$ coefficients, via a composite φ of three maps

$$\Lambda_\mathbb{C} \longrightarrow A(GL)(\iota)_\mathbb{C} \longrightarrow A(GL)_\mathbb{C} \longrightarrow \mathcal{H}$$

in which the first map is the isomorphism from Proposition 4.26, the second is inclusion, and the third is the quotient map from Proposition 4.32.

Theorem 4.35. *The above composite φ is a Hopf algebra isomorphism, sending*

$$\begin{array}{rcl} h_n & \longmapsto & \sum_{\lambda \in \operatorname{Par}_n} 1_{J_\lambda}, \\ e_n & \longmapsto & q^{\binom{n}{2}} 1_{J_{(1^n)}}, \\ p_n & \longmapsto & \sum_{\lambda \in \operatorname{Par}_n} (q;q)_{\ell(\lambda)} 1_{J_\lambda}, \end{array}$$

where

$$(x;q)_\ell := (1-x)(1-qx)(1-q^2x)\cdots(1-q^{\ell-1}x).$$

Proof. That φ is a Hopf morphism follows because it is a composite of three such morphisms. We claim that once one shows the formula for the (nonzero) image of $\varphi(p_n)$ given above is correct, then this will already show φ is an isomorphism, by the following argument. Note first that $\Lambda_\mathbb{C}$ and $\mathcal{H}$ both have dimension $|\operatorname{Par}_n|$ for their n^{th} homogeneous components, so it suffices to show that the graded map φ is injective. On the other hand, both $\Lambda_\mathbb{C}$ and $\mathcal{H}$ are (graded, connected, finite type) *self-dual* Hopf algebras, so Theorem 3.7 says that each is the symmetric algebra on its space of primitive elements. Thus it suffices to check that φ is injective when restricted to their subspaces of primitives. For $\Lambda_\mathbb{C}$, by Corollary 3.8 the primitives are spanned by $\{p_1, p_2, \ldots\}$, with only one basis element in each degree. Hence φ is injective on the subspace of primitives if and only if it does not annihilate any p_n.

Thus it only remains to show the above formulas for the images of h_n, e_n, p_n under φ. This is clear for h_n, since Proposition 4.26 shows that it maps under the first two composites to the indicator function 1_{GL_n} which then restricts to the sum of indicators $\sum_{\lambda \in \operatorname{Par}_n} 1_{J_\lambda}$ in $\mathcal{H}$. For e_n, p_n, we resort to generating functions. Let $\tilde{h}_n, \tilde{e_n}, \tilde{p}_n$ denote the three putative images in $\mathcal{H}$ of h_n, e_n, p_n, appearing on the right side in the theorem, and define generating functions in $\mathcal{H}[[t]]$

$$\tilde{H}(t) := \sum_{n \geq 0} \tilde{h}_n t^n, \quad \tilde{E}(t) := \sum_{n \geq 0} \tilde{e}_n t^n, \quad \tilde{P}(t) := \sum_{n \geq 0} \tilde{p}_{n+1} t^n.$$

We wish to show that φ maps $H(t), E(t), P(t)$ in $\Lambda[[t]]$ to these three generating functions. Since we have already shown this is correct for $H(t)$, by (2.10), (2.20), it suffices to check that in $\mathcal{H}[[t]]$ one has

$$\begin{array}{rcl} \tilde{H}(t)\tilde{E}(-t) = 1, & \text{or equivalently,} & \sum_{k=0}^n (-1)^k \tilde{e}_k \tilde{h}_{n-k} = \delta_{0,n} \\ \tilde{H}'(t)\tilde{E}(-t) = \tilde{P}(t), & \text{or equivalently,} & \sum_{k=0}^n (-1)^k (n-k) \tilde{e}_k \tilde{h}_{n-k} = \tilde{p}_n. \end{array}$$

Thus it would be helpful to evaluate the class function $\tilde{e}_k \tilde{h}_{n-k}$. Note that a unipotent g in GL_n having ℓ Jordan blocks has an ℓ-dimensional 1-eigenspace, so that

$$(\tilde{e}_k \tilde{h}_{n-k})(g) = q^{\binom{k}{2}} \cdot (1_{J_{(1^k)}} \cdot \tilde{h}_{n-k})(g) = q^{\binom{k}{2}} \begin{bmatrix} \ell \\ k \end{bmatrix}_q$$

where

$$\begin{bmatrix} \ell \\ k \end{bmatrix}_q = \frac{(q;q)_\ell}{(q;q)_k (q;q)_{\ell-k}}$$

is the *q-binomial coefficient* counting k-dimensional $\mathbb{F}_q$-subspaces V of an ℓ-dimensional $\mathbb{F}_q$-vector space; see, e.g., [72, §1.7]. Thus one needs for $\ell \geq 1$ that

$$\sum_{k=0}^{\ell} (-1)^k q^{\binom{k}{2}} \begin{bmatrix} \ell \\ k \end{bmatrix}_q = 0, \tag{4.34}$$

$$\sum_{k=0}^{\ell} (-1)^k (n-k) q^{\binom{k}{2}} \begin{bmatrix} \ell \\ k \end{bmatrix}_q = (q;q)_\ell. \tag{4.35}$$

Identity (4.34) comes from setting $x = 1$ in the *q-binomial theorem* [72, Exer. 3.119]:

$$\sum_{k=0}^{\ell} (-1)^k q^{\binom{k}{2}} \begin{bmatrix} \ell \\ k \end{bmatrix}_q x^{\ell-k} = (x-1)(x-q)(x-q^2)\cdots(x-q^{\ell-1}).$$

Identity (4.35) comes from taking $\frac{d}{dx}$ in the q-binomial theorem, then setting $x = 1$, and finally adding $(n-\ell)$ times (4.34). $\square$

We next indicate, without proof, how $\mathcal{H}$ relates to the classical Hall algebra.

Definition 4.36. The usual *Hall algebra*, or what Schiffmann [63, §2.3] calls *Steinitz's classical Hall algebra* (see also Macdonald [49, Chap. II]), has $\mathbb{Z}$-basis elements $\{u_\lambda\}_{\lambda \in \mathrm{Par}}$, with the multiplicative structure constants $g^\lambda_{\mu,\nu}(p)$ in

$$u_\mu u_\nu = \sum_\lambda g^\lambda_{\mu,\nu}(p)\ u_\lambda$$

defined as follows: fix a finite abelian p-group L of *type* λ, meaning that

$$L \cong \bigoplus_{i=1}^{\ell(\lambda)} \mathbb{Z}/p^{\lambda_i}\mathbb{Z},$$

and let $g^\lambda_{\mu,\nu}(p)$ be the number of subgroups M of L of type μ, for which the quotient $N := L/M$ is of type ν. In other words, $g^\lambda_{\mu,\nu}(p)$ counts, for a fixed abelian p-group L of type λ, the number of short exact sequences $0 \to M \to L \to N \to 0$ in which M, N have types μ, ν, respectively.

We claim that when one takes the finite field $\mathbb{F}_q$ of order $q = p$ a *prime*, the map

$$u_\lambda \longmapsto 1_{J_\lambda} \tag{4.36}$$

gives an isomorphism from this classical Hall algebra to the $\mathbb{Z}$-algebra $\mathcal{H}_\mathbb{Z} \subset \mathcal{H}$. The key point is *Hall's Theorem*, a non-obvious statement for which Macdonald includes two proofs in [49, Chap. II], one of them due to Zelevinsky[18]. To state it, we first recall some notions about discrete valuation rings.

Definition 4.37. A *discrete valuation ring (DVR)* $\mathfrak{o}$ is a principal ideal domain having only one maximal ideal $\mathfrak{m}$, with quotient $k = \mathfrak{o}/\mathfrak{m}$ called its *residue field*.

The structure theorem for finitely generated modules over a PID implies that an $\mathfrak{o}$-module L with finite composition series of composition length n must have $L \cong \bigoplus_{i=1}^{\ell(\lambda)} \mathfrak{o}/\mathfrak{m}^{\lambda_i}$ for some partition λ of n; say L has *type* λ in this situation.

Here are the two crucial examples for us.

Example 4.38. For any field $\mathbb{F}$, the power series ring $\mathfrak{o} = \mathbb{F}[[t]]$ is a DVR with maximal ideal $\mathfrak{m} = (t)$ and residue field $k = \mathfrak{o}/\mathfrak{m} = \mathbb{F}[[t]]/(t) \cong \mathbb{F}$. An $\mathfrak{o}$-module L of type λ is an $\mathbb{F}$-vector space together with an $\mathbb{F}$-linear transformation T that acts on M nilpotently (so that $g := T + 1$ acts unipotently) with Jordan blocks of sizes given by λ: each summand $\mathfrak{o}/\mathfrak{m}^{\lambda_i} = \mathbb{F}[[t]]/(t^{\lambda_i})$ of L has an $\mathbb{F}$-basis $\{1, t, t^2, \ldots, t^{\lambda_i - 1}\}$ on which the map T that multiplies by t acts as a nilpotent Jordan block of size λ_i. Note also that, in this setting, $\mathfrak{o}$-submodules are the same as T-stable (or g-stable) $\mathbb{F}$-subspaces.

[18] See also [63, Thm. 2.6, Prop. 2.7] for quick proofs of part of it, similar to Zelevinsky's.

Example 4.39. The ring of p-adic integers $\mathfrak{o} = \mathbb{Z}_p$ is a DVR with maximal ideal $\mathfrak{m} = (p)$ and residue field $k = \mathfrak{o}/\mathfrak{m} = \mathbb{Z}_p/p\mathbb{Z}_p \cong \mathbb{Z}/p\mathbb{Z}$. An $\mathfrak{o}$-module L of type λ is an abelian p-group of type λ: for each summand, $\mathfrak{o}/\mathfrak{m}^{\lambda_i} = \mathbb{Z}_p/p^{\lambda_i}\mathbb{Z}_p \cong \mathbb{Z}/p^{\lambda_i}\mathbb{Z}$. Note also that, in this setting, $\mathfrak{o}$-submodules are the same as subgroups.

One last notation: $n(\lambda) := \sum_{i \geq 1} (i-1)\lambda_i$, for λ in Par. Hall's Theorem is as follows.

Theorem 4.40. *Assume $\mathfrak{o}$ is a DVR with maximal ideal $\mathfrak{m}$, and that its residue field $k = \mathfrak{o}/\mathfrak{m}$ is finite of cardinality q. Fix an $\mathfrak{o}$-module L of type λ. Then the number of $\mathfrak{o}$-submodules M of type μ for which the quotient $N = L/M$ is of type ν can be written as the specialization*

$$[g^{\lambda}_{\mu,\nu}(t)]_{t=q}$$

of a polynomial $g^{\lambda}_{\mu,\nu}(t)$ in $\mathbb{Z}[t]$, called the **Hall polynomial**.

Furthermore, the Hall polynomial $g^{\lambda}_{\mu,\nu}(t)$ has degree at most $n(\lambda) - (n(\mu) + n(\nu))$, and its coefficient of $t^{n(\lambda)-(n(\mu)+n(\nu))}$ is the Littlewood-Richardson coefficient $c^{\lambda}_{\mu,\nu}$.

Comparing what Hall's Theorem says in Examples 4.38 and 4.39, shows that the map (4.36) gives the desired isomorphism from the classical Hall algebra to $\mathcal{H}_{\mathbb{Z}}$.

We close this section with some remarks on the vast literature on Hall algebras that we will *not* discuss here.

Remark 4.41. Macdonald's version of Hall's Theorem [49, (4.3)] is stronger than Theorem 4.40, and useful for certain applications: he shows that $g^{\lambda}_{\mu,\nu}(t)$ is the zero polynomial whenever the Littlewood-Richardson coefficient $c^{\lambda}_{\mu,\nu}$ is zero.

Remark 4.42. Zelevinsky in [81, Chaps 10, 11] uses the isomorphism $\Lambda_{\mathbb{C}} \to \mathcal{H}$ to derive J. Green's formula for the value of any irreducible character χ of GL_n on any unipotent class J_λ. The answer involves values of irreducible characters of $\mathfrak{S}_n$ along with *Green's polynomials* $Q^{\lambda}_{\mu}(q)$ (see Macdonald [49, §III.7]; they are denoted $Q(\lambda, \mu)$ by Zelevinsky), which express the images under the isomorphism of Theorem 4.35 of the symmetric function basis $\{p_\mu\}$ in terms of the basis $\{1_{J_\lambda}\}$.

Remark 4.43. The Hall polynomials $g^{\lambda}_{\mu,\nu}(t)$ also essentially give the multiplicative structure constants for $\Lambda(\mathbf{x})[t]$ with respect to its basis of *Hall-Littlewood symmetric functions* $P_\lambda = P_\lambda(\mathbf{x}; t)$:

$$P_\mu P_\nu = \sum_{\lambda} t^{n(\lambda)-(n(\mu)+n(\nu))} g^{\lambda}_{\mu,\nu}(t^{-1}) P_\lambda.$$

See Macdonald [49, §III.3].

Remark 4.44. Schiffmann [63] discusses self-dual Hopf algebras which vastly generalize the classical Hall algebra called *Ringel-Hall algebras*, associated to abelian categories which are hereditary. Examples come from categories of nilpotent representations of quivers; the quiver having exactly one node and one arc recovers the classical Hall algebra $\mathcal{H}_{\mathbb{Z}}$ discussed above.

Remark 4.45. The general linear groups $GL_n(\mathbb{F}_q)$ are one of four families of so-called *classical groups*. Progress has been made on extending Zelevinsky's PSH theory to the other families:

(a) Work of Thiem and Vinroot [80] shows that the tower $\{G_*\}$ of *finite unitary groups* $U_n(\mathbb{F}_{q^2})$ give rise to another positive self-dual Hopf algebra $A = \bigoplus_{n \geq 0} R(U_n(\mathbb{F}_{q^2}))$, in which the role of Harish-Chandra induction is played by *Deligne-Lusztig induction*. In this theory, character and degree formulas for $U_n(\mathbb{F}_{q^2})$ are related to those of $GL_n(\mathbb{F}_q)$ by substituting $q \mapsto -q$, along with appropriate scalings by ± 1, a phenomenon sometimes called *Ennola duality*. See also [73, §4].

(b) van Leeuwen [44] has studied $\bigoplus_{n \geq 0} R\left(Sp_{2n}\left(\mathbb{F}_q\right)\right)$, $\bigoplus_{n \geq 0} R\left(O_{2n}\left(\mathbb{F}_q\right)\right)$ and $\bigoplus_{n \geq 0} R\left(U_n\left(\mathbb{F}_{q^2}\right)\right)$ not as Hopf algebras, but rather as so-called *twisted PSH-modules* over the PSH-algebra $A(GL)$ (a "deformed" version of the older notion of Hopf modules). He classified these PSH-modules axiomatically similarly to Zelevinsky's above classification of PSH-algebras.

(c) In a recent honors thesis [68], Shelley-Abrahamson defined yet another variation of the concept of Hopf modules, named 2-*compatible Hopf modules*, and identified $\bigoplus_{n \geq 0} R\left(Sp_{2n}\left(\mathbb{F}_q\right)\right)$ and $\bigoplus_{n \geq 0} R\left(O_{2n+1}\left(\mathbb{F}_q\right)\right)$ as such modules over $A(GL)$.

5. Quasisymmetric functions and P-partitions

We discuss here our next important example of a Hopf algebra arising in combinatorics: the *quasisymmetric functions* of Gessel [27], with roots in work of Stanley [69] on P-partitions.

5.1. Definitions, and Hopf structure.

The definitions of quasisymmetric functions require a totally ordered variable set. Usually we will use a variable set denoted $\mathbf{x} = (x_1, x_2, \ldots)$ with the usual ordering $x_1 < x_2 < \ldots$. However, it is good to have some flexibility in changing the ordering, which is why we make the following definition.

Definition 5.1. Given any totally ordered set I, create a totally ordered variable set $\{x_i\}_{i \in I}$, and then let $R(\{x_i\}_{i\in I})$ denote the power series of bounded degree in $\{x_i\}_{i\in I}$ having coefficients in $\mathbf{k}$.

The *quasisymmetric functions* $\mathrm{QSym} := \mathrm{QSym}(\{x_i\}_{i\in I})$ will be the $\mathbf{k}$-submodule consisting of the elements f in $R(\{x_i\}_{i\in I})$ that have the same coefficient on the monomials $x_{i_1}^{\alpha_1} \cdots x_{i_\ell}^{\alpha_\ell}$ and $x_{j_1}^{\alpha_1} \cdots x_{j_\ell}^{\alpha_\ell}$ whenever both $i_1 < \cdots < i_\ell$ and $j_1 < \cdots < j_\ell$ in the total order on I. We write $\mathrm{QSym}_{\mathbf{k}}$ instead of QSym to stress the choice of base ring $\mathbf{k}$.

It immediately follows from this definition that $\mathrm{QSym}(\{x_i\}_{i\in I})$ forms a free $\mathbf{k}$-submodule of $R(\{x_i\}_{i\in I})$, having as $\mathbf{k}$-basis elements the *monomial quasisymmetric functions*

$$M_\alpha(\{x_i\}_{i\in I}) := \sum_{i_1 < \cdots < i_\ell \text{ in } I} x_{i_1}^{\alpha_1} \cdots x_{i_\ell}^{\alpha_\ell}$$

for all compositions α satisfying $\ell(\alpha) \leq |I|$. When I is infinite, this means that the M_α for all compositions α form a basis of $\mathrm{QSym}(\{x_i\}_{i\in I})$. Note that $\mathrm{QSym} = \bigoplus_{n\geq 0} \mathrm{QSym}_n$ is a graded $\mathbf{k}$-module of finite type, where QSym_n is the subspace of quasisymmetric functions which are homogeneous of degree n. Letting Comp denote the set of all compositions α, and Comp_n the compositions α of n (that is, compositions whose parts sum to n), the subset $\{M_\alpha\}_{\alpha \in \mathrm{Comp}_n}$ gives a $\mathbf{k}$-basis for QSym_n.

Example 5.2. Taking the variable set $\mathbf{x} = (x_1 < x_2 < \cdots)$ to define $\mathrm{QSym} = \mathrm{QSym}(\mathbf{x})$, for $n = 0, 1, 2, 3$, one has these basis elements in QSym_n:

$$\begin{aligned}
M_{()} = M_\varnothing &= 1 \\
\\
M_{(1)} &= x_1 + x_2 + x_3 + \cdots && = m_{(1)} = s_{(1)} = e_1 = h_1 = p_1 \\
\\
M_{(2)} &= x_1^2 + x_2^2 + x_3^2 + \cdots && = m_{(2)} = p_2 \\
M_{(1,1)} &= x_1x_2 + x_1x_3 + x_2x_3 + \cdots && = m_{(1,1)} = e_2 \\
\\
M_{(3)} &= x_1^3 + x_2^3 + x_3^3 + \cdots && = m_{(3)} = p_3 \\
M_{(2,1)} &= x_1^2x_2 + x_1^2x_3 + x_2^2x_3 + \cdots \\
M_{(1,2)} &= x_1x_2^2 + x_1x_3^2 + x_2x_3^2 + \cdots \\
M_{(1,1,1)} &= x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + \cdots && = m_{(1,1,1)} = e_3
\end{aligned}$$

It is not obvious that QSym is a subalgebra of $R(\mathbf{x})$, but we will show this momentarily. For example,

$$\begin{aligned}
M_{(a)}M_{(b,c)} &= (x_1^a + x_2^a + x_3^a + \cdots)(x_1^bx_2^c + x_1^bx_3^c + x_2^bx_3^c + \cdots) \\
&= x_1^{a+b}x_2^c + \cdots + x_1^bx_3^{a+c} + \cdots + x_1^ax_2^bx_3^c + \cdots + x_1^bx_2^ax_3^c + \cdots + x_1^bx_2^cx_3^a + \cdots \\
&= M_{(a+b,c)} + M_{(b,a+c)} + M_{(a,b,c)} + M_{(b,a,c)} + M_{(b,c,a)}
\end{aligned}$$

Proposition 5.3. *For any infinite totally ordered set I, one has that $\mathrm{QSym} = \mathrm{QSym}(\{x_i\}_{i\in I})$ is a $\mathbf{k}$-subalgebra of $R(\{x_i\}_{i\in I})$, with multiplication in the $\{M_\alpha\}$-basis as follows: Fix three pairwise disjoint chain posets $(i_1 < \cdots < i_\ell)$, $(j_1 < \cdots < j_m)$ and $(k_1 < k_2 < \cdots)$. Now, if $\alpha = (\alpha_1, \ldots, \alpha_\ell), \beta = (\beta_1, \ldots, \beta_m)$ then*

$$M_\alpha M_\beta = \sum_f M_{\mathrm{wt}\, f} \tag{5.1}$$

in which the sum is over all $p \in \mathbb{N}$ and all maps f from the disjoint union of two chains to a chain

$$(i_1 < \cdots < i_\ell) \sqcup (j_1 < \cdots < j_m) \xrightarrow{f} (k_1 < \cdots < k_p) \tag{5.2}$$

which are both surjective and strictly order-preserving ($x < y$ implies $f(x) < f(y)$), and where the composition $\mathrm{wt}(f) := (\mathrm{wt}_1(f), \ldots, \mathrm{wt}_p(f))$ *is defined by* $\mathrm{wt}_s(f) := \sum_{i_u \in f^{-1}(k_s)} \alpha_u + \sum_{j_v \in f^{-1}(k_s)} \beta_v$.

In particular, all such algebras are isomorphic to a single algebra QSym*, defined as having* $\mathbf{k}$*-basis* $\{M_\alpha\}_{\alpha \in \mathrm{Comp}}$ *and with multiplication defined* $\mathbf{k}$*-linearly by* (5.1)*. The isomorphism sends* $M_\alpha \longmapsto M_\alpha(\{x_i\}_{i \in I})$*.*

Proof. Formula (5.1) comes from considering a typical product of two monomials in the expansion of $M_\alpha M_\beta$:

$$\left(x_{i_1}^{\alpha_1} \cdots x_{i_\ell}^{\alpha_\ell}\right)\left(x_{j_1}^{\beta_1} \cdots x_{j_m}^{\beta_m}\right) = x_{k_1}^{\gamma_1} \cdots x_{k_p}^{\gamma_p}$$

for subscript sequences $i_1 < \cdots < i_\ell$ and $j_1 < \cdots < j_m$ and $k_1 < \cdots < k_p$ with

$$\{k_1, \ldots, k_p\} = \{i_1, \ldots, i_\ell\} \cup \{j_1, \ldots, j_m\}.$$

Thinking of $\{i_s\}$ and $\{j_t\}$ as disjoint sets, multiplication gives a surjective map f as in (5.2), with $\gamma_s = \mathrm{wt}_s(f)$.

Once one has this multiplication rule (5.1), the last assertion follows. $\square$

The comultiplication of QSym will extend the one that we defined for Λ. That is, one considers the linear order from (2.7) on two sets of variables $(\mathbf{x}, \mathbf{y}) = (x_1 < x_2 < \ldots < y_1 < y_2 < \ldots)$, and notes that

$$\mathrm{QSym}(\mathbf{x}, \mathbf{y}) \subset \mathrm{QSym}(\mathbf{x}) \otimes \mathrm{QSym}(\mathbf{y})$$

so that one can define $\mathrm{QSym} \xrightarrow{\Delta} \mathrm{QSym} \otimes \mathrm{QSym}$ as the composite of the maps in the bottom row here:

$$\begin{array}{ccccccc} & & R(\mathbf{x},\mathbf{y}) & \cong & R(\mathbf{x}) \otimes R(\mathbf{y}) & & \\ & & \cup & & \cup & & \\ \mathrm{QSym} & \cong & \mathrm{QSym}(\mathbf{x},\mathbf{y}) & \hookrightarrow & \mathrm{QSym}(\mathbf{x}) \otimes \mathrm{QSym}(\mathbf{y}) & \cong & \mathrm{QSym} \otimes \mathrm{QSym} \\ f & \longmapsto & f(\mathbf{x},\mathbf{y}) = f(x_1, x_2, \ldots, y_1, y_2, \ldots) & & & & \end{array} \tag{5.3}$$

Here, $f(\mathbf{x}, \mathbf{y})$ is formally defined as the image of f under the algebra isomorphism $\mathrm{QSym} \to \mathrm{QSym}(\mathbf{x}, \mathbf{y})$ defined in Proposition 5.3.

Example 5.4. For example,

$$\begin{aligned} \Delta M_{(a,b,c)} &= M_{(a,b,c)}(x_1, x_2, \ldots, y_1, y_2, \ldots) \\ &= x_1^a x_2^b x_3^c + x_1^a x_2^b x_4^c + \cdots \\ &\quad + x_1^a x_2^b \cdot y_1^c + x_1^a x_2^b \cdot y_2^c + \cdots \\ &\quad + x_1^a \cdot y_1^b y_2^c + x_1^a \cdot y_1^b y_3^c + \cdots \\ &\quad + y_1^a y_2^b y_3^c + y_1^a y_2^b y_4^c + \cdots \\ &= M_{(a,b,c)}(\mathbf{x}) + M_{(a,b)}(\mathbf{x}) M_{(c)}(\mathbf{y}) + M_{(a)}(\mathbf{x}) M_{(b,c)}(\mathbf{y}) + M_{(a,b,c)}(\mathbf{y}) \\ &= M_{(a,b,c)} \otimes 1 + M_{(a,b)} \otimes M_{(c)} + M_{(a)} \otimes M_{(b,c)} + 1 \otimes M_{(a,b,c)} \end{aligned}$$

Defining the *concatenation* $\beta \cdot \gamma$ of two compositions $\beta = (\beta_1, \ldots, \beta_r), \gamma = (\gamma_1, \ldots, \gamma_s)$ to be the composition $(\beta_1, \ldots, \beta_r, \gamma_1, \ldots, \gamma_s)$, one has the following description of the coproduct in the $\{M_\alpha\}$ basis.

