Symmetric Functions

Alistair Savage

Department of Mathematics and Statistics

University of Ottawa

This work is licensed under a
Creative Commons Attribution-ShareAlike 4.0 International License

Contents

Preface 4

1 Preliminaries 5
- 1.1 The symmetric group . . . 5
- 1.2 Partitions and compositions . . . 6
- 1.3 Graded rings . . . 10
- 1.4 Formal power series . . . 14

2 The ring of symmetric functions 20
- 2.1 Symmetric polynomials . . . 20
- 2.2 Symmetric functions . . . 23
- 2.3 Tableaux . . . 27
- 2.4 Elementary symmetric functions . . . 29
- 2.5 Complete homogeneous symmetric functions . . . 33
- 2.6 Power sums