Proposition 5.5. *For a composition* $\alpha = (\alpha_1, \ldots, \alpha_\ell)$*, one has*

$$\Delta M_\alpha = \sum_{k=0}^{\ell} M_{(\alpha_1, \ldots, \alpha_k)} \otimes M_{(\alpha_{k+1}, \ldots, \alpha_\ell)} = \sum_{\substack{(\beta,\gamma): \\ \beta \cdot \gamma = \alpha}} M_\beta \otimes M_\gamma$$

Proof. This comes from expressing a monomial in $\Delta M_\alpha = M_\alpha(\mathbf{x}, \mathbf{y})$ uniquely in the form $x_{i_1}^{\alpha_1} \cdots x_{i_k}^{\alpha_k} \cdot y_{j_1}^{\alpha_{k+1}} \cdots y_{j_{\ell-k}}^{\alpha_\ell}$ for some $k \in \{0, 1, \ldots, n\}$ and some subscripts $i_1 < \cdots < i_k$ and $j_1 < \cdots < j_{\ell-k}$. $\square$

Proposition 5.6. *The quasisymmetric functions* QSym *form a graded connected Hopf algebra of finite type, which is commutative, and contains the symmetric functions* Λ *as a Hopf subalgebra.*

Proof. To prove coassociativity of Δ, we need to be slightly careful. It seems reasonable to argue by $(\Delta \otimes 1) \circ \Delta f = f(\mathbf{x}, \mathbf{y}, \mathbf{z}) = (1 \otimes \Delta) \circ \Delta f$ as in the case of Λ, but this would now require further justification, as terms like $f(\mathbf{x}, \mathbf{y})$ and $f(\mathbf{x}, \mathbf{y}, \mathbf{z})$ are no longer directly defined as evaluations of f on some sequences (but rather are defined as images of f under certain homomorphisms). However, it is very easy to see that Δ is coassociative by checking $(\Delta \otimes 1) \circ \Delta = (1 \otimes \Delta) \circ \Delta$ on the $\{M_\alpha\}$ basis: Proposition 5.5 yields

$$\begin{aligned}
((\Delta \otimes 1) \circ \Delta) M_\alpha &= \sum_{k=0}^{\ell} \Delta(M_{(\alpha_1,\ldots,\alpha_k)}) \otimes M_{(\alpha_{k+1},\ldots,\alpha_\ell)} \\
&= \sum_{k=0}^{\ell} \left(\sum_{i=0}^{k} M_{(\alpha_1,\ldots,\alpha_i)} \otimes M_{(\alpha_{i+1},\ldots,\alpha_k)} \right) \otimes M_{(\alpha_{k+1},\ldots,\alpha_\ell)} \\
&= \sum_{k=0}^{\ell} \sum_{i=0}^{k} M_{(\alpha_1,\ldots,\alpha_i)} \otimes M_{(\alpha_{i+1},\ldots,\alpha_k)} \otimes M_{(\alpha_{k+1},\ldots,\alpha_\ell)}
\end{aligned}$$

and the same expression for $((1 \otimes \Delta) \circ \Delta) M_\alpha$.

The coproduct Δ of QSym is an algebra morphism because it is defined as a composite of algebra morphisms in the bottom row of (5.3). To prove that the restriction of Δ to the subring Λ of QSym is the comultiplication of QSym, it thus is enough to check that it sends the elementary symmetric function e_n to $\sum_{i=0}^{n} e_i \otimes e_{n-i}$ for every $n \in \mathbb{N}$. This again follows from Proposition 5.5, since $e_n = M_{(1,1,\ldots,1)}$ (with n times 1).

The counit is as usual for a graded connected coalgebra, and just as in the case of Λ, sends a quasisymmetric function $f(\mathbf{x})$ to its constant term $f(0,0,\ldots)$. This is an evaluation, and hence an algebra map. Hence QSym forms a bialgebra, and as it is graded, connected, and of finite type, also a Hopf algebra by Proposition 1.30. $\square$

We will identify the antipode in QSym shortly, but we first deal with another slightly subtle issue. In addition to the counit evaluation $\epsilon(f) = f(0,0,\ldots)$, starting in Section 6.1, we will want to specialize elements in QSym($\mathbf{x}$) by making other variable substitutions, in which all but a finite list of variables are set to zero. We justify this here.

Proposition 5.7. *Fix a totally ordered set I, a $\mathbf{k}$-algebra A, a finite list of variables $x_{i_1}, \ldots, x_{i_m}$, say with $i_1 < \ldots < i_m$ in I, and an ordered list of elements $(a_1, \ldots, a_m) \in A^m$.*

Then there is a well-defined evaluation homomorphism

$$\begin{array}{rcl}
\mathrm{QSym}(\{x_i\}_{i \in I}) & \longrightarrow & A \\
f & \longmapsto & [f]_{\substack{x_{i_1}=a_1,\ldots,x_{i_m}=a_m \\ x_j=0 \text{ for } j \notin \{i_1,\ldots,i_m\}}}
\end{array}.$$

Furthermore, this homomorphism depends only upon the list $(a_1, \ldots, a_m)$, as it coincides with the following:

$$\begin{array}{rcl}
\mathrm{QSym}(\{x_i\}_{i \in I}) \cong \mathrm{QSym}(x_1, x_2, \ldots) & \longrightarrow & A \\
f(x_1, x_2, \ldots) & \longmapsto & f(a_1, \ldots, a_m, 0, 0 \ldots).
\end{array}$$

(This latter statement is stated for the case when I is infinite; otherwise, read "$x_1, x_2, \ldots, x_{|I|}$" for "$x_1, x_2, \ldots$", and interpret $(a_1, \ldots, a_m, 0, 0 \ldots)$ as an $|I|$-tuple.)

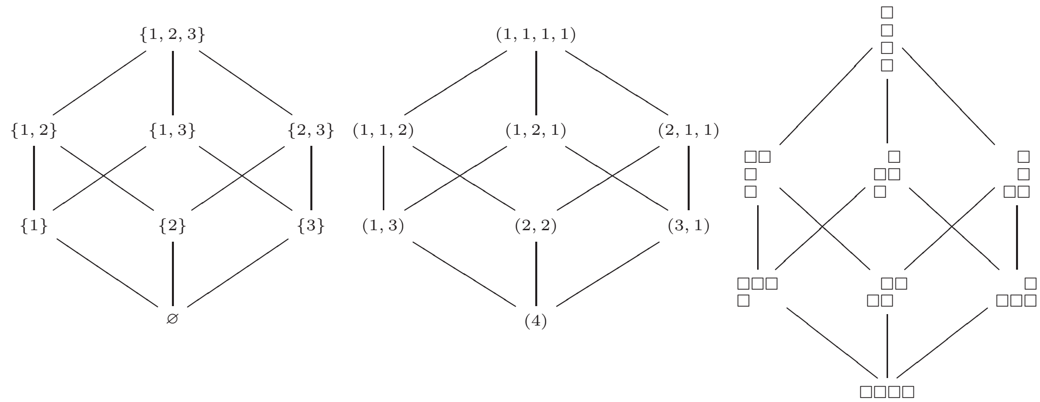
Proof. One already can make sense of evaluating $x_{i_1} = a_1, \ldots, x_{i_m} = a_m$ and $x_j = 0$ for $j \notin \{i_1, \ldots, i_m\}$ in the ambient ring $R(\{x_i\}_{i \in I})$ containing $\mathrm{QSym}(\{x_i\}_{i \in I})$, since a power series f of bounded degree will have finitely many monomials that only involve the variables $x_{i_1}, \ldots, x_{i_m}$. The last assertion follows from quasisymmetry of f, and is perhaps checked most easily when $f = M_\alpha(\{x_i\}_{i \in I})$ for some α. $\square$

The antipode in QSym has a reasonably simple expression in the $\{M_\alpha\}$ basis, but requiring a definition.

Definition 5.8. For α, β in Comp_n, say that α *refines* β or β *coarsens* α if, informally, one can obtain β from α by combining some of its adjacent parts. Alternatively, one has a bijection $\mathrm{Comp}_n \to 2^{[n-1]}$ where $[n-1] := \{1, 2, \ldots, n-1\}$ which sends $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ having length $\ell(\alpha) = \ell$ to its subset of partial sums

$$D(\alpha) := \{\alpha_1, \alpha_1 + \alpha_2, \ldots, \alpha_1 + \cdots + \alpha_{\ell-1}\},$$

and this sends the refinement ordering to the inclusion ordering on the Boolean algebra $2^{[n-1]}$ (to be more precise: a composition α refines a composition β if and only if $D(\alpha) \supset D(\beta)$). There is also a bijection sending α to its *ribbon* diagram: the skew diagram λ/μ having rows of sizes $(\alpha_1, \ldots, \alpha_\ell)$ read from bottom to top with exactly one column of overlap between adjacent rows. These bijections and the refinement partial order are illustrated here for $n = 4$:



Given $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, its *reverse* composition is $\mathrm{rev}(\alpha) = (\alpha_\ell, \alpha_{\ell-1}, \ldots, \alpha_2, \alpha_1)$. Note that $\alpha \mapsto \mathrm{rev}(\alpha)$ is a poset automorphism for the refinement ordering.

Theorem 5.9. *For any composition α in* Comp*,*

$$S(M_\alpha) = (-1)^{\ell(\alpha)} \sum_{\substack{\gamma \in \mathrm{Comp}: \\ \gamma \text{ coarsens } \mathrm{rev}(\alpha)}} M_\gamma$$

For example,

$$S(M_{(a,b,c)}) = -\left(M_{(c,b,a)} + M_{(b+c,a)} + M_{(c,a+b)} + M_{(a+b+c)}\right)$$

Proof. We give Ehrenborg's proof[19] [24, Prop. 3.4] via induction on $\ell = \ell(\alpha)$. One has easy base cases when $\ell(\alpha) = 0$, where $S(M_\varnothing) = S(1) = 1 = (-1)^0 M_{\mathrm{rev}(\varnothing)}$, and when $\ell(\alpha) = 1$, where $M_{(n)}$ is primitive by Proposition 5.5, so Proposition 1.31 shows $S(M_{(n)}) = -M_{(n)} = (-1)^1 M_{\mathrm{rev}((n))}$.

For the inductive step, apply the inductive definition of S from the proof of Proposition 1.30:

$$\begin{aligned} S(M_{(\alpha_1,\ldots,\alpha_\ell)}) &= -\sum_{i=0}^{\ell-1} S(M_{(\alpha_1,\ldots,\alpha_i)}) M_{(\alpha_{i+1},\ldots,\alpha_\ell)} \\ &= \sum_{i=0}^{\ell-1} \sum_{\substack{\beta \text{ coarsening} \\ (\alpha_i,\alpha_{i-1},\ldots,\alpha_1)}} (-1)^{i+1} M_\beta M_{(\alpha_{i+1},\ldots,\alpha_\ell)} \end{aligned}$$

The idea will be to cancel terms of opposite sign that appear in the expansions of the products $M_\beta M_{(\alpha_{i+1},\ldots,\alpha_\ell)}$. Note that each composition β appearing above has first part β_1 of the form $\alpha_i + \alpha_{i-1} + \cdots + \alpha_h$ for some $h \leq i$ (unless $\beta = \varnothing$), and hence each term M_γ in the expansion of the product $M_\beta M_{(\alpha_{i+1},\ldots,\alpha_\ell)}$ has γ_1 (that is, the first entry of γ) a sum that can take one of these three forms:

- $\alpha_i + \alpha_{i-1} + \cdots + \alpha_h$,
- $\alpha_{i+1} + (\alpha_i + \alpha_{i-1} + \cdots + \alpha_h)$,
- α_{i+1}.

Say that the *type* of γ is i in the first case, and $i+1$ in the second two cases[20]; in other words, the type is the largest subscript k on a part α_k which was combined in the sum γ_1. It is not hard to see that a given γ for which the type k is strictly smaller than ℓ arises from exactly two pairs $(\beta, \gamma), (\beta', \gamma)$, having opposite signs $(-1)^k$ and $(-1)^{k+1}$ in the above sum[21]. For example, if $\alpha = (\alpha_1, \ldots, \alpha_8)$, then the composition

[19] A different proof was given by Malvenuto and Reutenauer [51, Cor. 2.3], and is sketched in Remark 5.31 below.

[20] We imagine that we label the terms obtained by expanding $M_\beta M_{(\alpha_{i+1},\ldots,\alpha_\ell)}$ by distinct labels, so that each term knows how exactly it was created (i.e., which i, which β and which map f as in (5.2) gave rise to it). Strictly speaking, it is these triples (i, β, f) that we should be assigning types to, not terms.

[21] Strictly speaking, this means that we have an involution on the set of our (i, β, f) triples having type smaller than ℓ, and this involution switches the sign of $(-1)^i M_{\mathrm{wt}\, f}$.

$\gamma = (\alpha_6 + \alpha_5 + \alpha_4, \alpha_3, \alpha_7, \alpha_8 + \alpha_2 + \alpha_1)$ of type 6 can arise from either of

$$\beta = (\alpha_6 + \alpha_5 + \alpha_4, \alpha_3, \alpha_2 + \alpha_1) \text{ with } i = 6 \text{ and sign } (-1)^7$$
$$\beta' = (\alpha_5 + \alpha_4, \alpha_3, \alpha_2 + \alpha_1) \text{ with } i = 5 \text{ and sign } (-1)^6.$$

Similarly, $\gamma = (\alpha_6, \alpha_5 + \alpha_4, \alpha_3, \alpha_7, \alpha_8 + \alpha_2 + \alpha_1)$ can arise from either of

$$\beta = (\alpha_6, \alpha_5 + \alpha_4, \alpha_3, \alpha_2 + \alpha_1) \text{ with } i = 6 \text{ and sign } (-1)^7$$
$$\beta' = (\alpha_5 + \alpha_4, \alpha_3, \alpha_2 + \alpha_1) \text{ with } i = 5 \text{ and sign } (-1)^6.$$

Thus one can cancel almost all the terms, excepting those with γ of type ℓ among the terms M_γ in the expansion of the last $(i = \ell - 1)$ summand $M_\beta M_{(\alpha_\ell)}$. A bit of thought shows that these are the γ coarsening $\mathrm{rev}(\alpha)$, and all have sign $(-1)^\ell$. $\square$

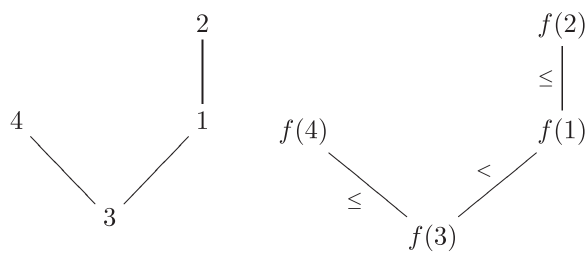
5.2. **The fundamental basis and P-partitions.** There is a second important basis for QSym which arose originally in Stanley's P-partition theory [69].

Definition 5.10. A *labelled poset* will here mean a partially ordered set P whose underlying set is some finite subset of the integers. A *P-partition* is a function $P \xrightarrow{f} \{1, 2, \ldots\}$ with the property that

- $i <_P j$ and $i <_{\mathbb{Z}} j$ implies $f(i) \leq f(j)$, and
- $i <_P j$ and $i >_{\mathbb{Z}} j$ implies $f(i) < f(j)$.

Denote by $\mathcal{A}(P)$ the set of all P-partitions f, and let $F_P(\mathbf{x}) := \sum_{f \in \mathcal{A}(P)} \mathbf{x}_f$ where $\mathbf{x}_f := \prod_{i \in P} x_{f(i)}$

Example 5.11. Depicted is a labelled poset P, along with the relations among the four values $f = (f(1), f(2), f(3), f(4))$ that define its P-partitions f:



The following is an important special case.

Proposition 5.12. *When P is a total or linear order $w = (w_1 < \ldots < w_n)$, the generating function $F_w(\mathbf{x})$ depends only upon the descent set*

$$\mathrm{Des}(w) := \{i : w_i >_{\mathbb{Z}} w_{i+1}\} \subset \{1, 2, \ldots, n-1\}$$

and its associated composition α in Comp_n having partial sums $D(\alpha) = \mathrm{Des}(w)$: one has that $F_w(\mathbf{x})$ equals the **fundamental quasisymmetric function**

$$L_\alpha := L_\alpha(\mathbf{x}) := \sum_{\substack{(1\leq) i_1 \leq \cdots \leq i_n: \\ i_j < i_{j+1} \text{ if } j \in D(\alpha)}} x_{i_1} \cdots x_{i_n} = \sum_{\substack{\beta \in \mathrm{Comp}_n: \\ \beta \text{ refines } \alpha}} M_\beta. \tag{5.4}$$

E.g., total order $w = 35142$ has $\mathrm{Des}(w) = \{2, 4\}$ and composition $\alpha = (2, 2, 1)$, so

$$\begin{aligned} F_{35142}(\mathbf{x}) &= \sum_{f(3) \leq f(5) < f(1) \leq f(4) < f(2)} x_{f(3)} x_{f(5)} x_{f(1)} x_{f(4)} x_{f(2)} \\ &= \sum_{i_1 \leq i_2 < i_3 \leq i_4 < i_5} x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} \\ &= L_{(2,2,1)} = M_{(2,2,1)} + M_{(2,1,1,1)} + M_{(1,1,2,1)} + M_{(1,1,1,1,1)}. \end{aligned}$$

Proof. Write $F_w(\mathbf{x})$ as a sum of monomials $x_{f(w_1)} \cdots x_{f(w_n)}$ over all sequences $f(w_1) \leq \cdots \leq f(w_n)$ having strict inequalities $f(w_i) < f(w_{i+1})$ whenever i is in $\mathrm{Des}(w)$. The underlying set $\{f(w_i)\}_{i=1}^n$ will equal $\{j_1 < \ldots < j_\ell\}$ with indices in increasing order having a multiplicity sequence $\beta = (\beta_1, \ldots, \beta_\ell)$ that gives a composition β refining α. $\square$

Example 5.13. The extreme cases for α in Comp_n give quasisymmetric functions L_α which are symmetric:

$$L_{(1^n)} = M_{(1^n)} = e_n,$$
$$L_{(n)} = \sum_{\alpha \in \mathrm{Comp}_n} M_\alpha = h_n$$

One sees that the $\{L_\alpha\}_{\alpha \in \mathrm{Comp}}$ are a $\mathbb{Z}$-basis for QSym, as inclusion-exclusion applied to (5.4) gives

$$M_\alpha = \sum_{\substack{\beta \in \mathrm{Comp}_n: \\ \beta \text{ refines } \alpha}} (-1)^{\ell(\beta)-\ell(\alpha)} L_\beta.$$

We need to make a technical observation, which will be used later.

Lemma 5.14. *Let $n \in \mathbb{N}$. Let α be a composition of n. Let I be an infinite totally ordered set. Then,*

$$L_\alpha\left(\{x_i\}_{i\in I}\right) = \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if } j \in D(\alpha)}} x_{i_1} x_{i_2} \cdots x_{i_n},$$

where $L_\alpha\left(\{x_i\}_{i\in I}\right)$ is defined as the image of L_α under the isomorphism $\mathrm{QSym} \to \mathrm{QSym}\left(\{x_i\}_{i\in I}\right)$ obtained in Proposition 5.3.

Proof. We cannot directly obtain the lemma by "evaluating" the sides of (5.4) at $\{x_i\}_{i\in I}$. However, we can notice that every composition $\beta = (\beta_1, \ldots, \beta_\ell)$ of n satisfies

$$M_\beta\left(\{x_i\}_{i\in I}\right) = \sum_{k_1 < \cdots < k_\ell \text{ in } I} x_{k_1}^{\beta_1} \cdots x_{k_\ell}^{\beta_\ell} = \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if and only if } j \in D(\beta)}} x_{i_1} x_{i_2} \cdots x_{i_n}. \tag{5.5}$$

Applying the ring homomorphism $\mathrm{QSym} \to \mathrm{QSym}\left(\{x_i\}_{i\in I}\right)$ to (5.4), we obtain

$$\begin{aligned} L_\alpha\left(\{x_i\}_{i\in I}\right) &= \sum_{\substack{\beta \in \mathrm{Comp}_n: \\ \beta \text{ refines } \alpha}} M_\beta\left(\{x_i\}_{i\in I}\right) \overset{(5.5)}{=} \sum_{\substack{\beta \in \mathrm{Comp}_n: \\ \beta \text{ refines } \alpha}} \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if and only if } j \in D(\beta)}} x_{i_1} x_{i_2} \cdots x_{i_n} \\ &= \sum_{\substack{\beta \in \mathrm{Comp}_n: \\ D(\alpha) \subset D(\beta)}} \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if and only if } j \in D(\beta)}} x_{i_1} x_{i_2} \cdots x_{i_n} \\ &= \sum_{\substack{Z \subset [n-1]: \\ D(\alpha) \subset Z}} \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if and only if } j \in Z}} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if } j \in D(\alpha)}} x_{i_1} x_{i_2} \cdots x_{i_n}. \end{aligned}$$

$\square$

The next proposition ([72, Cor. 7.19.5], [48, Cor. 3.3.24]) is an algebraic shadow of Stanley's main lemma [72, Thm. 7.19.4] in P-partition theory. It expands any $F_P(\mathbf{x})$ in the $\{L_\alpha\}$ basis, as a sum over the set $\mathcal{L}(P)$ of all *linear extensions* w of P, that is, the set of all extensions of P to a linear order. E.g., the poset P from Example 5.11 has $\mathcal{L}(P) = \{3124, 3142, 3412\}$.

Theorem 5.15. *For any labelled poset P,*

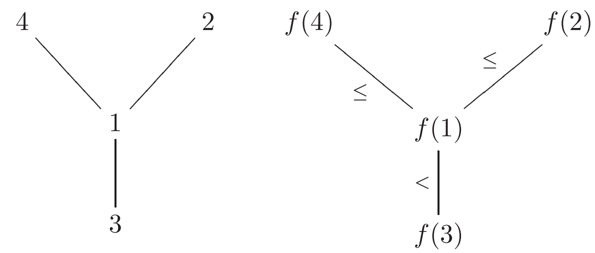
$$F_P(\mathbf{x}) = \sum_{w \in \mathcal{L}(P)} F_w(\mathbf{x}).$$

Proof. We give Gessel's proof [27, Thm. 1], via induction on the number of pairs i, j which are incomparable in P. When this quantity is 0, then P is itself a linear order w, so that $\mathcal{L}(P) = \{w\}$ and there is nothing to prove.

In the inductive step, let i, j be incomparable elements. Consider the two posets $P_{i<j}$ and $P_{j<i}$ which are obtained from P by adding in an order relation between i and j, and then taking the transitive closure; it is not hard to see that these transitive closures cannot contain a cycle, so that these really do define two posets. The result then follows by induction applied to $P_{i<j}, P_{j<i}$, once one notices that $\mathcal{L}(P) = \mathcal{L}(P_{i<j}) \sqcup \mathcal{L}(P_{j<i})$ since every linear extension w of P either has i before j or vice-versa, and $\mathcal{A}(P) = \mathcal{A}(P_{i<j}) \sqcup \mathcal{A}(P_{j<i})$ since, assuming that $i <_{\mathbb{Z}} j$ without loss of generality, every f in $\mathcal{A}(P)$ either satisfies $f(i) \leq f(j)$ or $f(i) > f(j)$. $\square$

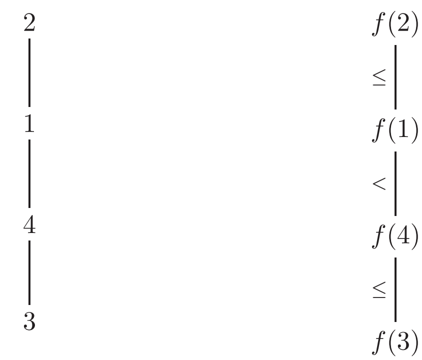
Example 5.16. To illustrate the induction in the above proof, consider the poset P from Example 5.11, having $\mathcal{L}(P) = \{3124, 3142, 3412\}$. Then choosing as incomparable pair $(i, j) = (1, 4)$, one has

$P_{i<j} =$



$\mathcal{L}(P_{i<j}) = \{3124, 3142\}$

$P_{j<i} =$



$\mathcal{L}(P_{j<i}) = \{3412\}$

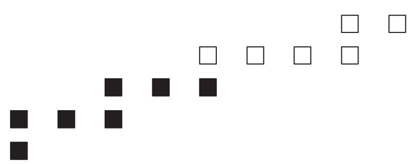
We next wish to describe the structure maps for the Hopf algebra QSym in the basis $\{L_\alpha\}$ of fundamental quasisymmetric functions. For this purpose, two more definitions are useful.

Definition 5.17. Given two nonempty compositions $\alpha = (\alpha_1, \ldots, \alpha_\ell), \beta = (\beta_1, \ldots, \beta_m)$, their *near-concatenation* is

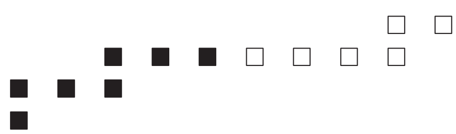
$$\alpha \odot \beta := (\alpha_1, \ldots, \alpha_{\ell-1}, \alpha_\ell + \beta_1, \beta_2, \ldots, \beta_m)$$

For example, the figure below depicts for $\alpha = (1, 3, 3)$ (black squares) and $\beta = (4, 2)$ (white squares) the concatenation and near-concatenation as ribbons:

$\alpha \cdot \beta =$

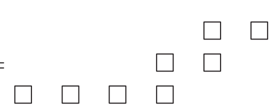


$\alpha \odot \beta =$

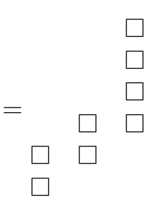


Lastly, given α in Comp_n, let $\omega(\alpha)$ be the unique composition in Comp_n whose partial sums $D(\omega(\alpha))$ form the complementary set within $[n-1]$ to the partial sums $D(\mathrm{rev}(\alpha))$; alternatively, one can check this means that the ribbon for $\omega(\alpha)$ is obtained from that of α by conjugation or transposing, that is, if $\alpha = \lambda/\mu$ then $\omega(\alpha) = \lambda^t/\mu^t$. E.g. if $\alpha = (4, 2, 2)$ so that $n = 8$, then $\mathrm{rev}(\alpha) = (2, 2, 4)$ has $D(\mathrm{rev}(\alpha)) = \{2, 4\} \subset [7]$, complementary to the set $\{1, 3, 5, 6, 7\}$ which are the partial sums for $\omega(\alpha) = (1, 2, 2, 1, 1, 1)$, and the ribbon diagrams of $\alpha, \omega(\alpha)$ are

$\alpha =$



$\omega(\alpha) =$



Proposition 5.18. *The structure maps for the Hopf algebra* QSym *in the basis* $\{L_\alpha\}$ *of fundamental quasisymmetric functions are as follows:*

$$\Delta L_\alpha = \sum_{\substack{(\beta,\gamma):\\ \beta\cdot\gamma=\alpha \text{ or } \beta\odot\gamma=\alpha}} L_\beta \otimes L_\gamma \tag{5.6}$$

$$L_\alpha L_\beta = \sum_{w \in w_\alpha \sqcup\!\sqcup\, w_\beta} L_{\gamma(w)} \tag{5.7}$$

$$S(L_\alpha) = (-1)^{|\alpha|} L_{\omega(\alpha)}. \tag{5.8}$$

Here we are making use of the following notations in (5.7) *(recall also Definition 1.39):*

- *A labelled linear order will mean a labelled poset P whose order $<_P$ is a total order. We will identify any labelled linear order P with the word (over the alphabet $\{1,2,3,\ldots\}$) obtained by writing down the elements of P in increasing order (with respect to the total order $<_P$). This way, every word (over the alphabet $\{1,2,3,\ldots\}$) which has no two equal letters becomes identified with a total labelled poset.*
- *w_α is any labelled linear order with underlying set $\{1,2,\ldots,|\alpha|\}$ such that* $\operatorname{Des}(w_\alpha) = D(\alpha)$.
- *w_β is any labelled linear order with underlying set $\{|\alpha|+1, |\alpha|+2, \ldots, |\alpha|+|\beta|\}$ such that* $\operatorname{Des}(w_\beta) = D(\beta)$.
- *$\gamma(w)$ is the unique composition of $|\alpha|+|\beta|$ with* $D(\gamma(w)) = \operatorname{Des}(w)$.

At first glance the formula (5.6) for ΔL_α might seem more complicated than the formula of Proposition 5.5 for ΔM_α. However, it is equally simple when viewed in terms of ribbon diagrams: it cuts the ribbon diagram α into two smaller ribbons β and γ, in all $|\alpha|+1$ possible ways, via *horizontal* cuts ($\beta\cdot\gamma=\alpha$) or *vertical* cuts ($\beta\odot\gamma=\alpha$). For example,

$$\begin{array}{ccccccc}
\Delta L_{(3,2)} & = 1 \otimes L_{(3,2)} & + L_{(1)} \otimes L_{(2,2)} & + L_{(2)} \otimes L_{(1,2)} & + L_{(3)} \otimes L_{(2)} & + L_{(3,1)} \otimes L_{(1)} & + L_{(3,2)} \otimes 1 \\
\begin{smallmatrix} & & \square\ \square \\ \underline{\square}\ \square & \square & \end{smallmatrix} & \begin{smallmatrix} & & \square\ \square \\ \square & |\square & \square \end{smallmatrix} & \begin{smallmatrix} & & \square\ \square \\ \square & \square & |\square \end{smallmatrix} & \begin{smallmatrix} & & \underline{\square}\ \square \\ \square & \square & \square \end{smallmatrix} & \begin{smallmatrix} & & \square\ |\square \\ \square & \square & \square \end{smallmatrix} & \begin{smallmatrix} & & \square\ \overline{\square} \\ \square & \square & \square \end{smallmatrix} &
\end{array}$$

Example 5.19. To multiply $L_{(1,1)}L_{(2)}$, one could pick $w_\alpha = 21$ and $w_\beta = 34$, and then

$$\begin{array}{llllllll}
L_{(1,1)}L_{(2)} = \sum_{w\in 21 \sqcup\!\sqcup\, 34} L_{\gamma(w)} & = L_{\gamma(2134)} & + L_{\gamma(2314)} & + L_{\gamma(3214)} & + L_{\gamma(2341)} & + L_{\gamma(3241)} & + L_{\gamma(3421)} \\
& = L_{(1,3)} & + L_{(2,2)} & + L_{(1,1,2)} & + L_{(3,1)} & + L_{(1,2,1)} & + L_{(2,1,1)}.
\end{array}$$

Proof of Proposition 5.18. To prove formula (5.6) for α in Comp_n, note that

$$\Delta L_\alpha = L_\alpha(\mathbf{x},\mathbf{y}) = \sum_{k=0}^{n} \sum_{\substack{1\le i_1\le\cdots\le i_k,\\ 1\le i_{k+1}\le\cdots\le i_n:\\ i_r<i_{r+1} \text{ for } r\in D(\alpha)\setminus\{k\}}} x_{i_1}\cdots x_{i_k}\cdot y_{i_{k+1}}\cdots y_{i_n}$$

by Lemma 5.14. One then realizes that the inner sums corresponding to values of k that lie (resp. do not lie) in $D(\alpha)\cup\{0,n\}$ correspond to the terms $L_\beta(\mathbf{x})L_\gamma(\mathbf{y})$ for pairs (β,γ) in which $\beta\cdot\gamma=\alpha$ (resp. $\beta\odot\gamma=\alpha$).

For formula (5.7), let P be the labelled poset which is the disjoint union of linear orders w_α, w_β. Then

$$L_\alpha L_\beta = F_{w_\alpha}(\mathbf{x})F_{w_\beta}(\mathbf{x}) = F_P(\mathbf{x}) = \sum_{w\in\mathcal{L}(P)} F_w(\mathbf{x}) = \sum_{w\in w_\alpha \sqcup\!\sqcup\, w_\beta} L_{\gamma(w)}$$

where the first equality used Proposition 5.12, the second equality comes from the definition of a P-partition, the third equality from Theorem 5.15, and the fourth from the equality $\mathcal{L}(P) = w_\alpha \sqcup\!\sqcup\, w_\beta$.

To prove formula (5.8), compute using Theorem 5.9 that

$$S(L_\alpha) = \sum_{\beta \text{ refining } \alpha} S(M_\beta) = \sum_{\substack{(\beta,\gamma):\\ \beta \text{ refines } \alpha,\\ \gamma \text{ coarsens } \operatorname{rev}(\beta)}} (-1)^{\ell(\beta)} M_\gamma = \sum_\gamma M_\gamma \sum_\beta (-1)^{\ell(\beta)}$$

in which the last inner sum is over β for which

$$D(\beta) \supset D(\alpha) \cup D(\operatorname{rev}(\gamma)).$$

The alternating signs make such inner sums vanish unless they have only the single term where $D(\beta) = [n-1]$ (that is, $\beta = (1^n)$). This happens exactly when $D(\mathrm{rev}(\gamma)) \cup D(\alpha) = [n-1]$ or equivalently, when $D(\mathrm{rev}(\gamma))$ contains the complement of $D(\alpha)$, that is, when $D(\gamma)$ contains the complement of $D(\mathrm{rev}(\alpha))$, that is, when γ refines $\omega(\alpha)$. Thus

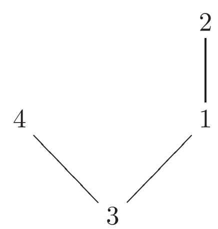
$$S(L_\alpha) = \sum_{\substack{\gamma \in \mathrm{Comp}_n: \\ \gamma \text{ refines } \omega(\alpha)}} M_\gamma \cdot (-1)^n = (-1)^{|\alpha|} L_{\omega(\alpha)}.$$

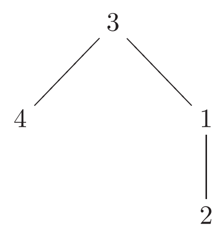
□

The antipode formula (5.8) for L_α leads to a general interpretation for the antipode of QSym acting on P-partition enumerators $F_P(\mathbf{x})$.

Definition 5.20. Given a labelled poset P on $\{1, 2, \ldots, n\}$, let the *opposite* or dual labelled poset P^{opp} have $i <_{P^{\mathrm{opp}}} j$ if and only if $j <_P i$.

For example,

$P =$ 

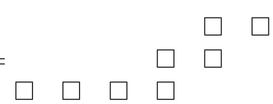
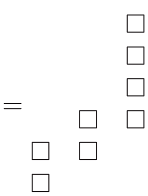
$P^{\mathrm{opp}} =$ 

The following observation is straightforward.

Proposition 5.21. *When P is a linear order corresponding to some permutation $w = (w_1, \ldots, w_n)$ in $\mathfrak{S}_n$, then $w^{\mathrm{opp}} = ww_0$ where $w_0 \in \mathfrak{S}_n$ is the permutation that swaps $i \leftrightarrow n+1-i$ (this is the so-called* **longest permutation***, thus named due to it having the highest "Coxeter length" among all permutations in $\mathfrak{S}_n$). Furthermore, in this situation one has $F_w(\mathbf{x}) = L_\alpha$, that is,* $\mathrm{Des}(w) = D(\alpha)$ *if and only if* $\mathrm{Des}(w^{\mathrm{opp}}) = D(\omega(\alpha))$*, that is $F_{w^{\mathrm{opp}}}(\mathbf{x}) = L_{\omega(\alpha)}$. Thus,*

$$S(F_w(\mathbf{x})) = (-1)^n F_{w^{\mathrm{opp}}}(\mathbf{x}).$$

For example, given the compositions considered earlier

$\alpha = (4, 2, 2) =$  $\omega(\alpha) = (1, 2, 2, 1, 1, 1) =$ 

if one picks $w = 1235 \cdot 47 \cdot 68$ (with descent positions marked by dots) having $\mathrm{Des}(w) = \{4, 6\} = D(\alpha)$, then $w^{\mathrm{opp}} = ww_0 = 8 \cdot 67 \cdot 45 \cdot 3 \cdot 2 \cdot 1$ has $\mathrm{Des}(w^{\mathrm{opp}}) = \{1, 3, 5, 6, 7\} = D(\omega(\alpha))$.

Corollary 5.22. *For any labelled poset P on $\{1, 2, \ldots, n\}$, one has*

$$S\left(F_P(\mathbf{x})\right) = (-1)^n F_{P^{\mathrm{opp}}}(\mathbf{x}).$$

Proof. Since S is linear, one can apply Theorem 5.15 and Proposition 5.21

$$S\left(F_P(\mathbf{x})\right) = \sum_{w \in \mathcal{L}(P)} S(F_w(\mathbf{x})) = \sum_{w \in \mathcal{L}(P)} (-1)^n F_{w^{\mathrm{opp}}}(\mathbf{x}) = (-1)^n F_{P^{\mathrm{opp}}}(\mathbf{x}).$$

as $\mathcal{L}(P^{\mathrm{opp}}) = \{w^{\mathrm{opp}} : w \in \mathcal{L}(P)\}$. □

Remark 5.23. Malvenuto and Reutenauer, in [52, Theorem 3.1], prove an even more general antipode formula, which encompasses our Corollary 5.22, Proposition 5.21, Theorem 5.9 and (5.8).

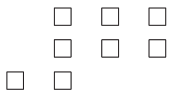
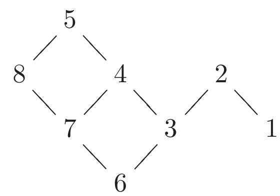
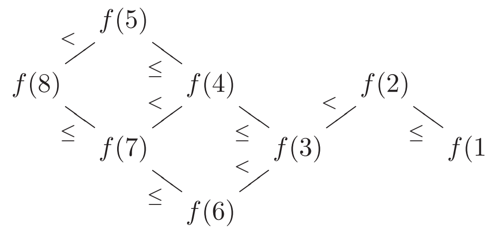
We remark on a special case of Corollary 5.22 to which we alluded earlier, related to skew Schur functions.

Corollary 5.24. *In Λ, the action of ω and the antipode S on skew Schur functions $s_{\lambda/\mu}$ are as follows:*

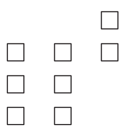
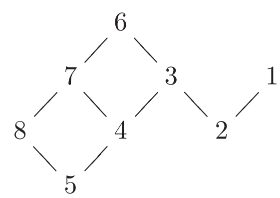
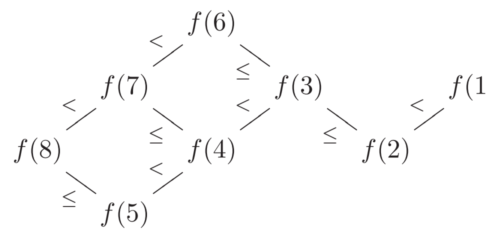
$$\omega(s_{\lambda/\mu}) = s_{\lambda^t/\mu^t} \tag{5.9}$$

$$S(s_{\lambda/\mu}) = (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t}. \tag{5.10}$$

Proof. Given a skew shape λ/μ, one can always create a labelled poset P which is its *skew Ferrers poset*, together with one of many *column-strict labellings*, in such a way that $F_P(\mathbf{x}) = s_{\lambda/\mu}(\mathbf{x})$. An example is shown here for $\lambda/\mu = (4,4,2)/(1,1,0)$:

$\lambda/\mu =$  $P =$  

The general definition is as follows: Let P be the set of all boxes of the skew diagram λ/μ. Label these boxes by the numbers $1, 2, \ldots, n$ (where $n = |\lambda/\mu|$) row by row from bottom to top (reading every row from left to right), and then define an order relation $<_P$ on P by requiring that every box be smaller (in P) than its right neighbor and smaller (in P) than its lower neighbor. It is not hard to see that in this situation, $F_{P^{\mathrm{opp}}}(\mathbf{x}) = \sum_T \mathbf{x}^{\mathrm{cont}(T)}$ as T ranges over all *reverse semistandard tableaux* or *column-strict plane partitions* of λ^t/μ^t:

$\lambda^t/\mu^t =$  $P^{\mathrm{opp}} =$  

But this means that $F_{P^{\mathrm{opp}}}(\mathbf{x}) = s_{\lambda^t/\mu^t}(\mathbf{x})$, since the fact that skew Schur functions lie in Λ implies that they can be defined either as generating functions for column-strict tableaux or reverse semistandard tableaux; see Remark 2.10 above, or [72, Prop. 7.10.4].

Thus we have

$$F_P(\mathbf{x}) = s_{\lambda/\mu}(\mathbf{x})$$
$$F_{P^{\mathrm{opp}}}(\mathbf{x}) = s_{\lambda^t/\mu^t}(\mathbf{x}).$$

Proposition 1.35(c) tell us that the antipode for QSym must specialize to the antipode for Λ (see also Remark 5.34 below), so (5.10) is a special case of Corollary 5.22. Then (5.9) follows from the relation (2.14) that $S(f) = (-1)^n \omega(f)$ for f in Λ_n. $\square$

Remark 5.25. Before leaving P-partitions temporarily, we mention two open questions about them.

The first is a conjecture of Stanley from his thesis [69]. As mentioned in the proof of Corollary 5.24, each skew Schur function $s_{\lambda/\mu}(\mathbf{x})$ is a special instance of P-partition enumerator $F_P(\mathbf{x})$.

Conjecture 5.26. *A labelled poset P has $F_P(\mathbf{x})$ symmetric, and not just quasisymmetric, if and only if P is a column-strict labelling of some skew Ferrers poset λ/μ.*

A somewhat weaker result in this direction was proven by Malvenuto in her thesis [50, Thm. 6.4], showing that if a labelled poset P has the stronger property that its set of linear extensions $\mathcal{L}(P)$ is a union of *plactic* or *Knuth equivalence classes*, then P must be a column-strict labelling of a skew Ferrers poset.

The next question is due to P. McNamara, and is suggested by the obvious factorizations of P-partition enumerators $F_{P_1 \sqcup P_2}(\mathbf{x}) = F_{P_1}(\mathbf{x}) F_{P_2}(\mathbf{x})$.

Question 5.27. Does a *connected* labelled poset P always have $F_P(\mathbf{x})$ *irreducible* within the ring QSym?

The phrasing of this question requires further comment. It is assumed here that $\mathbf{x} = (x_1, x_2, \ldots)$ is infinite; for example when P is a 2-element chain labelled "against the grain" (i.e., the bigger element of the chain has the smaller label), then $F_P(\mathbf{x}) = e_2(\mathbf{x})$ is irreducible, but its specialization to two variables $\mathbf{x} = (x_1, x_2)$ is $e_2(x_1, x_2) = x_1 x_2$, which is reducible. If one wishes to work in finitely many variables $\mathbf{x} = (x_1, \ldots, x_m)$ one can perhaps assume that m is at least $|P| + 1$.

When working in QSym = QSym($\mathbf{x}$) in infinitely many variables, it is perhaps not so clear where factorizations occur. For example, if f lies in QSym and factors $f = g \cdot h$ with g, h in $R(\mathbf{x})$, does this imply that g, h also lie in QSym? The answer is "Yes", but this is not obvious, and was proven by P. Pylyavskyy in [60, Chap. 11].

One also might wonder whether QSym is a unique factorization domain, but this follows from the result of M. Hazewinkel [28] who proved a conjecture of Ditters that QSym(:= $\mathrm{QSym}_{\mathbb{Z}}$) is a polynomial algebra; earlier Malvenuto and Reutenauer [51, Cor. 2.2] had shown that $\mathrm{QSym}_{\mathbb{Q}}$ is a polynomial algebra. In fact, one can find polynomial generators $\{P_\alpha\}$ for $\mathrm{QSym}_{\mathbb{Q}}$ as a subset of the dual basis to the $\mathbb{Q}$-basis $\{\pi_\alpha\}$ for $\mathrm{NSym}_{\mathbb{Q}}$ which comes from taking products $\pi_\alpha := \pi_{\alpha_1} \cdots \pi_{\alpha_\ell}$ of the elements $\{\pi_n\}$ defined in Remark 5.31 below. Specifically, one takes those P_α for which the composition α is a *Lyndon composition*.

An affirmative answer to Question 5.27 is known at least in the special case where P is a connected column-strict labelling of a skew Ferrers diagram, that is, when $F_P(\mathbf{x}) = s_{\lambda/\mu}(\mathbf{x})$ for some connected skew diagram λ/μ; see [9].

5.3. **The Hopf algebra** NSym **dual to** QSym**.** We introduce here the dual Hopf algebra to QSym.

Definition 5.28. Let NSym := QSym^o, with dual pairing $\mathrm{NSym} \otimes \mathrm{QSym} \xrightarrow{(\cdot,\cdot)} \mathbf{k}$. Let $\{H_\alpha\}$ be the $\mathbf{k}$-basis of NSym dual to the $\mathbf{k}$-basis $\{M_\alpha\}$ of QSym, so that

$$(H_\alpha, M_\beta) = \delta_{\alpha,\beta}.$$

When the base ring $\mathbf{k}$ is not clear from the context, we write $\mathrm{NSym}_{\mathbf{k}}$ in lieu of NSym.

Theorem 5.29. *Letting $H_n := H_{(n)}$ for $n = 0, 1, 2, \ldots$, with $H_0 = 1$, one has that*

$$\mathrm{NSym} \cong \mathbf{k}\langle H_1, H_2, \ldots \rangle,$$

the free associative (but not commutative) algebra on generators $\{H_1, H_2, \ldots\}$ with coproduct determined by[22]

$$\Delta H_n = \sum_{i+j=n} H_i \otimes H_j. \tag{5.11}$$

Proof. Since Proposition 5.5 asserts that $\Delta M_\alpha = \sum_{(\beta,\gamma):\beta\cdot\gamma=\alpha} M_\beta \otimes M_\gamma$, and since $\{H_\alpha\}$ are dual to $\{M_\alpha\}$, one concludes that for any compositions β, γ, one has

$$H_\beta H_\gamma = H_{\beta\cdot\gamma}$$

Iterating this gives

$$H_\alpha = H_{(\alpha_1,\ldots,\alpha_\ell)} = H_{\alpha_1} \cdots H_{\alpha_\ell}.$$

Since the H_α are a $\mathbf{k}$-basis for NSym, this shows $\mathrm{NSym} \cong \mathbf{k}\langle H_1, H_2, \ldots \rangle$.

Note that $H_n = H_{(n)}$ is dual to $M_{(n)}$, so to understand ΔH_n, one should understand how $M_{(n)}$ can appear as a term in the product $M_\alpha M_\beta$. By (5.1) this occurs only if $\alpha = (i), \beta = (j)$ where $i + j = n$, where

$$M_{(i)} M_{(j)} = M_{(i+j)} + M_{(i,j)} + M_{(j,i)}$$

(where the $M_{(i,j)}$ and $M_{(j,i)}$ addends have to be disregarded if one of i and j is 0). By duality, this implies the formula (5.11). $\square$

[22] The abbreviated summation indexing $\sum_{i+j=n} t_{i,j}$ used here is intended to mean

$$\sum_{\substack{(i,j):\\ 0 \le i,j \le n,\\ i+j=n}} t_{i,j}.$$

Corollary 5.30. *The algebra homomorphism defined by*

$$\begin{array}{rcl} \mathrm{NSym} & \xrightarrow{\pi} & \Lambda \\ H_n & \longmapsto & h_n \end{array}$$

is a Hopf algebra surjection, and adjoint to the inclusion $\Lambda \overset{i}{\hookrightarrow} \mathrm{QSym}$.

Proof. As an algebra map π may be identified with the surjection $T(V) \to \mathrm{Sym}(V)$ from the tensor algebra on a graded free $\mathbf{k}$-module V with basis $\{H_1, H_2, \ldots\}$ to the symmetric algebra on V, since

$$\mathrm{NSym} \cong \mathbf{k}\langle H_1, H_2, \ldots \rangle$$
$$\Lambda \cong \mathbf{k}[h_1, h_2, \ldots]$$

As (5.11) and Proposition 2.18(iii) assert that

$$\Delta H_n = \sum_{i+j=n} H_i \otimes H_j$$
$$\Delta h_n = \sum_{i+j=n} h_i \otimes h_j$$

this map π is also a bialgebra morphism, and hence a Hopf morphism by Proposition 1.35.

To check π is adjoint to i, let $\lambda(\alpha)$ denote the partition which is the weakly decreasing rearrangement of the composition α, and note that the bases $\{H_\alpha\}$ of NSym and $\{m_\lambda\}$ of Λ satisfy

$$(\pi(H_\alpha), m_\lambda) = (h_{\lambda(\alpha)}, m_\lambda) = \left\{ \begin{matrix} 1 & \text{if } \lambda(\alpha) = \lambda \\ 0 & \text{otherwise} \end{matrix} \right\} = \left(H_\alpha, \sum_{\beta : \lambda(\beta) = \lambda} M_\beta \right) = (H_\alpha, i(m_\lambda)).$$

□

Remark 5.31. For those who prefer generating functions to sign-reversing involutions, we sketch here Malvenuto and Reutenauer's elegant proof [51, Cor. 2.3] of the antipode formula (Theorem 5.9). One needs to know that when A is a $\mathbf{k}$-algebra (possibly noncommutative) with $\mathbf{k}$ of characteristic zero, in the ring of power series $A[[t]]$ where t commutes with all of A, one still has familiar facts, such as

$$a(t) = \log b(t) \quad \text{if and only if} \quad b(t) = \exp a(t)$$

and whenever $a(t), b(t)$ commute in $A[[t]]$, one has

$$\exp\left(a(t) + b(t)\right) = \exp a(t) \exp b(t) \tag{5.12}$$
$$\log\left(a(t) b(t)\right) = \log a(t) + \log b(t) \tag{5.13}$$

Start by assuming WLOG that $\mathbf{k} = \mathbb{Z}$ (as $\mathrm{NSym}_{\mathbf{k}} = \mathrm{NSym}_{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbf{k}$ in the general case). Now, define in $\mathrm{NSym}_{\mathbb{Q}} = \mathrm{NSym} \otimes_{\mathbb{Z}} \mathbb{Q}$ the elements $\{\pi_1, \pi_2, \ldots\}$ via generating functions in $\mathrm{NSym}_{\mathbb{Q}}[[t]]$:

$$\begin{aligned} H(t) &:= \sum_{n \geq 0} H_n t^n, \\ \pi(t) &:= \sum_{n \geq 1} \pi_n t^n = \log H(t) \end{aligned} \tag{5.14}$$

One first checks that this makes each π_n primitive, via a computation in the ring $(\mathrm{NSym}_{\mathbb{Q}} \otimes \mathrm{NSym}_{\mathbb{Q}})[[t]]$ (into which we "embed" the ring $(\mathrm{NSym}_{\mathbb{Q}}[[t]]) \otimes_{\mathbb{Q}[[t]]} (\mathrm{NSym}_{\mathbb{Q}}[[t]])$ via the canonical ring homomorphism from the latter into the former [23]):

$$\begin{aligned}\Delta\pi(t) &= \log \sum_{n\geq 0} \Delta(H_n)t^n = \log \sum_{n\geq 0} \left(\sum_{i+j=n} H_i \otimes H_j \right) t^n \\ &= \log \left(\left(\sum_{i\geq 0} H_i t^i \right) \otimes \left(\sum_{j\geq 0} H_j t^j \right) \right) = \log \left(\left(\sum_{i\geq 0} H_i t^i \otimes 1 \right) \left(1 \otimes \sum_{j\geq 0} H_j t^j \right) \right) \\ &\overset{(5.13)}{=} \log H(t) \otimes 1 + 1 \otimes \log H(t) = \pi(t) \otimes 1 + 1 \otimes \pi(t).\end{aligned}$$

Comparing coefficients in this equality yields $\Delta(\pi_n) = \pi_n \otimes 1 + 1 \otimes \pi_n$. Thus $S(\pi_n) = -\pi_n$, by Proposition 1.31. This allows one to determine $S(H_n)$ and $S(H_\alpha)$, after one first inverts the relation (5.14) to get that $H(t) = \exp \pi(t)$, and hence

$$\begin{aligned}S(H(t)) = S(\exp \pi(t)) = \exp S(\pi(t)) = \exp(-\pi(t)) &\overset{(5.12)}{=} (\exp \pi(t))^{-1} \\ = H(t)^{-1} = \left(1 + H_1 t + H_2 t^2 + \cdots\right)^{-1}.&\end{aligned}$$

Upon expanding the right side, and comparing coefficients of t^n, this gives

$$S(H_n) = \sum_{\beta \in \mathrm{Comp}_n} (-1)^{\ell(\beta)} H_\beta$$

and hence

$$S(H_\alpha) = S(H_{\alpha_\ell}) \cdots S(H_{\alpha_2}) S(H_{\alpha_1}) = \sum_{\substack{\gamma: \\ \gamma \text{ refines } \operatorname{rev}(\alpha)}} (-1)^{\ell(\gamma)} H_\gamma.$$

As $S_{\mathrm{NSym}}, S_{\mathrm{QSym}}$ are adjoint, and $\{H_\alpha\}, \{M_\alpha\}$ are dual bases, this is equivalent to Theorem 5.9:

$$S(M_\alpha) = (-1)^{\ell(\alpha)} \sum_{\substack{\gamma: \\ \gamma \text{ coarsens } \operatorname{rev}(\alpha)}} M_\gamma$$

(because if μ and ν are two compositions, then μ coarsens ν if and only if $\operatorname{rev}(\mu)$ coarsens $\operatorname{rev}(\nu)$).

We next explore the basis for NSym dual to the $\{L_\alpha\}$ in QSym.

Definition 5.32. Define the *noncommutative ribbon functions* $\{R_\alpha\}$ to be the $\mathbf{k}$-basis of NSym dual to the fundamental basis $\{L_\alpha\}$ of QSym, so that $(R_\alpha, L_\beta) = \delta_{\alpha,\beta}$.

Theorem 5.33. *One has that*

$$H_\alpha = \sum_{\beta \text{ coarsens } \alpha} R_\beta \tag{5.15}$$

$$R_\alpha = \sum_{\beta \text{ coarsens } \alpha} (-1)^{\ell(\beta)-\ell(\alpha)} H_\beta \tag{5.16}$$

and the surjection $\mathrm{NSym} \xrightarrow{\pi} \Lambda$ *sends* $R_\alpha \longmapsto s_\alpha$*, the skew Schur function associated to the ribbon* α*. Furthermore,*

$$R_\alpha R_\beta = R_{\alpha\cdot\beta} + R_{\alpha\odot\beta} \qquad \text{if } \alpha \text{ and } \beta \text{ are nonempty} \tag{5.17}$$

$$S(R_\alpha) = (-1)^{|\alpha|} R_{\omega(\alpha)} \tag{5.18}$$

Finally, $R_\varnothing$ *is the multiplicative identity of* NSym.

Proof. For the first assertion, note that

$$H_\alpha = \sum_\beta (H_\alpha, L_\beta) R_\beta = \sum_\beta \left(H_\alpha, \sum_{\substack{\gamma: \\ \gamma \text{ refines } \beta}} M_\gamma \right) R_\beta = \sum_{\substack{\beta: \\ \beta \text{ coarsens } \alpha}} R_\beta.$$

The second assertion follows from the first by inclusion-exclusion.

[23] This ring homomorphism fails to be injective, whence the "embed" stands in quotation marks.

Write α as $(\alpha_1, \ldots, \alpha_\ell)$. To show that $\pi(R_\alpha) = s_\alpha$, we instead examine $\pi(H_\alpha)$:

$$\pi(H_\alpha) = \pi(h_{\alpha_1} \cdots h_{\alpha_\ell}) = h_{\alpha_1} \cdots h_{\alpha_\ell} = s_{(\alpha_1)} \cdots s_{(\alpha_\ell)} = s_{(\alpha_1)\oplus\cdots\oplus(\alpha_\ell)}$$

where $(\alpha_1) \oplus \cdots \oplus (\alpha_\ell)$ is some skew shape which is a horizontal strip having rows of lengths $\alpha_1, \ldots, \alpha_\ell$ from bottom to top. We claim

$$s_{(\alpha_1)\oplus\cdots\oplus(\alpha_\ell)} = \sum_{\substack{\beta: \\ \beta \text{ coarsens } \alpha}} s_\beta$$

because column-strict tableaux T of shape $(\alpha_1) \oplus \cdots \oplus (\alpha_\ell)$ biject to column-strict tableaux T' of some ribbon β coarsening α, as follows: let a_i, b_i denote the leftmost, rightmost entries of the i^{th} row from the bottom in T, of length α_i, and

- if $b_i \leq a_{i+1}$, merge parts α_i, α_{i+1} in β, and concatenate the rows of length α_i, α_{i+1} in T', or
- if $b_i > a_{i+1}$, do not merge parts α_i, α_{i+1} in β, and let these two rows overlap in one column in T'

E.g., if $\alpha = (3, 3, 2, 3, 2)$, this T of shape $(\alpha_1) \oplus \cdots \oplus (\alpha_\ell)$ maps to this T' of shape $\beta = (3, 8, 2)$:

$$T = \begin{array}{ccccccccccccc}
 & & & & & & & & & & & 3 & 4 \\
 & & & & & & & & 4 & 4 & 5 & & \\
 & & & & & & 4 & 4 & & & & & \\
 & & & 2 & 2 & 3 & & & & & & & \\
1 & 1 & 3 & & & & & & & & & &
\end{array}
\longmapsto
\begin{array}{ccccccccccc}
 & & & & & & & & & 3 & 4 \\
 & & 2 & 2 & 3 & 4 & 4 & 4 & 4 & 5 & \\
1 & 1 & 3 & & & & & & & &
\end{array}$$

The reverse bijection breaks the rows of T' into the rows of T of lengths dictated by the parts of α. Having shown $\pi(H_\alpha) = \sum_{\beta:\beta \text{ coarsens } \alpha} s_\beta$, the relation (5.15) and inclusion-exclusion show $\pi(R_\alpha) = s_\alpha$.

Finally, (5.17) and (5.18) follow from (5.6) and (5.8) by duality. $\square$

Remark 5.34. Since the maps

$$\begin{tikzcd}
\mathrm{NSym} \arrow[dr, "\pi"'] & & \mathrm{QSym} \\
 & \Lambda \arrow[ur, hook, "i"'] &
\end{tikzcd}$$

are Hopf morphisms, they must respect the antipodes $S_\Lambda, S_{\mathrm{QSym}}, S_{\mathrm{NSym}}$, but it is interesting to compare them explicitly using the fundamental basis for QSym and the ribbon basis for NSym.

On one hand (5.8) shows that $S_{\mathrm{QSym}}(L_\alpha) = (-1)^{|\alpha|} L_{\omega(\alpha)}$ extends the map S_Λ since $L_{(1^n)} = e_n$ and $L_{(n)} = h_n$, as observed in Example 5.13, and $\omega((n)) = (1^n)$.

On the other hand, (5.18) shows that $S_{\mathrm{NSym}}(R_\alpha) = (-1)^{|\alpha|} R_{\omega(\alpha)}$ lifts the map S_Λ to S_{NSym}: Theorem 5.33 showed that R_α lifts the skew Schur function s_α, while (2.15) asserted that $S(s_{\lambda/\mu}) = (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t}$, and a ribbon $\alpha = \lambda/\mu$ has $\omega(\alpha) = \lambda^t/\mu^t$.

5.4. **Polynomial generators for** QSym **and Lyndon words.** Perhaps to be filled in later....

5.5. **Application: Multiple zeta values and Hoffman's stuffle conjecture.** Perhaps to be filled in later....

6. Aguiar-Bergeron-Sottile character theory Part I: QSym as a terminal object

It turns out that the universal mapping property of NSym as a free associative algebra leads via duality to a universal property for its dual QSym, elegantly explaining several combinatorial invariants that take the form of quasisymmetric or symmetric functions:

- Ehrenborg's quasisymmetric function of a *ranked poset* [24],
- Stanley's *chromatic* symmetric function of a *graph* [71],
- the quasisymmetric function of a *matroid* considered in [11]

6.1. **Characters and the universal property.**

Definition 6.1. Given a Hopf algebra A over $\mathbf{k}$, a *character* is an algebra morphism $A \xrightarrow{\zeta} \mathbf{k}$, that is,

- $\zeta(1_A) = 1_{\mathbf{k}}$,
- ζ is $\mathbf{k}$-linear, and
- $\zeta(ab) = \zeta(a)\zeta(b)$ for a, b in A.

Example 6.2. A particularly important character for $A = \mathrm{QSym}$ is defined as follows:

$$\begin{array}{rcl} \mathrm{QSym} & \xrightarrow{\zeta_Q} & \mathbf{k} \\ f(\mathbf{x}) & \longmapsto & f(1,0,0,\ldots) = [f(\mathbf{x})]_{x_1=1, x_2=x_3=\cdots=0} \,. \end{array}$$

Hence,

$$\zeta_Q(M_\alpha) = \zeta_Q(L_\alpha) = \begin{cases} 1 & \text{if } \alpha = (n) \text{ for some } n, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, the restriction $\zeta_Q|_{\mathrm{QSym}_n}$ coincides with the functional H_n in $\mathrm{NSym}_n = \mathrm{Hom}_{\mathbf{k}}(\mathrm{QSym}_n, \mathbf{k})$: one has for f in QSym_n that $\zeta_Q(f) = (H_n, f)$.

It is worth remarking that there is nothing special about setting $x_1 = 1$ and $x_2 = x_3 = \cdots = 0$: for quasisymmetric f, we could have defined the same character ζ_Q by picking any variable, say x_n, and sending

$$f(\mathbf{x}) \longmapsto [f(\mathbf{x})]_{\substack{x_n=1, \text{ and} \\ x_m=0 \text{ for } m\neq n}} \,.$$

This character $\mathrm{QSym} \xrightarrow{\zeta_Q} \mathbf{k}$ has a certain universal property.

Theorem 6.3. *A connected graded Hopf algebra A of finite type together with a character $A \xrightarrow{\zeta} \mathbf{k}$ induces a unique graded Hopf morphism $A \xrightarrow{\Psi} \mathrm{QSym}$ making this diagram commute:*

$$\begin{tikzcd} A \arrow[rr, "\Psi"] \arrow[dr, "\zeta"'] & & \mathrm{QSym} \arrow[dl, "\zeta_Q"] \\ & \mathbf{k} & \end{tikzcd} \tag{6.1}$$

Furthermore, Ψ has this formula on elements of A_n:

$$\Psi(a) = \sum_{\alpha \in \mathrm{Comp}_n} \zeta_\alpha(a) M_\alpha \tag{6.2}$$

where for $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, the map ζ_α is the composite

$$A_n \xrightarrow{\Delta^{(\ell-1)}} A^{\otimes \ell} \xrightarrow{\pi_\alpha} A_{\alpha_1} \otimes \cdots \otimes A_{\alpha_\ell} \xrightarrow{\zeta^{\otimes \ell}} \mathbf{k}$$

in which $A^{\otimes \ell} \xrightarrow{\pi_\alpha} A_{\alpha_1} \otimes \cdots \otimes A_{\alpha_\ell}$ is the canonical projection.

Proof. One argues that Ψ is unique, and has formula (6.2), using only that ζ is $\mathbf{k}$-linear and sends 1 to 1 and that Ψ is a graded $\mathbf{k}$-*coalgebra* map making (6.1) commute. Equivalently, consider the adjoint $\mathbf{k}$-*algebra* map

$$\mathrm{NSym} = \mathrm{QSym}^o \xrightarrow{\Psi^*} A^o.$$

Commutativity of (6.1) implies that for a in A_n,

$$(\Psi^*(H_n), a) = (H_n, \Psi(a)) = \zeta_Q(\Psi(a)) = \zeta(a)$$

where the second equality used Example 6.2. In other words, $\Psi^*(H_n)$ is the element of A^o defined as the following functional on A:

$$\Psi^*(H_n)(a) = \begin{cases} \zeta(a) & \text{if } a \in A_n, \\ 0 & \text{if } a \in A_m \text{ for some } m \neq n. \end{cases} \tag{6.3}$$

By the universal property for $\mathrm{NSym} \cong \mathbf{k}\langle H_1, H_2, \ldots \rangle$ as free associative $\mathbf{k}$-algebra, we see that any choice of a $\mathbf{k}$-linear map $A \xrightarrow{\zeta} \mathbf{k}$ uniquely produces a $\mathbf{k}$-algebra map $\Psi^* : \mathrm{QSym}^o \to A^o$ which satisfies (6.3) for all $n \geq 1$. It is easy to see that this Ψ^* then automatically satisfies (6.3) for $n = 0$ as well if ζ sends 1 to 1 (it is here that we use $\zeta(1) = 1$ and the connectedness of A). Hence, any given $\mathbf{k}$-linear map $A \xrightarrow{\zeta} \mathbf{k}$ sending 1 to 1 uniquely produces a $\mathbf{k}$-algebra map $\Psi^* : \mathrm{QSym}^o \to A^o$ which satisfies (6.3) for all $n \geq 0$. Formula (6.2) follows as

$$\Psi(a) = \sum_{\alpha \in \mathrm{Comp}} (H_\alpha, \Psi(a)) \; M_\alpha$$

and for a composition $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, one has

$$\begin{aligned} (H_\alpha, \Psi(a)) = (\Psi^*(H_\alpha), a) &= (\Psi^*(H_{\alpha_1}) \cdots \Psi^*(H_{\alpha_\ell}), a) \\ &= (\Psi^*(H_{\alpha_1}) \otimes \cdots \otimes \Psi^*(H_{\alpha_\ell}), (\pi_\alpha \circ \Delta^{(\ell-1)})(a)) = \zeta_\alpha(a) \end{aligned}$$

using (6.3) and the definition of ζ_α.

We wish to show that if, in addition, A is a Hopf algebra and $A \xrightarrow{\zeta} \mathbf{k}$ is a character (algebra map), then $A \xrightarrow{\Psi} \mathrm{QSym}$ will be an algebra map, that is, the two maps $A \otimes A \longrightarrow \mathrm{QSym}$ given by $\Psi \circ m$ and $m \circ (\Psi \otimes \Psi)$ coincide. To see this, consider these two diagrams having the two maps in question as the composites of their top rows:

$$\begin{tikzcd} A \otimes A \arrow[r, "m"] \arrow[dr, "\zeta\otimes\zeta"'] & A \arrow[r, "\Psi"] \arrow[d, "\zeta"] & \mathrm{QSym} \arrow[dl, "\zeta_Q"] \\ & \mathbf{k} & \end{tikzcd} \qquad \begin{tikzcd} A \otimes A \arrow[r, "\Psi\otimes\Psi"] \arrow[dr, "\zeta\otimes\zeta"'] & \mathrm{QSym}^{\otimes 2} \arrow[r, "m"] \arrow[d, "\zeta_Q\otimes\zeta_Q"] & \mathrm{QSym} \arrow[dl, "\zeta_Q"] \\ & \mathbf{k} & \end{tikzcd} \tag{6.4}$$

The fact that ζ, ζ_Q are algebra maps makes the above diagrams commute, so that applying the uniqueness in the first part of the proof to the character $A \otimes A \xrightarrow{\zeta\otimes\zeta} \mathbf{k}$ proves the desired equality $\Psi \circ m = m \circ (\Psi \otimes \Psi)$. $\square$

Remark 6.4. When one assumes in addition that A is cocommutative, it follows that the image of Ψ will lie in the subalgebra $\Lambda \subset \mathrm{QSym}$, e.g. from the explicit formula (6.2) and the fact that one will have $\zeta_\alpha = \zeta_\beta$ whenever β is a rearrangement of α. In other words, the character $\Lambda \xrightarrow{\zeta_\Lambda} \mathbf{k}$ defined by restricting ζ_Q to Λ, or by

$$\zeta_\Lambda(m_\lambda) = \begin{cases} 1 & \text{if } \lambda = (n) \text{ for some } n, \\ 0 & \text{otherwise,} \end{cases}$$

has a universal property as terminal object with respect to characters on cocommutative co- or Hopf algebras.

We close this section by discussing a well-known polynomiality and reciprocity phenomenon; see, e.g., Humpert and Martin [37, Prop. 2.2], Stanley [71, §4].

Definition 6.5. For a field $\mathbf{k}$, the *binomial Hopf algebra* is the polynomial algebra $\mathbf{k}[m]$ in a single variable m, with a Hopf algebra structure transported from the symmetric algebra $\mathrm{Sym}\left(\mathbf{k}^1\right)$ (which is a Hopf algebra by virtue of Example 1.18, applied to $V = \mathbf{k}^1$) along the isomorphism $\mathrm{Sym}\left(\mathbf{k}^1\right) \to \mathbf{k}[m]$ which sends the standard basis element of $\mathbf{k}^1$ to m. Thus the element m is primitive; that is, $\Delta m = 1 \otimes m + m \otimes 1$ and $S(m) = -m$. As S is an algebra anti-endomorphism by Proposition 1.26 and $\mathbf{k}[m]$ is commutative, one has $S(g)(m) = g(-m)$ for all polynomials $g(m)$ in $\mathbf{k}[m]$.

Definition 6.6. For an element $f(\mathbf{x})$ in QSym and a nonnegative integer m, let $\mathrm{ps}^1(f)(m)$ denote the element of $\mathbf{k}$ obtained by *principal specialization at* $q = 1$

$$\begin{aligned}\mathrm{ps}^1(f)(m) &= [f(\mathbf{x})]_{\substack{x_1=x_2=\cdots=x_m=1,\\ x_{m+1}=x_{m+2}=\cdots=0}} \\ &= f(\underbrace{1,1,\ldots,1}_{m \text{ ones}},0,0,\ldots).\end{aligned}$$

Proposition 6.7. *Assume that the field* $\mathbf{k}$ *has characteristic* 0*. The map* ps^1 *has the following properties.*

(i) *There is a unique polynomial in* $\mathbf{k}[m]$ *which agrees for each nonnegative integer* m *with* $\mathrm{ps}^1(f)(m)$*, and which, by abuse of notation, we will also denote* $\mathrm{ps}^1(f)(m)$*. If* f *lies in* QSym_n*, then* $\mathrm{ps}^1(f)(m)$ *is a polynomial of degree at most* n*, taking these values on* M_α, L_α *for* $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ *in* Comp_n*:*

$$\begin{aligned}\mathrm{ps}^1(M_\alpha)(m) &= \binom{m}{\ell}, \\ \mathrm{ps}^1(L_\alpha)(m) &= \binom{m-\ell+n}{n}.\end{aligned}$$

(ii) *The map* $\mathrm{QSym} \xrightarrow{\mathrm{ps}^1} \mathbf{k}[m]$ *is a Hopf morphism into the binomial Hopf algebra.*

(iii) *For all* m *in* $\mathbb{Z}$ *and* f *in* QSym *one has*

$$\zeta_Q^{\star m}(f) = \mathrm{ps}^1(f)(m).$$

In particular, one also has

$$\zeta_Q^{\star(-m)}(f) = \mathrm{ps}^1(S(f))(m) = \mathrm{ps}^1(f)(-m).$$

(iv) *For a graded Hopf algebra* A *of finite type with a character* $A \xrightarrow{\zeta} \mathbf{k}$*, and any element* a *in* A_n*, the polynomial* $\mathrm{ps}^1(\Psi(a))(m)$ *in* $\mathbf{k}[m]$ *has degree at most* m*, and when specialized to* m *in* $\mathbb{Z}$ *satisfies*

$$\zeta^{\star m}(a) = \mathrm{ps}^1(\Psi(a))(m).$$

Proof. To prove assertion (i), note that one has

$$\begin{aligned}\mathrm{ps}^1(M_\alpha)(m) = M_\alpha(1,1,\ldots,1,0,0,\ldots) &= \sum_{1 \le i_1 < \cdots < i_\ell \le m} \left[x_{i_1}^{\alpha_1} \cdots x_{i_\ell}^{\alpha_\ell}\right]_{x_j=1} = \binom{m}{\ell} \\ \mathrm{ps}^1(L_\alpha)(m) = L_\alpha(1,1,\ldots,1,0,0,\ldots) &= \sum_{\substack{1 \le i_1 \le \cdots \le i_n \le m \\ i_k < i_{k+1} \text{ if } k \in D(\alpha)}} \left[x_{i_1} \cdots x_{i_n}\right]_{x_j=1} \\ &= |\{1 \le j_1 \le j_2 \le \cdots \le j_n \le m-\ell+1\}| = \binom{m-\ell+n}{n}.\end{aligned}$$

As $\{M_\alpha\}_{\alpha \in \mathrm{Comp}_n}$ form a basis for QSym_n, and $\binom{m}{\ell}$ is a polynomial function in m of degree $\ell (\le n)$, one concludes that for f in QSym_n one has that $\mathrm{ps}^1(f)(m)$ is a polynomial function in m of degree at most n. The polynomial giving rise to this function is unique, since infinitely many of its values are fixed.

To prove assertion (ii), note that ps^1 is an algebra morphism because it is an evaluation homomorphism. To check that it is a coalgebra morphism, it suffices to check $\Delta \circ \mathrm{ps}^1 = (\mathrm{ps}^1 \otimes \mathrm{ps}^1) \circ \Delta$ on each M_α for $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ in Comp_n. Using the Vandermonde summation $\binom{A+B}{\ell} = \sum_k \binom{A}{k}\binom{B}{\ell-k}$, one has

$$(\Delta \circ \mathrm{ps}^1)(M_\alpha) = \Delta \binom{m}{\ell} = \binom{m \otimes 1 + 1 \otimes m}{\ell} = \sum_{k=0}^{\ell} \binom{m \otimes 1}{k} \binom{1 \otimes m}{\ell - k} = \sum_{k=0}^{\ell} \binom{m}{k} \otimes \binom{m}{\ell-k}$$

while at the same time

$$\left((\mathrm{ps}^1 \otimes \mathrm{ps}^1) \circ \Delta\right)(M_\alpha) = \sum_{k=0}^{\ell} \mathrm{ps}^1(M_{(\alpha_1,\ldots,\alpha_k)}) \otimes \mathrm{ps}^1(M_{(\alpha_{k+1},\ldots,\alpha_\ell)}) = \sum_{k=0}^{\ell} \binom{m}{k} \otimes \binom{m}{\ell-k}.$$

Thus ps^1 is a bialgebra map, and hence also a Hopf map, by Proposition 1.35(c).

For assertion (iii), first assume m lies in $\{0,1,2,\ldots\}$. Since $\zeta_Q(f) = f(1,0,0,\ldots)$, one has

$$\begin{aligned}
\zeta_Q^{\star m}(f) &= \zeta_Q^{\otimes m} \circ \Delta^{(m-1)} f(\mathbf{x}) = \zeta_Q^{\otimes m}\left(f(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}) \right) \\
&= \left[f(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}) \right]_{\substack{x_1^{(1)} = x_1^{(2)} = \cdots = x_1^{(m)} = 1, \\ x_2^{(j)} = x_3^{(j)} = \cdots = 0 \text{ for all } j}} \\
&= f(1,0,0,\ldots,1,0,0,\ldots,\cdots,1,0,0,\ldots) = f(\underbrace{1,1,\ldots,1}_{m \text{ ones}},0,0,\ldots) = \mathrm{ps}^1(f)(m)
\end{aligned}$$

But then Proposition 1.35(a) also implies

$$\begin{aligned}
\zeta_Q^{\star(-m)}(f) &= \left(\zeta_Q^{\star(-1)}\right)^{\star m}(f) = (\zeta_Q \circ S)^{\star m}(f) = \zeta_Q^{\star m}(S(f)) \\
&= \mathrm{ps}^1(S(f))(m) = S(\mathrm{ps}^1(f))(m) = \mathrm{ps}^1(f)(-m).
\end{aligned}$$

For assertion (iv), note that

$$\zeta^{\star m}(a) = (\zeta_Q \circ \Psi)^{\star m}(a) = (\zeta_Q^{\star m})(\Psi(a)) = \mathrm{ps}^1(\Psi(a))(m).$$

where the three equalities come from (6.1), Proposition 1.35(a), and assertion (iii) above, respectively. ☐

Remark 6.8. Aguiar, Bergeron and Sottile give a very cute (third) proof of the QSym antipode formula Theorem 5.9, via Theorem 6.3, in [4, Example 4.8]. They apply Theorem 6.3 to the *coopposite coalgebra* QSym^{cop} and its character $\zeta_Q^{\star^{-1}}$. One can show that the map $\mathrm{QSym} \xrightarrow{\Psi} \mathrm{QSym}^{cop}$ induced by $\zeta_Q^{\star^{-1}}$ is $\Psi = S$, the antipode of QSym, because S is a coalgebra anti-endomorphism satisfying $\zeta_Q^{\star^{-1}} = \zeta \circ S$. They then use the formula (6.2) for $\Psi = S$ (together with the polynomiality Proposition 6.7) to derive Theorem 5.9.

6.2. **Example: Ehrenborg's quasisymmetric function of a ranked poset.** Here we consider incidence algebras, coalgebras and Hopf algebras generally, and then particularize to the case of graded posets, to recover Ehrenborg's interesting quasisymmetric function invariant via Theorem 6.3.

6.2.1. *Incidence algebras, coalgebras, Hopf algebras.*

Definition 6.9. Given a family $\mathcal{P}$ of finite partially ordered sets P, let $\mathbf{k}[\mathcal{P}]$ denote the free $\mathbf{k}$-module whose basis consists of symbols $[P]$ corresponding to isomorphism classes of posets P in $\mathcal{P}$.

We will assume throughout that each P in $\mathcal{P}$ is *bounded*, that is, it has a unique minimal element $\hat{0} := \hat{0}_P$ and unique maximal element $\hat{1} := \hat{1}_P$. In particular, $P \neq \varnothing$, although it is allowed that $|P| = 1$, so that $\hat{0} = \hat{1}$; denote this isomorphism class of posets with one element by $[o]$.

If $\mathcal{P}$ is closed under taking intervals

$$[x,y] := [x,y]_P := \{z \in P : x \leq_P z \leq_P y\}$$

then one can easily that the following coproduct and counit endow $\mathbf{k}[\mathcal{P}]$ with the structure of a coalgebra, called the *(reduced) incidence coalgebra*:

$$\begin{aligned}
\Delta[P] &:= \sum_{x \in P} [\hat{0}, x] \otimes [x, \hat{1}], \\
\epsilon[P] &:= \begin{cases} 1 & \text{if } |P| = 1 \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

The dual algebra $\mathbf{k}[\mathcal{P}]^o$ is generally called the *(reduced) incidence algebra* for the family $\mathcal{P}$. It contains the important element $\mathbf{k}[\mathcal{P}] \xrightarrow{\zeta} \mathbf{k}$, called the ζ*-function* that takes the value $\zeta[P] = 1$ for all P.

If $\mathcal{P}$ (is not empty and) satisfies the further property of being *hereditary* in the sense that for every P_1, P_2 in $\mathcal{P}$, the *Cartesian product poset* $P_1 \times P_2$ with componentwise partial order is also in $\mathcal{P}$, then one can check that the following product and unit endow $\mathbf{k}[\mathcal{P}]$ with the structure of a (commutative) algebra:

$$\begin{aligned}
[P_1] \cdot [P_2] &:= m([P_1] \otimes [P_2]) := [P_1 \times P_2], \\
1_{\mathbf{k}[\mathcal{P}]} &:= [o].
\end{aligned}$$

Proposition 6.10. *For any hereditary family $\mathcal{P}$ of finite posets, $\mathbf{k}[\mathcal{P}]$ is a bialgebra, and even a Hopf algebra with antipode S given as in Theorem 1.15 (Takeuchi's formula):*

$$S[P] = \sum_{k \geq 0} (-1)^k \sum_{\hat{0}=x_0<\cdots<x_k=\hat{1}} [x_0, x_1] \cdots [x_{k-1}, x_k].$$

Proof. Checking the commutativity of the pentagonal diagram in (1.8) amounts to the fact that, for any $(x_1, x_2) <_{P_1 \times P_2} (y_1, y_2)$, one has a poset isomorphism

$$[(x_1, x_2)\ ,\ (y_1, y_2)]_{P_1 \times P_2} \cong [x_1, y_1]_{P_1} \times [x_2, y_2]_{P_2}.$$

Commutativity of the remaining diagrams in (1.8) is straightforward, and so $\mathbf{k}[\mathcal{P}]$ is a bialgebra. But then Remark 1.34 implies that it is a Hopf algebra, with antipode S as in the theorem, because the map $f := 1_{\mathbf{k}[\mathcal{P}]} - u\epsilon$ (sending the class $[o]$ to 0, and fixing all other $[P]$) is locally $\star$-nilpotent:

$$f^{\star k}[P] = \sum_{\hat{0}=x_0<\cdots<x_k=\hat{1}} [x_0, x_1] \cdots [x_{k-1}, x_k]$$

will vanish due to an empty sum whenever k exceeds the maximum length of a chain in the finite poset P. ☐

It is perhaps worth remarking how this generalizes the Möbius function formula of P. Hall. Note that the zeta function $\mathbf{k}[\mathcal{P}] \xrightarrow{\zeta} \mathbf{k}$ is a *character*, that is, an algebra morphism. Proposition 1.35(a) then tells us that ζ should have a convolutional inverse $\mathbf{k}[\mathcal{P}] \xrightarrow{\mu=\zeta^{\star-1}} \mathbf{k}$, traditionally called the *Möbius function*, with the formula $\mu = \zeta^{\star-1} = \zeta \circ S$. Rewriting this via the antipode formula for S given in Proposition 6.10 yields P. Hall's formula.

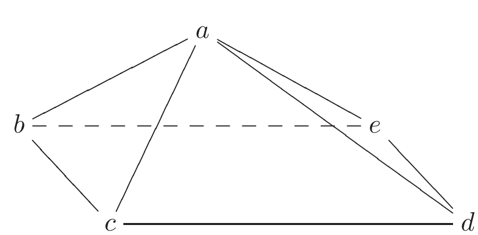
Corollary 6.11. *For a finite bounded poset P, one has*

$$\mu[P] = \sum_{k \geq 0} (-1)^k |\{\textit{chains } \hat{0} = x_0 < \cdots < x_k = \hat{1} \textit{ in } P\}|.$$

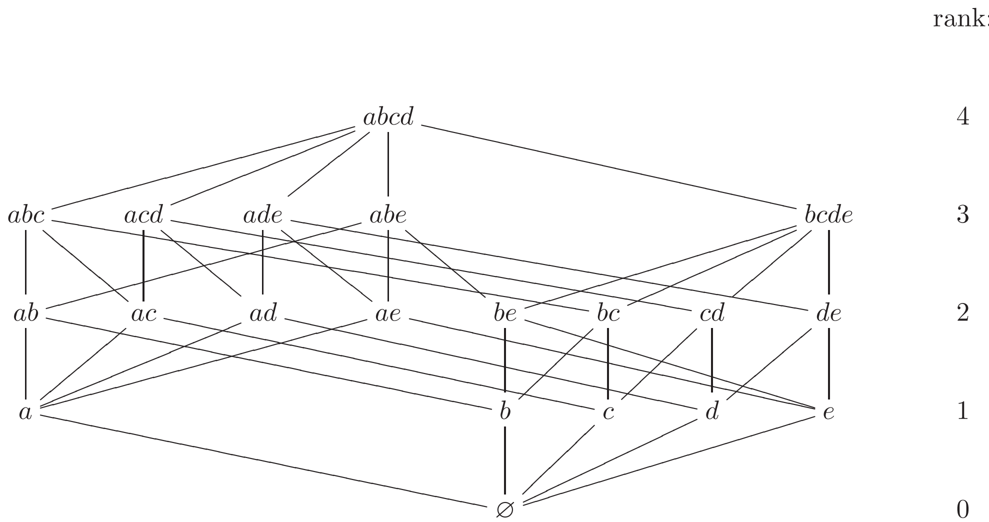
6.2.2. *The incidence Hopf algebras for ranked posets and Ehrenborg's function.*

Definition 6.12. Take $\mathcal{P}$ to be the class of bounded *ranked* finite posets P, that is, those for which all maximal chains from $\hat{0}$ to $\hat{1}$ have the same length $r(P)$. This is a hereditary class, as it implies that any interval is $[x, y]_P$ is also ranked, and the product of two bounded ranked posets is also bounded and ranked. It also uniquely defines a *rank function* $P \xrightarrow{r} \mathbb{N}$ in which $r(\hat{0}) = 0$ and $r(x)$ is the length of any maximal chain from $\hat{0}$ to x.

Example 6.13. Consider a pyramid with apex vertex a over a square base with vertices b, c, d, e:



Ordering its faces by inclusion gives a bounded ranked poset P, where the rank of an element is one more than the dimension of the face it represents:



Definition 6.14. *Ehrenborg's quasisymmetric function* $\Psi[P]$ for a bounded ranked poset P is the image of $[P]$ under the map $\mathbf{k}[\mathcal{P}] \xrightarrow{\Psi} \mathrm{QSym}$ induced by the zeta function $\mathbf{k}[\mathcal{P}] \xrightarrow{\zeta} \mathbf{k}$ as a character, via Theorem 6.3.

The quasisymmetric function $\Psi[P]$ captures several interesting combinatorial invariants of P; see Stanley [72, Chap. 3] for more background on these notions.

Definition 6.15. Let P be a bounded ranked poset P of rank $r(P) := r(\hat{1})$. Define its *rank-generating function*

$$RGF(P,q) := \sum_{p \in P} q^{r(p)},$$

its *characteristic polynomial*

$$\chi(P,q) := \sum_{p \in P} \mu(\hat{0}, p) q^{r(p)}$$

(where $\mu(p,q)$ is shorthand for $\mu([p,q])$), its *zeta polynomial*

$$Z(P,m) = |\{\text{multichains } \hat{0} \leq_P p_1 \leq_P \cdots \leq_P p_{m-1} \leq_P \hat{1}\}| \tag{6.5}$$

$$= \sum_{s=0}^{r(P)-1} \binom{m}{s+1} |\{\text{chains } \hat{0} < p_1 < \cdots < p_s < \hat{1}\}| \tag{6.6}$$

and for a subset $S \subset \{1, 2, \ldots, r-1\}$, its *flag number* f_S, as a component of its *flag f-vector* $(f_S)_{S \subset [r-1]}$

$$f_S = |\{\text{chains } \hat{0} <_P p_1 <_P \cdots <_P p_s <_P \hat{1} \text{ with } \{r(p_1), \ldots, r(p_s)\} = S\}|,$$

as well as the *flag h-vector* entry h_T given by $f_S = \sum_{T \subset S} h_T$, or by inclusion-exclusion, $h_S = \sum_{T \subset S} (-1)^{|S \setminus T|} f_T$.

Example 6.16. For the poset P in Example 6.13, one has $RGF(P,q) = 1 + 5q + 8q^2 + 5q^3 + q^4$. Since P is the poset of faces of a polytope, the Möbius function values for its intervals are easily predicted: $\mu(x,y) = (-1)^{r[x,y]}$, that is, P is an *Eulerian ranked poset*; see Stanley [72, §3.16]. Hence its characteristic polynomial is trivially related to the rank generating function, sending $q \mapsto -q$, that is,

$$\chi(P,q) = RGF(P,-q) = 1 - 5q + 8q^2 - 5q^3 + q^4.$$

Its flag f-vector and h-vector entries are given in the following table.

S	f_S		h_S
$\varnothing$	1		1
$\{1\}$	5	$5-1=$	4
$\{2\}$	8	$8-1=$	7
$\{3\}$	5	$5-1=$	4
$\{1,2\}$	16	$16-(5+8)+1=$	4
$\{1,3\}$	16	$16-(5+5)+1=$	7
$\{2,3\}$	16	$16-(5+8)+1=$	4
$\{1,2,3\}$	32	$32-(16+16+16)+(5+8+5)-1=$	1

and using (6.6), its zeta polynomial is

$$Z(P,m)=1\binom{m}{1}+(5+8+5)\binom{m}{2}+(16+16+16)\binom{m}{3}+32\binom{m}{4}=\frac{m^2(2m-1)(2m+1)}{3}.$$

Theorem 6.17. *Ehrenborg's quasisymmetric function $\Psi[P]$ for a bounded ranked poset P encodes*

(i) *the flag f-vector entries f_S and flag h-vector entries h_S as its M_α and L_α expansion coefficients*[24] *:*

$$\Psi[P]=\sum_\alpha f_{D(\alpha)}(P)\ M_\alpha=\sum_\alpha h_{D(\alpha)}(P)\ L_\alpha,$$

(ii) *the zeta polynomial as the specialization from Definition 6.6*

$$Z(P,m)=\mathrm{ps}^1(\Psi[P])(m)=[\Psi[P]]_{\substack{x_1=x_2=\cdots=x_m=1,\\ x_{m+1}=x_{m+2}=\cdots=0}}, \textit{and}$$

(iii) *the rank-generating function as the specialization*

$$RGF(P,q)=[\Psi[P]]_{\substack{x_1=q,x_2=1,\\ x_3=x_4=\cdots=0}},$$

(iv) *the characteristic polynomial as the convolution*

$$\chi(P,q)=((\psi_q\circ S)\star\zeta_Q)\circ\Psi[P]$$

where $\mathrm{QSym}\xrightarrow{\psi_q}\mathbf{k}[q]$ *maps* $f(\mathbf{x})\longmapsto f(q,0,0,\ldots)$.

Proof. In assertion (i), the expansion $\Psi[P]=\sum_\alpha f_{D(\alpha)}(P)\,M_\alpha$ is (6.2), since $\zeta_\alpha[P]=f_{D(\alpha)}(P)$. The L_α expansion follows by inclusion-exclusion, as $L_\alpha=\sum_{\beta:D(\beta)\supset D(\alpha)}M_\beta$ and $f_S(P)=\sum_{T\subset S}h_T$.

Assertion (ii) is immediate from Proposition 6.7(iv), since $Z(P,m)=\zeta^{\star m}[P]$.

Assertion (iii) can be deduced from assertion (i), but it is perhaps more fun and in the spirit of things to proceed as follows. Note that $\psi_q(M_\alpha)=q^n$ for $\alpha=(n)$, and $\psi_q(M_\alpha)$ vanishes for all other $\alpha\neq(n)$ in Comp_n. Hence for a bounded ranked poset P one has

$$(\psi_q\circ\Psi)[P]=q^{r(P)}. \tag{6.7}$$

Consequently, using (1.16) one can compute

$$\begin{aligned}RGF(P,q)&=\sum_{p\in P}q^{r(p)}\cdot 1=\sum_{p\in P}q^{r([\hat0,p])}\cdot\zeta[p,\hat1]\overset{(6.7),(6.1)}{=}\sum_{p\in P}(\psi_q\circ\Psi)[\hat0,p]\cdot(\zeta_Q\circ\Psi)[p,\hat1]\\&\overset{(1.16)}{=}(\psi_q\star\zeta_Q)(\Psi[P])=(\psi_q\otimes\zeta_Q)\,(\Delta\Psi[P])\\&=[\Psi[P](\mathbf{x},\mathbf{y})]_{\substack{x_1=q,x_2=x_3=\cdots=0\\ y_1=1,y_2=y_3=\cdots=0}}=[\Psi[P](\mathbf{x})]_{\substack{x_1=q,x_2=1,\\ x_3=x_4=\cdots=0}}\end{aligned}$$

Similarly, for assertion (iv) first note that Proposition 6.10 and Corollary 6.11 let one calculate that

$$\begin{aligned}(\psi_q\circ\Psi\circ S)[P]&=\sum_k(-1)^k\sum_{\hat0=x_0<\cdots<x_k=\hat1}(\psi_q\circ\Psi)([x_0,x_1])\cdots(\psi_q\circ\Psi)([x_{k-1},x_k])\\&\overset{(6.7)}{=}\sum_k(-1)^k\sum_{\hat0=x_0<\cdots<x_k=\hat1}q^{r(P)}=\mu(\hat0,\hat1)q^{r(P)}.\end{aligned}$$

[24] In fact, Ehrenborg *defined* $\Psi[P]$ in [24, Defn. 4.1] via this M_α expansion, and then showed that it gave a Hopf morphism.

This is used in the penultimate equality here:

$$((\psi_q \circ S) \star \zeta_Q) \circ \Psi[P] \overset{(1.16)}{=} ((\psi_q \circ S \circ \Psi) \star (\zeta_Q \circ \Psi))[P] = ((\psi_q \circ \Psi \circ S) \star \zeta)[P]$$
$$= \sum_{p \in P} (\psi_q \circ \Psi \circ S)[\hat{0}, p] \cdot \zeta[p, \hat{1}] = \sum_{p \in P} \mu[\hat{0}, p] q^{r(p)} = \chi(P, q).$$

□

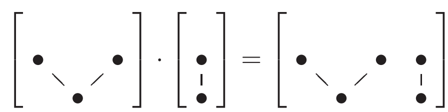
6.3. **Example: Stanley's chromatic symmetric function of a graph.** We introduce the *chromatic Hopf algebra of graphs* and an associated character ζ so that the map Ψ from Theorem 6.3 sends a graph G to Stanley's *chromatic symmetric function* of G. Then principal specialization ps^1 sends this to the *chromatic polynomial* of the graph.

6.3.1. *The chromatic Hopf algebra of graphs.*

Definition 6.18. The *chromatic Hopf algebra* (see Schmitt [65, §3.2]) $\mathcal{G}$ is a free $\mathbf{k}$-module whose $\mathbf{k}$-basis elements $[G]$ are indexed by isomorphism classes of (finite) simple graphs $G = (V, E)$. Define for $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$ the multiplication

$$[G_1] \cdot [G_2] := [G_1 \sqcup G_2]$$

where $[G_1 \sqcup G_2]$ denote the isomorphism class of the disjoint union, on vertex set $V = V_1 \sqcup V_2$ which is a disjoint union of copies of their vertex sets V_1, V_2, with edge set $E = E_1 \sqcup E_2$. For example,



Thus the class $[\varnothing]$ of the empty graph $\varnothing$ having $V = \varnothing, E = \varnothing$ is a unit element.

Given a subset $V' \subset V$, the *subgraph induced on vertex set* V' is defined as the graph $G|_{V'} := (V', E')$ with edge set $E' = \{e \in E : e = \{v_1, v_2\} \subset V'\}$. This lets one define a comultiplication

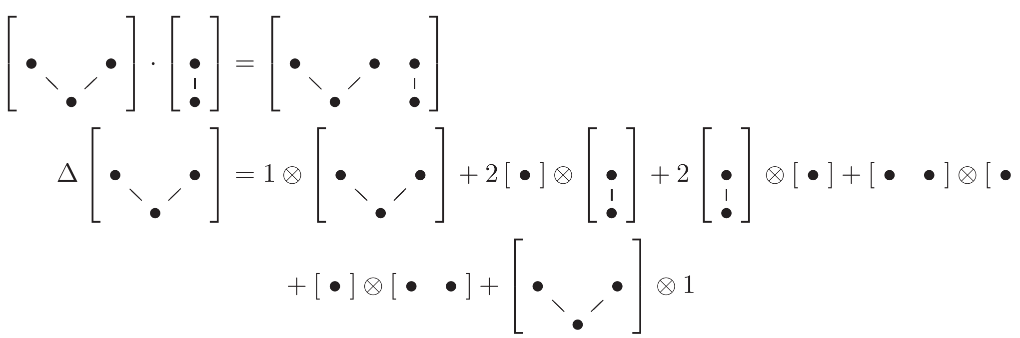
$$\Delta[G] := \sum_{(V_1, V_2): V_1 \sqcup V_2 = V} [G|_{V_1}] \otimes [G|_{V_2}].$$

Define a counit

$$\epsilon[G] := \begin{cases} 1 & \text{if } G = \varnothing \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 6.19. *The above maps endow $\mathcal{G}$ with the structure of a graded connected finite type Hopf algebra over $\mathbf{k}$, which is both commutative and cocommutative.*

Example 6.20. Here are some examples of these structure maps:



Proof. The associativity of the multiplication and comultiplication should be clear as

$$m^{(2)}([G_1] \otimes [G_2] \otimes [G_3]) = [G_1 \sqcup G_2 \sqcup G_3]$$
$$\Delta^{(2)}[G] = \sum_{\substack{(V_1, V_2, V_3): \\ V = V_1 \sqcup V_2 \sqcup V_3}} [G|_{V_1}] \otimes [G|_{V_2}] \otimes [G|_{V_3}].$$

Checking the unit and counit conditions are straightforward. Commutativity of the pentagonal bialgebra diagram in (1.8) comes down to check that, given graphs G_1, G_2 on disjoint vertex sets V_1, V_2 , when one applies to $[G_1] \otimes [G_2]$ either the composite $\Delta \circ m$ or the composite $(m \otimes m) \circ (1 \otimes T \otimes 1) \circ (\Delta \otimes \Delta)$, the result is the same:

$$\sum_{\substack{(V_{11}, V_{12}, V_{21}, V_{22}): \\ V_1 = V_{11} \sqcup V_{12} \\ V_2 = V_{21} \sqcup V_{22}}} [G_1|_{V_{11}} \sqcup G_2|_{V_{21}}] \otimes [G_1|_{V_{12}} \sqcup G_2|_{V_{22}}].$$

Letting $\mathcal{G}_n$ be the $\mathbf{k}$-span of $[G]$ having n vertices makes $\mathcal{G}$ a bialgebra which is graded, connected, and of finite type, and hence also a Hopf algebra by Proposition 1.30. Cocommutativity should be clear, and commutativity follows from the graph isomorphism $G_1 \sqcup G_2 \cong G_2 \sqcup G_1$. $\square$

Remark 6.21. Humpert and Martin [37, Theorem 3.1] gave the following expansion for the antipode in the chromatic Hopf algebra, containing fewer terms than Takeuchi's general formula (1.15): given a graph $G = (V, E)$, one has

$$S[G] = \sum_F (-1)^{|V| - \mathrm{rank}(F)} \operatorname{acyc}(G/F) [G_{V,F}].$$

Here F runs over all subsets of edges that form *flats* in the graphic matroid for G, meaning that if $e = \{v, v'\}$ is an edge in E for which one has a path of edges in F connecting v to v', then e also lies in F. Here G/F denotes the quotient graph in which all of the edges of F have been *contracted*, while $\operatorname{acyc}(G/F)$ denotes its number of *acyclic orientations*, and $G_{V,F} := (V, F)$ as a simple graph.

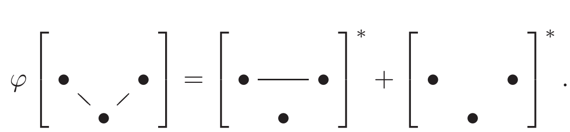
It turns out that the chromatic Hopf algebra $\mathcal{G}$ is self-dual. In the dual Hopf algebra $\mathcal{G}^o$, let $\{[G]^*\}$ denote the dual basis elements, so that $([H]^*, [G]) = \delta_{[H],[G]}$. To describe the structure maps in $\mathcal{G}^o$ explicitly, for graphs H, H_1, H_2 one has

$$\begin{aligned} \Delta[H]^* &= \sum_{\substack{(V_1, V_2): \\ V = V_1 \sqcup V_2 \\ H = H|_{V_1} \sqcup H|_{V_2}}} [H|_{V_1}]^* \otimes [H|_{V_2}]^* \\ [H_1]^*[H_2]^* &= \sum_{\substack{H = (V_1 \sqcup V_2, E) \\ H|_{V_1} = H_1 \\ H|_{V_2} = H_2}} [H]^*. \end{aligned} \tag{6.8}$$

Proposition 6.22. *One has a Hopf isomorphism* $\mathcal{G} \xrightarrow{\varphi} \mathcal{G}^o$ *defined by*

$$[G] \longmapsto \sum_{\substack{H = (V, E'): \\ E' \cap E = \varnothing}} [H]^*.$$

For example, this isomorphism maps



Proof. First note that φ is a $\mathbf{k}$-module isomorphism via triangularity: one has $H = (V, E')$ with $E' \cap E = \varnothing$ if and only if H is an edge subgraph of the *complementary graph* $\overline{G}$ to G on the same vertex set V.

One can then check that for graphs G_1, G_2 on vertex sets V_1, V_2, the fact that $\varphi([G_1][G_2]) = \varphi[G_1]\varphi[G_2]$ amounts, using (6.8), to both being a sum of $[H]^*$ over graphs H on $V_1 \sqcup V_2$ that share no edges with $G_1 \sqcup G_2$.

Similarly, one can check that the fact that $\Delta\varphi[G] = (\varphi \otimes \varphi)(\Delta[G])$ amounts, using (6.8), to both being a sum of $[H_1]^* \otimes [H_2]^*$ over triples (H, H_1, H_2) where H is a graph on the same vertex set V as G but sharing no edges with G, and with $H = H_1 \sqcup H_2$. $\square$

6.3.2. *Stanley's chromatic symmetric function of a graph.*

Definition 6.23. *Stanley's chromatic symmetric function* $\Psi[G]$ for a simple graph $G = (V, E)$ is the image of $[G]$ under the map $\mathcal{G} \xrightarrow{\Psi} \mathrm{QSym}$ induced via Theorem 6.3 from the *edge-free character* $\mathcal{G} \xrightarrow{\zeta} \mathbf{k}$

$$\zeta[G] = \begin{cases} 1 & \text{if } G \text{ has no edges, that is, } G \text{ is an independent/stable set of vertices,} \\ 0 & \text{otherwise.} \end{cases} \tag{6.9}$$

Note that, because $\mathcal{G}$ is cocommutative, $\Psi[G]$ is symmetric and not just quasisymmetric; see Remark 6.4.

Recall that for a graph $G = (V, E)$, a (vertex-)coloring $f : V \to \{1, 2, \ldots\}$ is called *proper* if no edge $e = \{v, v'\}$ in E has $f(v) = f(v')$.

Proposition 6.24. *For a graph $G = (V, E)$, the symmetric function $\Psi[G]$ has the expansion* [25]

$$\Psi[G] = \sum_{\substack{\textit{proper colorings} \\ f:V\to\{1,2,\ldots\}}} \mathbf{x}_f$$

where $\mathbf{x}_f := \prod_{v\in V} x_{f(v)}$. In particular, its specialization from Proposition 6.6 gives the chromatic polynomial of G:

$$\mathrm{ps}^1 \Psi[G](m) = \chi_G(m) = |\{\textit{proper colorings } f : V \to \{1, 2, \ldots, m\}\}|$$

Proof. The iterated coproduct $\mathcal{G} \xrightarrow{\Delta^{(\ell-1)}} \mathcal{G}^{\otimes \ell}$ sends

$$[G] \longmapsto \sum_{\substack{(V_1,\ldots,V_\ell): \\ V=V_1\sqcup\cdots\sqcup V_\ell}} [G|_{V_1}] \otimes \cdots \otimes [G|_{V_\ell}]$$

and the map $\zeta^{\otimes \ell}$ sends the element on the right to 1 or 0, depending upon whether each $V_i \subset V$ is a stable set or not, that is, whether the assignment of color i to the vertices in V_i gives a proper coloring. Thus formula (6.2) shows that the coefficient ζ_α of $x_1^{\alpha_1} \cdots x_\ell^{\alpha_\ell}$ in $\Psi[G]$ counts the proper colorings f in which $|f^{-1}(i)| = \alpha_i$ for each i. ☐

Example 6.25. For the complete graph K_n on n vertices, one has

$$\begin{aligned} \Psi[K_n] &= n! e_n \\ \mathrm{ps}^1(\Psi[K_n])(m) &= n! e_n(\underbrace{1, 1, \ldots, 1}_{m \text{ ones}}) = n! \binom{m}{n} \\ &= m(m-1)\cdots(m-(n-1)) = \chi(K_n, m). \end{aligned}$$

In particular, the single vertex graph K_1 has $\Psi[K_1] = e_1$, and since the Hopf morphism Ψ is in particular an algebra morphism, a graph $K_1^{\sqcup n}$ having n isolated vertices and no edges will have $\Psi[K_1^{\sqcup n}] = e_1^n$.

As a slightly more interesting example, the graph P_3 which is a path having three vertices and two edges will have

$$\Psi[P_3] = m_{(2,1)} + 6m_{(1,1,1)} = e_2 e_1 + 3e_3$$

One might wonder, based on the previous examples, when $\Psi[G]$ is *e-positive*, that is, when does its unique expansion in the $\{e_\lambda\}$ basis for Λ have nonnegative coefficients? This is an even stronger assertion than *s-positivity*, that is, having nonnegative coefficients for the expansion in terms of Schur functions $\{s_\lambda\}$, since each e_λ is s-positive. This weaker property fails, starting with the *claw graph* $K_{3,1}$, which has

$$\Psi[K_{3,1}] = s_{(3,1)} - s_{(2,2)} + 5s_{(2,1,1)} + 8s_{(1,1,1,1)}.$$

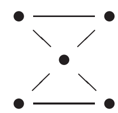
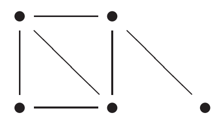
On the other hand, a result of Gasharov [26] shows that one at least has s-positivity for $\Psi[\mathrm{inc}(P)]$ where $\mathrm{inc}(P)$ is the *incomparability graph* of a poset which is $(\mathbf{3}+\mathbf{1})$-free; we refer the reader to Stanley [71, §5] for a discussion of the following conjecture, which remains open[26]:

[25] In fact, Stanley *defined* $\Psi[G]$ in [71, Defn. 2.1] via this expansion.

[26] A recent refinement for incomparability graphs of posets which are both $(\mathbf{3}+\mathbf{1})$- and $(\mathbf{2}+\mathbf{2})$-free, also known as *unit interval orders* is discussed by Shareshian and Wachs [67].

Conjecture 6.26. *For any* $(\mathbf{3}+\mathbf{1})$*-free poset* P*, the incomparability graph* $\operatorname{inc}(P)$ *has* $\Psi[\operatorname{inc}(P)]$ *an* e*-positive symmetric function.*

Here is another question about $\Psi[G]$: how well does it distinguish nonisomorphic graphs? Stanley gave this example of two graphs G_1, G_2 having $\Psi[G_1] = \Psi[G_2]$:

$G_1 =$  $G_2 =$ 

At least $\Psi[G]$ appears to do better at distinguishing *trees*, much better than its specialization, the chromatic polynomial $\chi(G, m)$, which takes the same value $m(m-1)^{n-1}$ on all trees with n vertices.

Question 6.27. Does the chromatic symmetric function distinguish trees?

It has been checked that the answer is affirmative for trees on 23 vertices or less. There are also interesting partial results on this question by Martin, Morin and Wagner [57].

We close this section with a few other properties of $\Psi[G]$ proven by Stanley which follow easily from the theory we have developed. For example, his work makes no explicit mention of the chromatic Hopf algebra $\mathcal{G}$, and the fact that Ψ is a Hopf morphism (although he certainly notes the trivial algebra morphism property $\Psi[G_1 \sqcup G_2]) = \Psi[G_1]\Psi[G_2]$). One property he proves is implicitly related to Ψ as a coalgebra morphism: he considers the effect on Ψ of the operator $\frac{\partial}{\partial p_1} : \Lambda_{\mathbb{Q}} \longrightarrow \Lambda_{\mathbb{Q}}$ which acts by first expressing a symmetric function $f = f(p_1, p_2, \ldots)$ as a polynomial in the power sums $\{p_n\}$, and then applies $\frac{\partial}{\partial p_1}$. It is not hard to see that $\frac{\partial}{\partial p_1}$ is the same as the skewing operator $s_{(1)}^{\perp} = p_1^{\perp}$: both act as derivations on $\Lambda_{\mathbb{Q}} = \mathbb{Q}[p_1, p_2, \ldots]$, and agree in their effect on each p_n, in that both send $p_1 \mapsto 1$, and both annihilate $p_2, p_3, \ldots$.

Proposition 6.28. *(Stanley* [71, Cor. 2.12(a)]*) For any graph* $G = (V, E)$*, one has*

$$\frac{\partial}{\partial p_1}\Psi[G] = \sum_{v \in V} \Psi[G|_{V \setminus v}].$$

Proof. One first computes

$$\Delta\Psi[G] = (\Psi \otimes \Psi)\Delta[G] = \sum_{\substack{(V_1, V_2):\\ V = V_1 \sqcup V_2}} \Psi[G|_{V_1}] \otimes \Psi[G|_{V_2}].$$

Since degree considerations force $(s_{(1)}, \Psi[G|_{V_1}]) = 0$ unless $|V_1| = 1$, in which case $\Psi[G|_{V_1}] = s_{(1)}$, one has

$$\frac{\partial}{\partial p_1}\Psi[G] = s_{(1)}^{\perp}\Psi[G] = \sum_{\substack{(V_1, V_2):\\ V = V_1 \sqcup V_2}} (s_{(1)}, \Psi[G|_{V_1}]) \cdot \Psi[G|_{V_2}] = \sum_{v \in V} \Psi[G|_{V \setminus v}].$$

□

Definition 6.29. Given a graph $G = (V, E)$, an acyclic orientation Ω of the edges E (that is, an orientation of each edge such that the resulting directed graph has no cycles), and a vertex-coloring $f : V \to \{1, 2, \ldots\}$, say that the pair (Ω, f) are *weakly compatible* if whenever Ω orients an edge $\{v, v'\}$ in E as $v \to v'$, one has $f(v) \leq f(v')$. Note that a *proper* vertex-coloring f of a graph $G = (V, E)$ is weakly compatible with a unique acyclic orientation Ω.

Proposition 6.30. *(Stanley* [71, Prop. 4.1, Thm. 4.2]*) The involution* ω *of* Λ *sends* $\Psi[G]$ *to* $\omega(\Psi[G]) = \sum_{(\Omega, f)} \mathbf{x}_f$ *in which the sum runs over weakly compatible pairs* (Ω, f) *of an acyclic orientation* Ω *and vertex-coloring* f*.*

Furthermore, the chromatic polynomial $\chi_G(m)$ *has the property that* $(-1)^{|V|}\chi(G, -m)$ *counts all such weakly compatible pairs* (Ω, f) *in which* $f : V \to \{1, 2, \ldots, m\}$ *is a vertex-*m*-coloring.*

Proof. As observed above, a proper coloring f is weakly compatible with a unique acyclic orientation Ω of G. Denote by P_Ω the poset on V which is the transitive closure of Ω, endowed with a *strict labelling* by integers, that is, $i <_P j$ implies $i >_{\mathbb{Z}} j$. Then proper colorings f that induce Ω are the same as P_Ω-partitions, so that

$$\Psi[G] = \sum_{\Omega} F_{P_\Omega}(\mathbf{x}). \tag{6.10}$$

Applying the antipode S and using Corollary 5.22 gives

$$\omega\left(\Psi[G]\right) = (-1)^{|V|} S\left(\Psi[G]\right) = \sum_{\Omega} F_{P_\Omega^{\mathrm{opp}}}(\mathbf{x}) = \sum_{(\Omega,f)} \mathbf{x}_f$$

where in the last line one sums over weakly compatible pairs as in the proposition. The last equality comes from the fact that since each P_Ω has been given a strict labelling, P_Ω^{opp} acquires a *weak (or natural) labelling*, that is $i <_{P_\Omega^{\mathrm{opp}}} j$ implies $i <_{\mathbb{Z}} j$.

The last assertion follows from Proposition 6.7(iii). □

Remark 6.31. The interpretation of $\chi(G,-m)$ in Proposition 6.30 is a much older result of Stanley [70]. The special case interpreting $\chi(G,-1)$ as $(-1)^{|V|}$ times the number of acyclic orientations of G has sometimes been called Stanley's *(-1)-color theorem*. It also follows (via Proposition 6.7) from Humpert and Martin's antipode formula for $\mathcal{G}$ discussed in Remark 6.21: taking ζ to be the character of $\mathcal{G}$ given in (6.9),

$$\chi(G,-1) = \zeta^{*(-1)}[G] = \zeta(S[G]) = \sum_{F} (-1)^{|V|-\mathrm{rank}(F)} \operatorname{acyc}(G/F) S[G_{V,F}] = (-1)^{|V|} \operatorname{acyc}(G)$$

where the last equality uses the vanishing of ζ on graphs that have edges, so only the $F = \varnothing$ term survives.

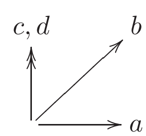
6.4. **Example: The quasisymmetric function of a matroid.** We introduce the *matroid-minor Hopf algebra* of Schmitt [64], and studied extensively by Crapo and Schmitt [16, 17, 18]. A very simple character ζ on this Hopf algebra will then give rise, via the map Ψ from Theorem 6.3, to the quasisymmetric function invariant of matroids from the work of Billera, Jia and the author [11].

6.4.1. *The matroid-minor Hopf algebra.* We begin by reviewing some notions from matroid theory; see Oxley [58] for background, undefined terms and unproven facts.

Definition 6.32. A *matroid* M of rank r on a (finite) ground set E is specified by a nonempty collection $\mathcal{B}(M)$ of r-element subsets of E with the following *exchange property*:

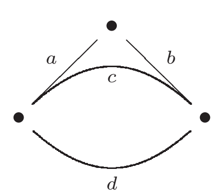
For any B, B' in $\mathcal{B}(M)$ and b in B, there exists b' in B' with $(B \setminus \{b\}) \cup \{b'\}$ in $\mathcal{B}(M)$.

Example 6.33. A matroid M is *represented* by a collection of vectors $E = \{e_1, \dots, e_n\}$ in a vector space if $\mathcal{B}(M)$ is the collection of subsets $B = \{e_{i_1}, \dots, e_{i_r}\}$ having the property that B forms a basis for the span of all of the vectors in E. For example, if $E = \{a, b, c, d\}$ are the four vectors $a = (1,0), b = (1,1), c = (0,1) = d$ in $\mathbb{R}^2$ depicted here



then $\mathcal{B}(M) = \{\{a,b\},\{a,c\},\{a,d\},\{b,c\},\{b,d\}\}$.

Example 6.34. A special case of matroids M represented by vectors are *graphic matroids*, coming from a graph $G = (V,E)$, with parallel edges and self-loops allowed. One represents these by vectors in $\mathbb{R}^V$ with standard basis $\{\epsilon_v\}_{v \in V}$ by associating to the edge $e = \{v, v'\}$ the vector $\epsilon_v - \epsilon_{v'}$. One can check (or see [58, §1.2]) that the bases B in $\mathcal{B}(M)$ correspond to the edge sets of *spanning forests* for G, that is, edge sets which are acyclic and contain one spanning tree for each connected component of G. For example, the graph $G = (V,E)$ shown below has the same matroid $\mathcal{B}(M)$ as the one represented by the vectors in Example 6.33:



whose spanning trees are the edge sets $\mathcal{B}(M) = \{\{a,b\},\{a,c\},\{a,d\},\{b,c\},\{b,d\}\}$.

To define the matroid-minor Hopf algebra one needs the basic matroid operations of *deletion* and *contraction*. These model the operations of deleting or contracting an edge in a graph. For configurations of vectors they model the deletion of a vector, or the passage to images in the quotient space modulo the span of a vector.

Definition 6.35. Given a matroid M of rank r and an element e of its ground set E, say that e is *loop* (resp. *coloop*) of M if e lies in no basis (resp. every basis) B in $\mathcal{B}(M)$. If e is not a coloop, the *deletion* $M \setminus e$ is a matroid of rank r on ground set $E \setminus \{e\}$ having bases

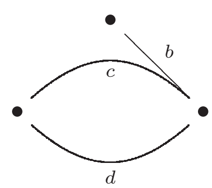
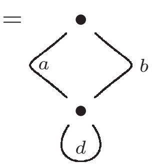
$$\mathcal{B}(M \setminus e) := \{B \in \mathcal{B}(M) : e \notin B\}. \tag{6.11}$$

If e is not a loop, the *contraction* M/e is a matroid of rank r on ground set $E \setminus \{e\}$ having bases

$$\mathcal{B}(M/e) := \{B \setminus \{e\} : e \in B \in \mathcal{B}(M)\}. \tag{6.12}$$

When e is a loop of M, then M/e has rank r instead of $r-1$ and one defines its bases as in (6.11) rather than (6.12); similarly, if e is a coloop of M then $M \setminus e$ has rank $r-1$ instead of r and one defines its bases as in (6.12) rather than (6.11).

Example 6.36. Starting with the graph G and its graphic matroid M from Example 6.34, the deletion $M \setminus a$ and contraction M/c correspond to the graphs $G \setminus a$ and G/c shown here:

$G \setminus a =$  $G/c =$ 

One has

- $\mathcal{B}(M \setminus a) = \{\{b, c\}, \{b, d\}\}$, so that b has become a coloop in $M \setminus a$, and
- $\mathcal{B}(M/c) = \{\{a\}, \{b\}\}$, so that d has become a loop in M/c.

Definition 6.37. Deletions and contractions commute with each other, leading to well-defined operations for subsets $A \subset E$ of the

- *restriction* $M|_A$ on ground set A, obtained by deleting all e of $E \setminus A$ in any order, and
- *quotient/contraction* M/A on ground set $E \setminus A$, obtained by contracting all e in A in any order.

We will also need the *direct sum* $M_1 \oplus M_2$, whose ground set $E = E_1 \sqcup E_2$ is the disjoint union of a copy of the ground sets E_1, E_2 for M_1, M_2, and having bases

$$\mathcal{B}(M_1 \oplus M_2) := \{B_1 \sqcup B_2 : B_i \in \mathcal{B}(M_i) \text{ for } i = 1, 2\}.$$

Lastly, say that two matroids M_1, M_2 are *isomorphic* if there is a bijection of their ground sets $E_1 \xrightarrow{\varphi} E_2$ having the property that $\varphi\mathcal{B}(M_1) = \mathcal{B}(M_2)$.

Now one can define the matroid-minor Hopf algebra, originally introduced by Schmitt [64, §15], and studied further by Crapo and Schmitt [16, 17, 18].

Definition 6.38. Let $\mathcal{M}$ have $\mathbf{k}$-basis elements $[M]$ indexed by isomorphism classes of matroids. Define the multiplication via

$$[M_1] \cdot [M_2] := [M_1 \oplus M_2]$$

so that the class $[\varnothing]$ of the *empty matroid* $\varnothing$ having empty ground set gives a unit. Define the comultiplication for M a matroid on ground set E via

$$\Delta[M] := \sum_{A \subset E} [M|_A] \otimes [M/A],$$

and a counit

$$\epsilon[M] := \begin{cases} 1 & \text{if } M = \varnothing \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 6.39. *The above maps endow $\mathcal{M}$ with the structure of a graded connected finite type Hopf algebra over $\mathbf{k}$, which is commutative.*

Proof. Checking the unit and counit conditions are straightforward. Associativity and commutativity of the multiplication follow because the direct sum operation $\oplus$ for matroids is associative, and commutative up to isomorphism. Coassociativity follows because for a matroid M on ground set E, one has this equality between the two candidates for $\Delta^{(2)}[M]$

$$\begin{aligned} &\sum_{\varnothing \subseteq A_1 \subseteq A_2 \subseteq E} [M|_{A_1}] \otimes [(M|_{A_2})/A_1] \otimes [M/A_2] \\ =& \sum_{\varnothing \subseteq A_1 \subseteq A_2 \subseteq E} [M|_{A_1}] \otimes [(M/A_1)|_{A_2 \setminus A_1}] \otimes [M/A_2] \end{aligned}$$

due to the matroid isomorphism $(M|_{A_2})/A_1 \cong (M/A_1)|_{A_2 \setminus A_1}$. Commutativity of the bialgebra diagram in (1.8) amounts to the fact that for a pair of matroids M_1, M_2 and subsets A_1, A_2 of their (disjoint) ground sets E_1, E_2, one has isomorphisms

$$\begin{aligned} M_1|_{A_1} \oplus M_2|_{A_2} &\cong (M_1 \oplus M_2)\,|_{A_1 \sqcup A_2}, \\ M_1/A_1 \oplus M_2/A_2 &\cong (M_1 \oplus M_2)\,/(A_1 \sqcup A_2). \end{aligned}$$

Letting $\mathcal{M}_n$ be the $\mathbf{k}$-span of $[M]$ for matroids whose ground set E has cardinality $|E| = n$, one can then easily check that $\mathcal{M}$ becomes a bialgebra which is graded, connected, and of finite type, hence also a Hopf algebra by Proposition 1.30. $\square$

6.4.2. *A quasisymmetric function for matroids.*

Definition 6.40. Define a character $\mathcal{M} \xrightarrow{\zeta} \mathbf{k}$ by

$$\zeta[M] = \begin{cases} 1 & \text{if } M \text{ has only one basis,} \\ 0 & \text{otherwise.} \end{cases}$$

It is easily checked that this is a character, that is, an algebra map $\mathcal{M} \xrightarrow{\zeta} \mathbf{k}$. Note that if M has only one basis, say $\mathcal{B}(M) = \{B\}$, then $B := \text{coloops}(M)$ is the set of coloops of M, and $E \setminus B = \text{loops}(M)$ is the set of loops of M. Equivalently, $M = \bigoplus_{e \in E} M|_{\{e\}}$ is the direct sum of matroids each having one element, each a coloop or loop.

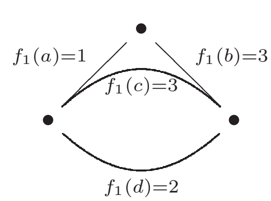
Define $\Psi[M]$ for a matroid M to be the image of $[M]$ under the map $\mathcal{M} \xrightarrow{\Psi} \text{QSym}$ induced via Theorem 6.3 from the above character ζ.

It turns out that $\Psi[M]$ is intimately related with greedy algorithms and finding minimum cost bases. A fundamental property of matroids (and one that characterizes them, in fact; see [58, §1.8]) is that no matter how one assigns costs $f : E \to \mathbb{R}$ to the elements of E, the following *greedy algorithm* (generalizing *Kruskal's algorithm* for finding minimum cost spanning trees) always succeeds in finding one basis B in $\mathcal{B}(M)$ achieving the minimum *total cost* $f(B) := \sum_{b \in B} f(b)$:

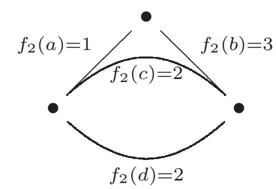
> Start with the empty subset $I_0 = \varnothing$ of E. For $j = 1, 2, \ldots, r$, having already defined the set I_{j-1}, let e be the element of $E \setminus I_{j-1}$ having the lowest cost $f(e)$ among all those for which $I_{j-1} \cup \{e\}$ is *independent*, that is, still a subset of at least one basis B in $\mathcal{B}(M)$. Then define $I_j := I_{j-1} \cup \{e\}$. Repeat this until $j = r$, and $B = I_r$ will be among the bases that achieve the minimum cost.

Definition 6.41. Say that a cost function $f : E \to \{1, 2, \ldots\}$ is *M-generic* if there is a *unique* basis B in $\mathcal{B}(M)$ achieving the minimum cost $f(B)$.

Example 6.42. For the graphic matroid M of Example 6.34, this cost function $f_1 : E \to \{1, 2, \ldots\}$



is M-generic, as it minimizes uniquely on the basis $\{a,d\}$, whereas this cost function $f_2 : E \to \{1,2,\ldots\}$



is *not* M-generic, as it achieves its minimum value on the two bases $\{a,c\},\{a,d\}$.

Proposition 6.43. *For a matroid M on ground set E, one has this expansion*[27]

$$\Psi[M] = \sum_{\substack{M\text{-generic} \\ f:E\to\{1,2,\ldots\}}} \mathbf{x}_f$$

where $\mathbf{x}_f := \prod_{e\in E} x_{f(e)}$. In particular, for $m \geq 0$, its specialization ps^1 from Definition 6.6 has this interpretation:

$$\mathrm{ps}^1\Psi[M](m) = |\{M\text{-generic } f : E \to \{1,2,\ldots,m\}\}|.$$

Proof. The iterated coproduct $\mathcal{M} \overset{\Delta^{(\ell-1)}}{\longrightarrow} \mathcal{M}^{\otimes \ell}$ sends

$$[M] \longmapsto \sum [M|_{A_1}] \otimes [(M|_{A_2})/A_1] \otimes \cdots \otimes [(M|_{A_\ell})/A_{\ell-1}]$$

where the sum is over flags of nested subsets

$$\varnothing = A_0 \subseteq A_1 \subseteq \cdots \subseteq A_{\ell-1} \subseteq A_\ell = E. \tag{6.13}$$

The map $\zeta^{\otimes \ell}$ sends each summand to 1 or 0, depending upon whether each $(M|_{A_j})/A_{j-1}$ has a unique basis or not. Thus formula (6.2) shows that the coefficient ζ_α of $x_{i_1}^{\alpha_1} \cdots x_{i_\ell}^{\alpha_\ell}$ in $\Psi[M]$ counts the flags of subsets in (6.13) for which $|A_j \setminus A_{j-1}| = \alpha_j$ and $(M|_{A_j})/A_{j-1}$ has a unique basis, for each j.

Given a flag as in (6.13), associate the cost function $f : E \to \{1,2,\ldots\}$ whose value on each element of $A_j \setminus A_{j-1}$ is i_j; conversely, given any cost function, say whose distinct values are $i_1 < \ldots < i_\ell$, one associates the flag having $A_j \setminus A_{j-1} = f^{-1}(i_j)$ for each j.

We will prove below, using induction on $s = 0,1,2\ldots,\ell$ the following **claim**: After having completed $\alpha_1 + \alpha_2 + \cdots + \alpha_s$ steps in the greedy algorithm (6.4.2), there is *a unique choice* for the independent set produced thus far, namely

$$I_{\alpha_1+\alpha_2+\cdots+\alpha_s} = \bigsqcup_{j=1}^{s} \mathrm{coloops}((M|_{A_j})/A_{j-1}), \tag{6.14}$$

if and only if each of the matroids $(M|_{A_j})/A_{j-1}$ for $j = 1,2,\ldots,s$ *has a unique basis.*

The case $s = \ell$ in this claim would show what we want, namely that f is M-generic, minimizing uniquely on the basis shown in (6.14) with $s = \ell$, if and only if each $(M|_{A_j})/A_{j-1}$ has a unique basis.

The assertion of the claim is trivially true for $s = 0$. In the inductive step, one may assume that

- the independent set $I_{\alpha_1+\alpha_2+\cdots+\alpha_{s-1}}$ takes the form in (6.14), replacing s by $s-1$,
- it is the unique f-minimizing basis for $M|_{A_{s-1}}$, and
- $(M|_{A_j})/A_{j-1}$ has a unique basis for $j = 1,2,\ldots,s-1$.

Since A_{s-1} exactly consists of all of the elements e of E whose costs $f(e)$ lie in the range $\{i_1,i_2,\ldots,i_{s-1}\}$, in the next α_s steps the algorithm will work in the quotient matroid M/A_{j-1} and attempt to augment $I_{\alpha_1+\alpha_2+\cdots+\alpha_{s-1}}$ using the next-cheapest elements, namely the elements of $A_s \setminus A_{s-1}$, which all have cost f equal to i_s. Thus the algorithm will have no choices about how to do this augmentation if and only if $(M|_{A_s})/A_{s-1}$ has a unique basis, namely its set of coloops, in which case the algorithm will choose to add all of these coloops, giving $I_{\alpha_1+\alpha_2+\cdots+\alpha_s}$ as described in (6.14). This completes the induction.

The last assertion follows from Proposition 6.7. ☐

[27] In fact, this expansion was the original definition of $\Psi[M]$ in [11, Defn. 1.1].

Example 6.44. If M has one basis then every function $f : E \to \{1, 2, \ldots\}$ is M-generic, and

$$\Psi[M] = \sum_{f:E\to\{1,2,\ldots\}} \mathbf{x}_f = (x_1 + x_2 + \cdots)^{|E|} = M_{(1)}^{|E|}.$$

Example 6.45. Let $U_{r,n}$ denote the *uniform matroid* of rank r on n elements E, having $\mathcal{B}(U_{r,n})$ equal to all of the r-element subsets of E.

As $U_{1,2}$ has $E = \{1, 2\}$ and $\mathcal{B} = \{\{1\}, \{2\}\}$, genericity means $f(1) \neq f(2)$, so

$$\Psi[U_{1,2}] = \sum_{\substack{(f(1),f(2)):\\ f(1)\neq f(2)}} x_{f(1)}x_{f(2)} = x_1x_2 + x_2x_1 + x_1x_3 + x_3x_1 + \cdots = 2M_{(1,1)}.$$

Similarly $U_{1,3}$ has $E = \{1, 2, 3\}$ with $\mathcal{B} = \{\{1\}, \{2\}, \{3\}\}$, and genericity means either that $f(1), f(2), f(3)$ are all distinct, or that two of them are the same and the third is smaller. This shows

$$\begin{aligned}\Psi[U_{1,3}] &= 3\sum_{i<j} x_ix_j^2 + 6\sum_{i<j<k} x_ix_jx_k\\ &= 3M_{(1,2)} + 6M_{(1,1,1)}\\ \mathrm{ps}^1\Psi[U_{1,3}](m) &= 3\binom{m}{2} + 6\binom{m}{3} = \frac{m(m-1)(2m-1)}{2}\end{aligned}$$

One can similarly analyze $U_{2,3}$ and check that

$$\begin{aligned}\Psi[U_{2,3}] &= 3M_{(2,1)} + 6M_{(1,1,1)}\\ \mathrm{ps}^1\Psi[U_{2,3}](m) &= 3\binom{m}{2} + 6\binom{m}{3} = \frac{m(m-1)(2m-1)}{2}\end{aligned}$$

These last examples illustrate the behavior of Ψ under the duality operation on matroids.

Definition 6.46. Given a matroid M of rank r on ground set E, its *dual* or *orthogonal matroid* $M^\perp$ is a matroid of rank $|E| - r$ on the same ground set E, having

$$\mathcal{B}(M^\perp) := \{E \setminus B\}_{B\in\mathcal{B}(M)}.$$

Here are a few examples of dual matroids.

Example 6.47. The dual of a uniform matroid is another uniform matroid:

$$U_{r,n}^\perp = U_{n-r,n}.$$

Example 6.48. If M is matroid of rank r represented by collection of vectors $E = \{e_1, \ldots, e_n\}$ in a vector space over some field $\mathbf{k}$, one can find a collection of vectors $\{e_1^\perp, \ldots, e_n^\perp\}$ that represent $M^\perp$ in the following way. Pick a basis for the span of the vectors $\{e_i\}_{i=1}^n$, and create a matrix A in $\mathbf{k}^{r\times n}$ whose columns express the e_i in terms of this basis. Then pick any matrix $A^\perp$ whose row space is the null space of A, and one finds that the columns $\{e_i^\perp\}_{i=1}^n$ of $A^\perp$ represent $M^\perp$. See Oxley [58, §2.2].

Example 6.49. Let $G = (V, E)$ be graph embedded in the plane with edge set E, giving rise to a graphic matroid M on ground set E. Let $G^\perp$ be a planar dual of G, so that, in particular, for each edge e in E, the graph $G^\perp$ has one edge $e^\perp$, crossing e transversely. Then the graphic matroid of $G^\perp$ is $M^\perp$. See Oxley [58, §2.3].

Proposition 6.50. *If* $\Psi[M] = \sum_\alpha c_\alpha M_\alpha$ *then* $\Psi[M^\perp] = \sum_\alpha c_\alpha M_{\mathrm{rev}(\alpha)}$.

Consequently, $\mathrm{ps}^1\Psi[M](m) = \mathrm{ps}^1\Psi[M^\perp](m)$.

Proof. This amounts to showing that for any composition $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, the cardinality of the set of M-generic f having $\mathbf{x}_f = \mathbf{x}^\alpha$ is the same as the cardinality of the set of $M^\perp$-generic $f^\perp$ having $\mathbf{x}_{f^\perp} = \mathbf{x}^{\mathrm{rev}(\alpha)}$. We claim that the map $f \longmapsto f^\perp$ in which $f^\perp(e) = \ell + 1 - f(e)$ gives a bijection between these sets. To see this, note that any basis B of M satisfies

$$f(B) + f(E \setminus B) = \sum_{e\in E} f(e) \tag{6.15}$$

$$f(E \setminus B) + f^\perp(E \setminus B) = (\ell + 1)(|E| - r), \tag{6.16}$$

where r denotes the rank of M. Thus B is f-minimizing if and only if $E \setminus B$ is f-maximizing (by (6.15)) if and only if $E \setminus B$ is $f^{\perp}$-minimizing (by (6.16)). Consequently f is M-generic if and only if $f^{\perp}$ is $M^{\perp}$-generic.

The last assertion follows, for example, from the calculation in Proposition 6.7(i) that $\mathrm{ps}^1(M_\alpha)(m) = \binom{m}{\ell(\alpha)}$ together with the fact that $\ell(\mathrm{rev}(\alpha)) = \ell(\alpha)$. $\square$

Just as (6.10) showed that Stanley's chromatic symmetric function of a graph has an expansion as a sum of P-partition enumerators for certain strictly labelled posets P, the same holds for $\Psi[M]$.

Definition 6.51. Given a matroid M on ground set E, and a basis B in $\mathcal{B}(M)$, define the *base-cobase* poset P_B to have $b < b'$ whenever b lies in B and b' lies in $E \setminus B$ and $(B \setminus \{b\}) \cup \{b'\}$ is in $\mathcal{B}(M)$.

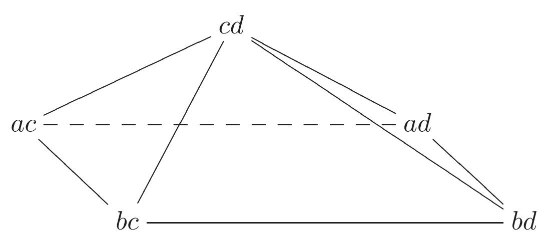
Proposition 6.52. *For any matroid M, one has $\Psi[M] = \sum_{B \in \mathcal{B}(M)} F_{(P_B, \mathrm{strict})}(\mathbf{x})$ where $F_{(P, \mathrm{strict})}(\mathbf{x})$ for a poset P means the P-partition enumerator for any strict labelling of P, i.e. a labelling such that the P-partitions satisfy $f(i) < f(j)$ whenever $i <_P j$.*

In particular, $\Psi[M]$ expands nonnegatively in the $\{L_\alpha\}$ basis.

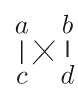
Proof. A basic result about matroids, due to Edmonds [23], describes the *edges* in the *matroid base polytope* which is the convex hull of all vectors $\{\sum_{b \in B} \epsilon_b\}_{B \in \mathcal{B}(M)}$ inside $\mathbb{R}^E$ with standard basis $\{\epsilon_e\}_{e \in E}$. He shows that all such edges connect two bases B, B' that differ by a single *basis exchange*, that is, $B' = (B \setminus \{b\}) \cup \{b'\}$ for some b in B and b' in $E \setminus B$.

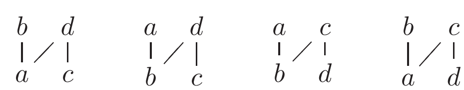
Polyhedral theory then says that a cost function f on E will minimize uniquely at B if and only if one has a strict increase $f(B) < f(B')$ along each such edge $B \to B'$ emanating from B, that is, if and only if $f(b) < f(b')$ whenever $b <_{P_B} b'$ in the base-cobase poset P_B, that is, f lies in $\mathcal{A}(P_B, \mathrm{strict})$. $\square$

Example 6.53. The graphic matroid from Example 6.34 has this matroid base polytope, with the bases B in $\mathcal{B}(M)$ labelling the vertices:

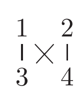


The base-cobase posets P_B for its five vertices B are as follows:





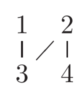
One can label the first of these five strictly as



and compute its strict P-partition enumerator from the linear extensions $\{3412, 3421, 4312, 4321\}$ as

$$L_{(2,2)} + L_{(2,1,1)} + L_{(1,1,2)} + L_{(1,1,1,1)}$$

while any of the last four can be labelled strictly as



and they each have an extra linear extension 3142 giving their strict P-partition enumerators as

$$L_{(2,2)} + L_{(2,1,1)} + L_{(1,1,2)} + L_{(1,1,1,1)} + L_{(1,2,1)}.$$

Hence one has

$$\Psi[M] = 5L_{(2,2)} + 5L_{(1,1,2)} + 4L_{(1,2,1)} + 5L_{(2,1,1)} + 5L_{(1,1,1,1)}.$$

As M is a graphic matroid for a self-dual planar graph, one has a matroid isomorphism $M \cong M^{\perp}$ (see Example 6.49), reflected in the fact that $\Psi[M]$ is invariant under the symmetry swapping $M_\alpha \leftrightarrow M_{\mathrm{rev}(\alpha)}$ (and simultaneously swapping $L_\alpha \leftrightarrow L_{\mathrm{rev}(\alpha)}$).

This P-partition expansion for $\Psi[M]$ also allows us to identify its image under the antipode of QSym.

Proposition 6.54. *For a matroid M on ground set E, one has*

$$S(\Psi[M]) = (-1)^{|E|} \sum_{f : E \to \{1,2,\ldots\}} |\{f\text{-maximizing bases } B\}| \cdot \mathbf{x}_f$$

and

$$\mathrm{ps}^1 \Psi[M](-m) = (-1)^{|E|} \sum_{f : E \to \{1,2,\ldots,m\}} |\{f\text{-maximizing bases } B\}|.$$

In particular, the expected number of f-maximizing bases among all cost functions $f : E \to \{1, 2, ..., m\}$ is $(-m)^{-|E|}\mathrm{ps}^1\Psi[M](-m)$.

Proof. Corollary 5.22 implies

$$S(\Psi[M]) = \sum_{B \in \mathcal{B}(M)} S(F_{(P_B,\mathrm{strict})}(\mathbf{x})) = (-1)^{|E|} \sum_{B \in \mathcal{B}(M)} F_{(P_B^{\mathrm{opp}},\mathrm{natural})}(\mathbf{x})$$

where $F_{(P,\mathrm{natural})}(\mathbf{x})$ is the enumerator for P-partitions in which P has been *naturally* labelled, so that they satisfy $f(i) \leq f(j)$ whenever $i <_P j$. When $P = P_B^{\mathrm{opp}}$, this is exactly the condition for f to achieve its maximum value at $f(B)$ (possibly not uniquely), that is, for f to lie in the *closed* normal cone to the vertex indexed by B in the matroid base polytope; compare this with the discussion in the proof of Proposition 6.52. Thus one has

$$S(\Psi[M]) = (-1)^{|E|} \sum_{\substack{(B,f):\\ B \in \mathcal{B}(M)\\ f \text{ maximizing at } B}} \mathbf{x}_f$$

which agrees with the statement of the proposition, after reversing the order of the summation.

The rest follows from Proposition 6.7. ☐

Example 6.55. We saw in Example 6.53 that the matroid M from Example 6.34 has

$$\Psi[M] = 5L_{(2,2)} + 5L_{(1,1,2)} + 4L_{(1,2,1)} + 5L_{(2,1,1)} + 5L_{(1,1,1,1)}.$$

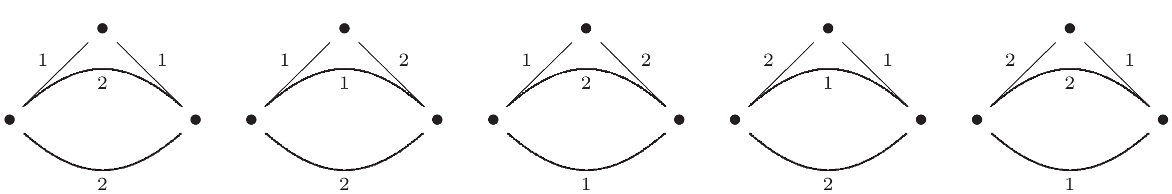
and therefore will have

$$\mathrm{ps}^1\Psi[M](m) = 5\binom{m-2+4}{4} + (5+4+5)\binom{m-3+4}{4} + 5\binom{m-4+4}{4} = \frac{m(m-1)(2m^2-2m+1)}{2}.$$

using $\mathrm{ps}^1(L_\alpha)(m) = \binom{m-\ell+|\alpha|}{|\alpha|}$ from Proposition 6.7 (i). Let us first do a reality-check on a few of its values with $m \geq 0$ using Proposition 6.43, and for negative m using Proposition 6.54:

m	-1	0	1	2
$\mathrm{ps}^1\Psi[M](m)$	5	0	0	5

When $m = 0$, interpreting the set of cost functions $f : E \to \{1, 2, \ldots, m\}$ as being empty explains why the value shown is 0. When $m = 1$, there is only one function $f : E \to \{1\}$, and it is not M-generic; any of the 5 bases in $\mathcal{B}(M)$ will minimize $f(B)$, explaining both why the value for $m = 1$ is 0, but also explaining the value of 5 for $m = -1$. The value of 5 for $m = 2$ counts these M-generic cost functions $f : E \to \{1, 2\}$:



Lastly, Proposition 6.54 predicts the expected number of f-minimizing bases for $f : E \to \{1, 2, \ldots, m\}$ as

$$(-m)^{-|E|}\mathrm{ps}^1\Psi[M](-m) = (-m)^{-4}\frac{m(m+1)(2m^2+2m+1)}{2} = \frac{(m+1)(2m^2+2m+1)}{2m^3},$$

whose limit as $m \to \infty$ is 1, consistent with the notion that "most" cost functions should be generic with respect to the bases of M, and maximize/minimize on a unique basis.

Remark 6.56. It is not coincidental that there is a similarity of results for Stanley's chromatic symmetric function of a graph $\Psi[G]$ and for the matroid quasisymmetric function $\Psi[M]$, such as the P-partition expansions (6.10) versus Proposition 6.52, and the reciprocity results Proposition 6.30 versus Proposition 6.54. It was noted in [11, §9] that one can associate a similar quasisymmetric function invariant to any *generalized permutohedra* in the sense of Postnikov [59]. Furthermore, recent work of Ardila and Aguiar [3] has shown that there is a Hopf algebra of such generalized permutohedra, arising from a *Hopf monoid* in the sense of Aguiar and Mahajan [5]. This Hopf algebra generalizes the chromatic Hopf algebra of graphs and the matroid-minor Hopf algebra, and its quasisymmetric function invariant derives as usual from Theorem 6.3. Their work [3] also provides a generalization of the chromatic Hopf algebra antipode formula of Humpert and Martin [37] discussed in Remark 6.21 above.

7. The Malvenuto-Reutenauer Hopf algebra of permutations

Like so many Hopf algebras we have seen, the *Malvenuto-Reutenauer Hopf algebra* FQSym can be thought of fruitfully in more than one way. One is that it gives a natural noncommutative lift of the quasisymmetric P-partition enumerators and the fundamental basis $\{L_\alpha\}$ of QSym, rendering their product and coproduct formulas even more natural.

7.1. **Definition and Hopf structure.**

Definition 7.1. Define $\mathrm{FQSym} = \bigoplus_{n\geq 0} \mathrm{FQSym}_n$ to be a graded $\mathbf{k}$-module in which FQSym_n has $\mathbf{k}$-basis $\{F_w\}_{w\in\mathfrak{S}_n}$ indexed by the permutations $w = (w_1, \ldots, w_n)$ in $\mathfrak{S}_n$.

We first attempt to lift the product and coproduct formulas (5.7), (5.6) in the $\{L_\alpha\}$ basis of QSym. We attempt to define a product for $u \in \mathfrak{S}_k, v \in \mathfrak{S}_\ell$

$$F_u F_v := \sum_{w \in u \,\sqcup\!\sqcup\, v[k]} F_w \tag{7.1}$$

where for $v = (v_1, \ldots, v_\ell)$ one sets $v[k] := (k + v_1, \ldots, k + v_\ell)$.

The coproduct will be defined using the notation of *standardization* $\mathrm{std}(\mathbf{i})$ of a word $\mathbf{i} = (i_1, \ldots, i_n)$ in some linearly ordered alphabet, which is the permutation in $\mathfrak{S}_n$ obtained by replacing all the occurrences of the smallest letter in $\mathbf{i}$ by the numbers $1, 2, \ldots, m_1$ from left to right, then replacing all occurrences of the next smallest letter by the numbers $m_1 + 1, m_1 + 2, \ldots, m_1 + m_2$, from left to right, etc.

Example 7.2. Considering words in the Roman alphabet $a < b < c < \cdots$

$$\begin{array}{rcccccccccc} \mathrm{std}(b & a & c & c & b & a & a & b & a & c & b) \\ = (5 & 1 & 9 & 10 & 6 & 2 & 3 & 7 & 4 & 11 & 8). \end{array}$$

Using this, define for $w = (w_1, \ldots, w_n)$ in $\mathfrak{S}_n$

$$\Delta F_w := \sum_{k=0}^{n} F_{\mathrm{std}(w_1,w_2,\ldots,w_k)} \otimes F_{\mathrm{std}(w_{k+1},w_{k+2},\ldots,w_n)}. \tag{7.2}$$

It is possible to check directly that the maps defined in (7.1) and (7.2) endow FQSym with the structure of a graded connected finite type Hopf algebra; see Hazewinkel, Gubareni, Kirichenko [29, Thm. 7.1.8]. However in justifying this here, we will follow the approach of Duchamp, Hivert and Thibon [22], which exhibits FQSym as a subalgebra of a larger ring of (noncommutative) power series of bounded degree in a totally ordered alphabet.

Definition 7.3. Given a totally ordered set I, create a totally ordered variable set $\{X_i\}_{i\in I}$, and the ring $R\langle\{X_i\}_{i\in I}\rangle$ of noncommutative power series of bounded degree in this alphabet. Many times, we will use a variable set $\mathbf{X} := (X_1 < X_2 < \cdots)$, and call the ring $R\langle\mathbf{X}\rangle$.

We first identify the algebra structure for FQSym as the subalgebra of finite type within $R\langle\{X_i\}_{i\in I}\rangle$ spanned by the elements

$$F_w = F_w(\{X_i\}_{i\in I}) := \sum_{\substack{\mathbf{i}=(i_1,\ldots,i_n):\\ \mathrm{std}(\mathbf{i})=w^{-1}}} \mathbf{X}_{\mathbf{i}}$$

where $\mathbf{X}_{\mathbf{i}} := X_{i_1} \cdots X_{i_n}$, as w ranges over $\bigcup_{n\geq 0} \mathfrak{S}_n$.

Example 7.4. For the alphabet $\mathbf{X} = (X_1 < X_2 < \cdots)$, in $R\langle\mathbf{X}\rangle$ one has

$$\begin{aligned}
F_1 &= \sum_{1 \le i} X_i = X_1 + X_2 + \cdots \\
F_{12} &= \sum_{1 \le i \le j} X_i X_j = X_1^2 + X_2^2 + \cdots + X_1X_2 + X_1X_3 + X_2X_3 + X_1X_4 + \cdots \\
F_{21} &= \sum_{1 \le i < j} X_j X_i = X_2X_1 + X_3X_1 + X_3X_2 + X_4X_1 + \cdots \\
F_{312} &= \sum_{\mathbf{i}:\mathrm{std}(\mathbf{i})=231} \mathbf{X}_{\mathbf{i}} = \sum_{1 \le i < j \le k} X_j X_k X_i \\
&= X_2^2X_1 + X_3^2X_1 + X_3^2X_2 + \cdots + X_2X_3X_1 + X_2X_4X_1 + \cdots
\end{aligned}$$

Proposition 7.5. *For any totally ordered infinite set I, the elements $\{F_w\}$ as w ranges over $\bigcup_{n\ge 0}\mathfrak{S}_n$ form the $\mathbf{k}$-basis for a subalgebra* $\mathrm{FQSym}(\{X_i\}_{i\in I})$ *of $R\langle\mathbf{X}\rangle$, which is graded connected and of finite type, having multiplication defined $\mathbf{k}$-linearly by (7.1).*

Consequently all such algebras are isomorphic to a single algebra FQSym*, having basis $\{F_w\}$ and multiplication given by the rule (7.1), with the isomorphism mapping $F_w \longmapsto F_w(\{X_i\}_{i\in I})$.*

For example,

$$\begin{aligned}
F_1F_{21} &= (X_1 + X_2 + X_3 + \cdots)(X_2X_1 + X_3X_1 + X_3X_2 + X_4X_1 + \cdots) \\
&= X_1 \cdot X_3X_2 + X_1 \cdot X_4X_2 + \cdots + X_1 \cdot X_2X_1 + X_2 \cdot X_3X_2 + X_2 \cdot X_4X_2 + \cdots \\
&\quad + X_2 \cdot X_3X_1 + X_2 \cdot X_4X_1 + \cdots + X_2 \cdot X_2X_1 + X_3 \cdot X_3X_1 + X_3 \cdot X_3X_2 + \cdots \\
&\quad + X_3 \cdot X_2X_1 + X_4 \cdot X_2X_1 + \cdots \\
&= \sum_{\mathbf{i}:\mathrm{std}(\mathbf{i})=132} \mathbf{X}_{\mathbf{i}} + \sum_{\mathbf{i}:\mathrm{std}(\mathbf{i})=231} \mathbf{X}_{\mathbf{i}} + \sum_{\mathbf{i}:\mathrm{std}(\mathbf{i})=321} \mathbf{X}_{\mathbf{i}} = F_{132} + F_{312} + F_{321} = \sum_{w \in 1 \sqcup\!\sqcup 32} F_w
\end{aligned}$$

Proof. The elements $\{F_w(\{X_i\}_{i\in I})\}$ are linearly independent as they are supported on disjoint monomials, and so form a $\mathbf{k}$-basis for their span. The fact that they multiply via rule (7.1) is the equivalence of conditions (i) and (iii) in the following Lemma 7.6, from which all the remaining assertions follow. □

Lemma 7.6. *For a triple of permutations*

$$\begin{aligned}
u &= (u_1, \ldots, u_k) \text{ in } \mathfrak{S}_k, \\
v &= (v_1, \ldots, v_{n-k}) \text{ in } \mathfrak{S}_{n-k}, \\
w &= (w_1, \ldots, w_n) \text{ in } \mathfrak{S}_n,
\end{aligned}$$

the following conditions are equivalent:

(i) *w^{-1} lies in the set $u^{-1} \sqcup\!\sqcup v^{-1}[k]$.*
(ii) *$u = \mathrm{std}(w_1, \ldots, w_k)$ and $v = \mathrm{std}(w_{k+1}, \ldots, w_n)$,*
(iii) *for some word $\mathbf{i} = (i_1, \ldots, i_n)$ with $\mathrm{std}(\mathbf{i}) = w$ one has $u = \mathrm{std}(i_1, \ldots, i_k)$ and $v = \mathrm{std}(i_{k+1}, \ldots, i_n)$.*

Proof of Lemma. The implication (ii) ⇒ (iii) is clear since $\mathrm{std}(w) = w$. The reverse implication (iii) ⇒ (ii) is best illustrated by example, e.g. considering Example 7.2 as concatenated, with $n = 11, k = 6, n - k = 5$:

$$\begin{array}{rccccccccccccc}
w = \mathrm{std} & (b & a & c & c & b & a & & | & a & b & a & c & b) \\
= & (5 & 1 & 9 & 10 & 6 & 2 & & | & 3 & 7 & 4 & 11 & 8)
\end{array}$$

$$\begin{array}{rccccccccccccc}
u = \mathrm{std} & (5 & 1 & 9 & 10 & 6 & 2 &) & v = \mathrm{std}(& 3 & 7 & 4 & 11 & 8) \\
= & (3 & 1 & 5 & 6 & 4 & 2 &) & = (& 1 & 3 & 2 & 5 & 4) \\
= \mathrm{std} & (b & a & c & c & b & a &) & = \mathrm{std}(& a & b & a & c & b)
\end{array}$$

The equivalence of (i) and (ii) is a fairly standard consequence of unique parabolic factorization $W = W^J W_J$ where $W = \mathfrak{S}_n$ and $W_J = \mathfrak{S}_k \times \mathfrak{S}_{n-k}$, so that W^J are the minimum-length coset representatives for cosets xW_J (that is, the permutations $x \in \mathfrak{S}_n$ satisfying $x_1 < \cdots < x_k$ and $x_{k+1} < \cdots < x_n$). One can uniquely express any w in W as $w = xy$ with x in W^J and y in W_J, which here means that $y = u \cdot v[k] = v[k] \cdot u$

for some u in $\mathfrak{S}_k$ and v in $\mathfrak{S}_{n-k}$. Therefore $w = xuv[k]$, if and only if $w^{-1} = u^{-1}v^{-1}[k]x^{-1}$, which means that w^{-1} is the shuffle of the sequences u^{-1} in positions $\{x_1, \ldots, x_k\}$ and $v^{-1}[k]$ in positions $\{x_{k+1}, \ldots, x_n\}$. $\square$

Example 7.7. To illustrate the equivalence of (i) and (ii) and the parabolic factorization in the preceding proof, let $n = 9$ and $k = 5$ with

$$\begin{aligned} w &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & | & 6 & 7 & 8 & 9 \\ 4 & 9 & 6 & 1 & 5 & | & 8 & 2 & 3 & 7 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & | & 6 & 7 & 8 & 9 \\ 1 & 4 & 5 & 6 & 9 & | & 2 & 3 & 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 1 & 3 \end{pmatrix} \begin{pmatrix} 6 & 7 & 8 & 9 \\ 9 & 6 & 7 & 8 \end{pmatrix} \\ &= x \cdot u \cdot v[k] \end{aligned}$$

$$\begin{aligned} w^{-1} &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 9 & 6 & 1 & 5 & 8 & 2 & 3 & 7 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 3 & 2 \end{pmatrix} \begin{pmatrix} 6 & 7 & 8 & 9 \\ 7 & 8 & 9 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \underline{1} & \underline{\underline{6}} & \underline{\underline{7}} & \underline{2} & \underline{3} & \underline{4} & \underline{\underline{8}} & \underline{\underline{9}} & \underline{5} \end{pmatrix} \\ &= u^{-1} \cdot v^{-1}[k] \cdot x^{-1} \end{aligned}$$

One can now use this to define a coalgebra structure on $R\langle \mathbf{X} \rangle$ as follows. Given the ordered variable set

$$(\mathbf{X}, \mathbf{Y}) := (X_1 < X_2 < \cdots < Y_1 < Y_2 < \cdots)$$

form the ring $R\langle \mathbf{X}, \mathbf{Y} \rangle$ and its quotient $R\langle \mathbf{X}, \mathbf{Y} \rangle / [\mathbf{X}, \mathbf{Y}]$ by the two-sided ideal generated by all commutators $[X_i, Y_j] = X_i Y_j - Y_j X_i$, in which one has forced the $\mathbf{X}$ variables to commute with $\mathbf{Y}$ variables. One can check that

$$R\langle \mathbf{X}, \mathbf{Y} \rangle / [\mathbf{X}, \mathbf{Y}] \cong R\langle \mathbf{X} \rangle \otimes R\langle \mathbf{Y} \rangle$$

giving a ring homomorphism $\mathrm{FQSym} \xrightarrow{\Delta} R\langle \mathbf{X} \rangle \otimes R\langle \mathbf{Y} \rangle$ which is the composite of these ring morphisms:

$$\begin{array}{ccl} \mathrm{FQSym} & \cong & \mathrm{FQSym}(\mathbf{X}, \mathbf{Y}) \hookrightarrow R\langle \mathbf{X}, \mathbf{Y} \rangle \longrightarrow R\langle \mathbf{X}, \mathbf{Y} \rangle / [\mathbf{X}, \mathbf{Y}] \cong R\langle \mathbf{X} \rangle \otimes R\langle \mathbf{Y} \rangle \\ f(\mathbf{X}) & \longmapsto & f(\mathbf{X}, \mathbf{Y}). \end{array} \tag{7.3}$$

Example 7.8. Recall from Example 7.4 that one has

$$F_{312} = \sum_{\mathbf{i} : \mathrm{std}(\mathbf{i}) = 231} \mathbf{X}_{\mathbf{i}} = \sum_{1 \le i < j \le k} X_j X_k X_i$$

and therefore its coproduct is

$$\begin{aligned} \Delta F_{312} &= F_{312}(X_1, X_2, \ldots, Y_1, Y_2, \ldots) \\ &= \sum_{i < j \le k} X_j X_k X_i + \sum_{\substack{i<j, \\ k}} X_j Y_k X_i + \sum_{\substack{i, \\ j \le k}} Y_j Y_k X_i + \sum_{i < j \le k} Y_j Y_k Y_i \\ &= \sum_{i < j \le k} X_j X_k X_i \cdot 1 + \sum_{\substack{i<j, \\ k}} X_j X_i \cdot Y_k + \sum_{\substack{i, \\ j \le k}} X_i \cdot Y_j Y_k + \sum_{i < j \le k} 1 \cdot Y_j Y_k Y_i \\ &= F_{312}(\mathbf{X}) \cdot 1 + F_{21}(\mathbf{X}) \cdot F_1(\mathbf{Y}) + F_1(\mathbf{X}) \cdot F_{12}(\mathbf{Y}) + 1 \cdot F_{312}(\mathbf{Y}) \\ &= F_{312} \otimes 1 + F_{21} \otimes F_1 + F_1 \otimes F_{12} + 1 \otimes F_{312} \end{aligned}$$

Proposition 7.9. *The image of the composite map in* (7.3) *lies in* $\mathrm{FQSym}(\mathbf{X}) \otimes \mathrm{FQSym}(\mathbf{Y})$*, giving rise to a coproduct*

$$\begin{array}{ccc} \mathrm{FQSym}(\mathbf{X}, \mathbf{Y}) & \xrightarrow{\Delta} & \mathrm{FQSym}(\mathbf{X}) \otimes \mathrm{FQSym}(\mathbf{Y}) \\ \| & & \| \\ \mathrm{FQSym} & & \mathrm{FQSym} \otimes \mathrm{FQSym} \end{array}$$

whose action on the $\{F_w\}$ *basis is given by the rule* (7.2)*. This endows* FQSym *with the structure of a graded connected finite type Hopf algebra.*

Proof. One has

$$\Delta F_w = F_w(\mathbf{X},\mathbf{Y}) = \sum_{\mathbf{i}:\mathrm{std}(\mathbf{i})=w^{-1}} (\mathbf{X},\mathbf{Y})_{\mathbf{i}} = \sum_{k=0}^{n} \sum_{\substack{(\mathbf{i},\mathbf{j}):\\ \mathrm{std}(\mathbf{i})=(\mathrm{std}(w_1,\ldots,w_k))^{-1}\\ \mathrm{std}(\mathbf{j})=(\mathrm{std}(w_{k+1},\ldots,w_n))^{-1}}} \mathbf{X}_{\mathbf{i}}\mathbf{Y}_{\mathbf{j}}$$
$$= \sum_{k=0}^{n} F_{\mathrm{std}(w_1,\ldots,w_k)}(\mathbf{X}) F_{\mathrm{std}(w_{k+1},\ldots,w_n)}(\mathbf{Y})$$

relying again on Lemma 7.6. □

Corollary 7.10. *The Hopf algebra* FQSym *is self-dual: the map sending* $F_w \longmapsto F_{w^{-1}}$ *gives an isomorphism* $\mathrm{FQSym} \longrightarrow \mathrm{FQSym}^o$*.*

Proof. One has

$$F_{u^{-1}}F_{v^{-1}} = \sum_{w^{-1} \in u^{-1} \sqcup v^{-1}[k]} F_{w^{-1}} = \sum_{\substack{w:\\ \mathrm{std}(w_1,\ldots,w_k)=u\\ \mathrm{std}(w_{k+1},\ldots,w_n)=v}} F_{w^{-1}}$$

via the equivalence of (i) and (ii) in Lemma 7.6. On the other hand, in FQSym^o, the dual $\mathbf{k}$-basis $\{G_w\}$ to the $\mathbf{k}$-basis $\{F_w\}$ for FQSym should have product formula

$$G_uG_v = \sum_{\substack{w:\\ \mathrm{std}(w_1,\ldots,w_k)=u\\ \mathrm{std}(w_{k+1},\ldots,w_n)=v}} G_w$$

coming from the coproduct formula (7.2) for FQSym in the $\{F_w\}$-basis. □

We can now be a bit more precise about the relations between the various algebras

$$\Lambda, \mathrm{QSym}, \mathrm{NSym}, \mathrm{FQSym}, R\langle\mathbf{X}\rangle, R(\mathbf{x}).$$

Not only does FQSym allow one to *lift* the Hopf structure of QSym, it dually allows one to *extend* the Hopf structure of NSym. To set up this duality, note that Corollary 7.10 motivates the choice of an inner product on FQSym in which

$$(F_u, F_v) := \delta_{u^{-1},v}.$$

We wish to identify the images of the ribbon basis $\{R_\alpha\}$ of NSym when included in FQSym.

Definition 7.11. Define the *free quasi-ribbon function*

$$\mathbf{R}_\alpha := \sum_{w:\mathrm{Des}(w)=D(\alpha)} F_{w^{-1}} = \sum_{\substack{(w,\mathbf{i}):\\ \mathrm{Des}(w)=D(\alpha)\\ \mathrm{std}(\mathbf{i})=w}} \mathbf{X}_{\mathbf{i}} = \sum_{\mathbf{i}:\mathrm{Des}(\mathbf{i})=D(\alpha)} \mathbf{X}_{\mathbf{i}}$$

where the w in the sums are supposed to belong to $\mathfrak{S}_{|\alpha|}$, and where the *descent set* of a sequence $\mathbf{i} = (i_1,\ldots,i_n)$ is defined by

$$\mathrm{Des}(\mathbf{i}) := \{j \in \{1,2,\ldots,n-1\} : i_j > i_{j+1}\} = \mathrm{Des}(\mathrm{std}(\mathbf{i})).$$

Alternatively, $\mathbf{R}_\alpha = \sum_T \mathbf{X}_T$ in which the sum is over column-strict tableaux of the ribbon skew shape α, and $\mathbf{X}_T = \mathbf{X}_{\mathbf{i}}$ in which $\mathbf{i}$ is the sequence of entries of T read in order from the southwest toward the northeast.

Example 7.12. Taking $\alpha = (1,3,2)$, with ribbon shape and column-strict fillings T as shown

$$\begin{array}{cccc} & & \square & \square \\ \square & \square & \square & \\ \square & & & \end{array} \qquad T = \begin{array}{ccccccc} & & & & i_5 & \leq & i_6 \\ & & & & \wedge & & \\ i_2 & \leq & i_3 & \leq & i_4 & & \\ \wedge & & & & & & \\ i_1 & & & & & & \end{array}$$

one has that

$$\mathbf{R}_{(1,3,2)} = \sum_{\substack{\mathbf{i}=(i_1,i_2,i_3,i_4,i_5,i_6):\\ \mathrm{Des}(\mathbf{i})=D(\alpha)=\{1,4\}}} \mathbf{X}_{\mathbf{i}} = \sum_{i_1>i_2\leq i_3\leq i_4>i_5\leq i_6} X_{i_1}X_{i_2}X_{i_3}X_{i_4}X_{i_5}X_{i_6} = \sum_T \mathbf{X}_T$$

Corollary 7.13. *The following surjection and injection are Hopf morphisms*

$$\begin{array}{rcl} \mathrm{FQSym} & \overset{\pi}{\twoheadrightarrow} & \mathrm{QSym} \\ F_w & \longmapsto & L_{\gamma(w)} \end{array}$$

$$\begin{array}{rcl} \mathrm{NSym} & \overset{\iota}{\hookrightarrow} & \mathrm{FQSym} \\ R_\alpha & \longmapsto & \mathbf{R}_\alpha \end{array}$$

and are also adjoint maps with respect to the above choice of inner product on FQSym *and the usual dual pairing between* NSym *and* QSym*. Furthermore, the second map ι lets one factor the surjection* $\mathrm{NSym} \twoheadrightarrow \Lambda$ *as follows*

$$\begin{array}{rcccl} \mathrm{NSym} & \rightarrow & \mathrm{FQSym} \hookrightarrow R\langle \mathbf{X} \rangle & \rightarrow & R(\mathbf{x}) \\ R_\alpha & \longmapsto & \mathbf{R}_\alpha & \longmapsto & s_\alpha(\mathbf{x}) \end{array}$$

through the abelianization $R\langle \mathbf{X} \rangle \twoheadrightarrow R(\mathbf{x})$ sending the noncommutative variable X_i to the commutative x_i.

Proof. The fact that $\mathrm{FQSym} \overset{\pi}{\twoheadrightarrow} \mathrm{QSym}$ is a Hopf map comes from checking that it respects the product (compare (5.7) and (7.1)) and the coproducts (compare (5.6) and (7.2)) then applying Proposition 1.35(c). It follows that $\mathrm{NSym} \overset{\iota}{\hookrightarrow} \mathrm{FQSym}$ is a Hopf map once we check the adjointness assertion, via the following calculation:

$$(\iota(R_\alpha), F_w) = (\mathbf{R}_\alpha, F_w) = \sum_{u : \mathrm{Des}(u) = D(\alpha)} (F_{u^{-1}}, F_w) = \left\{ \begin{array}{ll} 1 & \text{if } \mathrm{Des}(w) = D(\alpha) \\ 0 & \text{otherwise} \end{array} \right\}$$
$$= (R_\alpha, L_{\gamma(w)}) = (R_\alpha, \pi(F_w)).$$

The last assertion is clear: the abelianization map sends the noncommutative tableau monomial $\mathbf{X}_T$ to the commutative tableau monomial $\mathbf{x}_T$. ☐

We summarize some of this picture as follows:

$$\begin{tikzcd} \mathrm{FQSym} \arrow[rr, leftrightarrow, "\mathrm{id}"] & & \mathrm{FQSym} \arrow[d, twoheadrightarrow, "\pi"] \\ \mathrm{NSym} \arrow[u, hookrightarrow, "\iota"'] \arrow[dr, "\pi"] & & \mathrm{QSym} \\ & \Lambda \arrow[ur, hookrightarrow] & \end{tikzcd}$$

(This is not a commutative diagram!)

8. 0-Hecke algebras

8.1. **Review of representation theory of finite-dimensional algebras.** Review the notions of indecomposables, simples, projectives, along with the theorems of Krull-Remak-Schmidt, of Jordan-Hölder, and the two kinds of Grothendieck groups dual to each other.

8.2. 0**-Hecke algebra representation theory.** Describe the simples and projectives, following Denton, Hivert, Schilling, Thiery on $\mathcal{J}$-trivial monoids.

8.3. **Nsym and Qsym as Grothendieck groups.** Give Krob and Thibon's interpretation of

- QSym and the Grothendieck group of composition series, and
- NSym and the Grothendieck group of projectives.

Remark 8.1. Mention P. McNamara's interpretation, in the case of *supersolvable lattices*, of the Ehrenborg quasisymmetric function as the composition series enumerator for an $H_n(0)$-action on the maximal chains

9. Aguiar-Bergeron-Sottile character theory Part II: Odd and even characters, subalgebras

10. Face enumeration, Eulerian posets, and cd-indices

Borrowing from Billera's ICM notes.

10.1. f-vectors, h-vectors.

10.2. flag f-vectors, flag h-vectors.

10.3. ab-indices and cd-indices.

11. Further topics

Some of these we may touch on in class, others are appropriate for student talks.

- Loday-Ronco Hopf algebra of planar binary trees
- Poirier-Reutenauer Hopf algebra of tableaux
- Reading Hopf algebra of Baxter permutations
- Hopf monoids, e.g. of Hopf algebra of generalized permutohedra, of matroids, of graphs, Stanley chromatic symmetric functions and Tutte polynomials
- Lam-Pylyavskyy Hopf algebra of set-valued tableaux
- Connes-Kreimer Hopf algebra and renormalization
- Noncommutative symmetric functions and $\Omega\Sigma\mathbb{C}P^\infty$
- Maschke's theorem and "integrals" for Hopf algebras
- Nichols-Zoeller structure theorem and group-like elements
- Cartier-Milnor-Moore structure theorem and primitive elements
- Quasi-triangular Hopf algebras and quantum groups
- The Steenrod algebra, its dual, and tree Hopf algebras
- Ringel-Hall algebras of quivers
- Ellis-Khovanov odd symmetric function Hopf algebras (see also Lauda-Russell)

Student talks given in class were:

(1) Al Garver, on Maschke's theorem for finite-dimensional Hopf algebras
(2) Jonathan Hahn, on the paper by Humpert and Martin.
(3) Emily Gunawan, on the paper by Lam, Lauve and Sottile.
(4) Jonas Karlsson, on the paper by Connes and Kreimer
(5) Thomas McConville, on Butcher's group and generalized Runge-Kutta methods.
(6) Cihan Bahran, on universal enveloping algebras and the Poincaré-Birkhoff-Witt theorem.
(7) Theodosios Douvropolos, on the Cartier-Milnor-Moore theorem.
(8) Alex Csar, on the Loday-Ronco Hopf algebra of binary trees
(9) Kevin Dilks, on Reading's Hopf algebra of (twisted) Baxter permutations
(10) Becky Patrias, on the paper by Lam and Pylyavskyy
(11) Meng Wu, on multiple zeta values and Hoffman's homomorphism from QSym

12. Some open problems and conjectures

- Is there a proof of the Assaf-McNamara skew Pieri rule that gives a resolution of Specht or Schur/Weyl modules whose character corresponds to $s_{\lambda/\mu}h_n$, whose terms model their alternating sum?
- Explicit antipodes in the Lam-Pylyavskyy Hopf algebras?
- P. McNamara's question: are P-partition enumerators irreducible for connected posets P?
- Stanley's question: are the only P-partition enumerators which are symmetric (not just quasisymmetric) those for which P is a skew-shape with a column-strict labelling?
- Does Stanley's chromatic symmetric function distinguish trees?
- Hoffman's stuffle conjecture
- Billera-Brenti's nonnegativity conjecture for the total cd-index of Bruhat intervals

Acknowledgements

The author thanks the following for helpful comments and/or teaching him about Hopf algebras: Marcelo Aguiar, Federico Ardila, Lou Billera, Richard Ehrenborg, Mark Haiman, Florent Hivert, Christophe Hohlweg, Jia Huang, Jang Soo Kim, Aaron Lauve, John Palmieri, Margie Readdy, Nathan Reading, Muge Taskin, Jean-Yves Thibon.

References

[1] E. Abe. Hopf algebras. *Cambridge Tracts in Mathematics* **74**. Cambridge University Press, Cambridge-New York, 1980.

[2] M. Aguiar, et. al. (28 authors). Supercharacters, symmetric functions in noncommuting variables, and related Hopf algebras. *Adv. Math.* **229** (2012), 2310–2337.

[3] M. Aguiar and F. Ardila. The Hopf monoid of generalized permutahedra, in preparation, 2012.

[4] M. Aguiar, N. Bergeron, and F. Sottile. Combinatorial Hopf algebras and generalized Dehn-Sommerville relations. *Compos. Math.* **142** (2006), 1 – 30. Newer version at `http://www.math.tamu.edu/~maguiar/CHalgebra.pdf`.

[5] M. Aguiar and S. Mahajan. Monoidal functors, species and Hopf algebras. *CRM Monograph Series* **29**. American Mathematical Society, Providence, RI, 2010.

[6] M. Aguiar and F. Sottile, Structure of the Malvenuto-Reutenauer Hopf algebra of permutations. *Adv. Math.* **191** (2005), 225–275.

[7] S. Assaf and P. McNamara. A Pieri rule for skew shapes. *J. Combin. Theory, Ser. A* **118** (2011) 277–290.

[8] A. Baker and B. Richter. Quasisymmetric functions from a topological point of view. *Math. Scand.* **103** (2008), 208–242.

[9] F. Barekat, V. Reiner, S. van Willigenburg, Corrigendum to "Coincidences among skew Schur functions" [*Adv. Math.* **216** (2007), 118–152], *Adv. Math.* **220**(2009), 1655–1656. See also a corrected version of this paper on `arXiv:math/0602634v4`.

[10] L.J. Billera. Flag enumeration in polytopes, Eulerian partially ordered sets and Coxeter groups. *Proceedings of the International Congress of Mathematicians* **IV**, 2389–2415, Hindustan Book Agency, New Delhi, 2010.

[11] L.J. Billera, N. Jia, and V. Reiner, A quasisymmetric function for matroids. *European J. Combin.* **30** (2009), 1727 – 1757.

[12] A. Björner, Some combinatorial and algebraic properties of Coxeter complexes and Tits buildings. *Adv. in Math.* **52** (1984), 173–212.

[13] D. Bump, Notes on representations of $GL(r)$ over a finite field, available at `http://math.stanford.edu/~bump/`.

[14] P. F. Cartier, A primer of Hopf algebras. Frontiers in number theory, physics, and geometry. II, 537–615, Springer, Berlin, 2007.

[15] V. Chari and A. Pressley, A guide to quantum groups. Cambridge University Press, Cambridge, 1994.

[16] H. Crapo and W. Schmitt, Primitive elements in the matroid-minor Hopf algebra. *J. Algebraic Combin.* **28** (2008), 43–64.

[17] ———, A unique factorization theorem for matroids. *J. Combin. Theory Ser. A* **112** (2005), 222–249.

[18] ———, A free subalgebra of the algebra of matroids. *European J. Combin.* **26** (2005), 1066–1085.

[19] S. Dascalescu, C. Nastasescu, S. Raianu, Hopf algebras. An introduction. *Monographs and Textbooks in Pure and Applied Mathematics* **235**. Marcel Dekker, Inc., New York, 2001.

[20] T. Denton, F. Hivert, A. Schilling, and N. Thiery. On the representation theory of finite J-trivial monoids. *Sém. Lothar. Combin.* **64** (2010/11), Art. B64d, 44 pp.

[21] P. Diaconis, A. Pang and A. Ram. Hopf algebras and Markov chains: Two examples and a theory. *J. Algebraic Combin.* **39**, Issue 3, May 2014, 527–585. A newer version is available at `http://math.stanford.edu/~amyp/v4hopfpowermarkovchains.pdf`.

[22] G. Duchamp, F. Hivert, and J.-Y. Thibon, Noncommutative symmetric functions VI. Free quasi-symmetric functions and related algebras. *Internat. J. Algebra Comput.* **12** (2002), 671–717.

[23] J. Edmonds, Submodular functions, matroids, and certain polyhedra, in: Combinatorial Structures and their Applications (Proc. Calgary Internat. Conf., Calgary, Alta., 1969), Gordon and Breach, New York, 1970, pp. 66–87; reprinted in Combinatorial optimization: Eureka, you shrink!, pp. 11–26, *Lecture Notes in Comput. Sci.* **2570**, Springer, Berlin, 2003.

[24] R. Ehrenborg. On posets and Hopf algebras. *Adv. Math.* **119** (1996), 1–25.

[25] A.E. Ellis and M. Khovanov. The Hopf algebra of odd symmetric functions. *Adv. Math.* **231** (2012), 965–999. A newer version is available as `arXiv:1107.5610v2`.

[26] V. Gasharov, Incomparability graphs of (3+1)-free posets are s-positive. Proceedings of the 6th Conference on Formal Power Series and Algebraic Combinatorics (New Brunswick, NJ, 1994). *Discrete Math.* **157** (1996), 193 – 197.

[27] I.M. Gessel, Multipartite P-partitions and inner products of skew Schur functions. Combinatorics and algebra (Boulder, Colo., 1983), 289–317, *Contemp. Math.* **34**, Amer. Math. Soc., Providence, RI, 1984.

[28] M. Hazewinkel, The algebra of quasi-symmetric functions is free over the integers. *Adv. Math.* **164** (2001), 283 – 300.

[29] M. Hazewinkel, N. Gubareni, and V.V. Kirichenko, Algebras, rings and modules. Lie algebras and Hopf algebras. *Mathematical Surveys and Monographs* **168**. American Mathematical Society, Providence, RI, 2010.

[30] R. Henderson, The Algebra Of Multiple Zeta Values. `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.227.5432`

[31] F. Hivert, An introduction to combinatorial Hopf algebras: examples and realizations. *Nato Advanced Study Institute School on Physics and Computer Science, 2005, october, 17–29, Cargese, France.* `http://www-igm.univ-mlv.fr/~hivert/PAPER/Cargese.pdf`

[32] F. Hivert, J.-C. Novelli and J.-Y. Thibon, Commutative combinatorial Hopf algebras. *J. Algebraic Combin.* **28** (2008), no. 1, 65–95. Also available as `arXiv:math/0605262v1`.

[33] ———, The algebra of binary search trees. *Theoret. Comput. Sci.* **339** (2005), no. 1, 129–165.

[34] ———, Trees, functional equations, and combinatorial Hopf algebras. *European J. Combin.* **29** (2008), no. 7, 1682–1695.

[35] M.E. Hoffman, Combinatorics of rooted trees and Hopf algebras. *Trans. AMS* **355** (2003), 3795–3811.

[36] ———, A character on the quasi-symmetric functions coming from multiple zeta values. *The Electronic Journal of Combinatorics* **15** (2008), R97.

[37] B. Humpert and J. Martin, The incidence Hopf algebra of graphs. *SIAM Journal on Discrete Mathematics* **26**, no. 2 (2012), 555-570. Also available as `arXiv:1012.4786v3`.

[38] G. James and M. Liebeck, Representations and characters of groups. 2nd edition, Cambridge University Press, Cambridge-New York, 2001.

[39] S.V. Kerov. Asymptotic representation theory of the symmetric group and its applications in analysis. *Translations of Mathematical Monographs* **219**. American Mathematical Society, Providence, RI, 2003.

[40] S.A. Joni and G.-C. Rota, Coalgebras and bialgebras in combinatorics. *Studies in Applied Mathematics* **61** (1979), 93–139.

[41] C. Kassel. Quantum groups. *Graduate Texts in Mathematics* **155**. Springer, Berlin, 1995.

[42] T. Lam, A. Lauve, and F. Sottile. Skew Littlewood-Richardson rules from Hopf Algebras. *DMTCS Proceedings, 22nd International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2010)* **2010**, 355–366.

[43] T. Lam and P. Pylyavskyy. Combinatorial Hopf algebras and K-homology of Grassmanians. *International Mathematics Research Notices*, **2007** (2007), rnm 125, 48 pages.

[44] M. van Leeuwen. An application of Hopf-Algebra techniques to representations of finite Classical Groups. *Journal of Algebra* **140**, Issue 1, 15 June 1991, pp. 210-246. Also available at `http://wwwmathlabo.univ-poitiers.fr/~maavl/pdf/Hopf.pdf`

[45] A. Liulevicius. Arrows, symmetries and representation rings. *Journal of Pure and Applied Algebra* **19** (1980), 259–273.

[46] J.-L. Loday and M.O. Ronco. Combinatorial Hopf algebras. Quanta of maths, *Clay Math. Proc.* **11**, 347–383, Amer. Math. Soc., Providence, RI, 2010.

[47] ———, Hopf algebra of the planar binary trees. *Adv. Math.* **139** (1998), no. 2, 293–309.

[48] K. Luoto, S. Mykytiuk, S. van Willigenburg, An introduction to quasisymmetric Schur functions – Hopf algebras, quasisymmetric functions, and Young composition tableaux. Springer, May 23, 2013. `http://www.math.ubc.ca/~steph/papers/QuasiSchurBook.pdf`

[49] I.G. Macdonald. Symmetric functions and Hall polynomials. 2nd edition, Oxford University Press, Oxford-New York, 1995.

[50] C. Malvenuto, Produits et coproduits des fonctions quasi-symétriques et de l'algèbre des descents, PhD dissertation, Univ. du Québéc à Montreal, 1993. `http://lacim.uqam.ca/publications_pdf/16.pdf`

[51] C. Malvenuto and C. Reutenauer, Duality between quasi-symmetric functions and the Solomon descent algebra. *J. Algebra* **177** (1995), 967–982.

[52] C. Malvenuto and C. Reutenauer, Plethysm and conjugation of quasi-symmetric functions. *Discrete Mathematics* **193**, Issues 13, 28 November 1998, 225–233.

[53] C. Malvenuto and C. Reutenauer, A self paired Hopf algebra on double posets and a LittlewoodRichardson rule. *Journal of Combinatorial Theory, Series A* **118** (2011), 1322–1333.

[54] D. Manchon, Hopf algebras, from basics to applications to renormalization. *Comptes Rendus des Rencontres Mathematiques de Glanon* 2001 (published in 2003). `arXiv preprint math/0408405`

[55] J. Milnor and J. Moore, On the structure of Hopf algebras. *The Annals of Mathematics, Second Series* **81**, No. 2 (Mar., 1965), 211–264.

[56] S. Montgomery, Hopf algebras and their actions on rings. *Regional Conference Series in Mathematics* **82**, Amer. Math. Soc., Providence, RI, 2010.

[57] J.L. Martin, M. Morin, J.D. Wagner, On distinguishing trees by their chromatic symmetric functions. *Journal of Combinatorial Theory, Series A* **115**, Issue 2, February 2008, 237–253.

[58] J. Oxley, Matroid theory. Oxford University Press, Oxford-New York, 1992.

[59] A. Postnikov, Permutohedra, associahedra, and beyond. *Int. Math. Res. Notices* **2009**, No. 6, pp. 1026–1106.

[60] P. Pylyavskyy, Comparing products of Schur functions and quasisymmetric functions, PhD dissertation, MIT, 2007.

[61] N. Reading. Lattice congruences, fans and Hopf algebras. *Journal of Combinatorial Theory, Series A* **110**, Issue 2, May 2005, pp. 237–273.

[62] B.E. Sagan, The symmetric group: representations, combinatorial algorithms, and symmetric functions. 2nd edition, Springer, New York-Berlin-Heidelberg 2001.

[63] O. Schiffmann. Lectures on Hall algebras. `arXiv:math/0611617`.

[64] W. Schmitt. Incidence Hopf algebras. *Journal of Pure and Applied Algebra* **96** (1994), 299–330.

[65] ———, Hopf algebras of combinatorial structures. *Canadian Journal of Mathematics* **45** (1993), 412–428.

[66] J.-P. Serre. Linear representations of finite groups. Springer, Berlin-Heidelberg-New York, 1977.

[67] J. Shareshian and M.L. Wachs, Chromatic quasisymmetric functions and Hessenberg varieties. In: A. Björner, F. Cohen, C. De Concini, C. Procesi, M. Salvetti (Eds.), Configuration Spaces, *Publications of the Scuola Normale Superiore* **14**, Springer, Berlin-Heidelberg-New York 2013.

[68] S. Shelley-Abrahamson, Hopf Modules and Representations of Finite Groups of Lie Type. Honors thesis, Stanford, May 2013. `https://math.stanford.edu/theses/Shelley-Abrahamson%20Honors%20Thesis.pdf`

[69] R.P. Stanley, Ordered structures and partitions. *Memoirs of the Amer. Math. Soc.* **119**, American Mathematical Society, Providence, R.I., 1972.

[70] ——— Acyclic orientations of graphs. *Discrete Math.* **5** (1973), 171–178.

[71] ———, A symmetric function generalization of the chromatic polynomial of a graph, *Adv. Math.* **111** (1995), 166 – 194.

[72] ———, Enumerative Combinatorics, Volumes 1 and 2. *Cambridge Studies in Advanced Mathematics*, **49** and **62**. Cambridge University Press, Cambridge, 1997 and 1999.

[73] Shishuo Fu, V. Reiner, Dennis Stanton, Nathaniel Thiem. The negative q-binomial. *The Electronic Journal of Combinatorics* **19**, Issue 1 (2012), P36.

[74] R. Steinberg, A geometric approach to the representations of the full linear group over a Galois field. *Trans. Amer. Math. Soc.* **71**, (1951), 274-282.

[75] J. Stembridge. A concise proof of the Littlewood-Richardson rule. *The Electronic Journal of Combinatorics* **9**, 2002, N5.

[76] M.E. Sweedler, Hopf algebras. W.A. Benjamin, New York, 1969.

[77] M. Takeuchi, Free Hopf algebras generated by coalgebras. *J. Math. Soc. Japan* **23** (1971), 561 – 582.

[78] H. Tamvakis, The theory of Schur polynomials revisited. *Enseign. Math.* **58** (2012), 147–163.

[79] J.-Y. Thibon. An Introduction to Noncommutative Symmetric Functions. Cargese lecture, October 2005. J.-P. Gazeau, J. Nesetril, B. Rovan (eds.): From Numbers and Languages to (Quantum) Cryptography, *NATO Security through Science Series: Information and Communication Security* **7**, IOS Press, 2007. Available at `http://igm.univ-mlv.fr/~jyt/ARTICLES/cargese_thibon.ps`

[80] N. Thiem and C.R. Vinroot, On the characteristic map of finite unitary groups. *Advances in Mathematics* **210**, Issue 2, 1 April 2007, pp. 707–732.

[81] A.V. Zelevinsky. Representations of finite classical groups: a Hopf algebra approach. *Lecture Notes in Mathematics* **869**. Springer-Verlag, Berlin-New York, 1981.

E-mail address: `darijgrinberg@gmail.com`

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

E-mail address: `reiner@math.umn.edu`

School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA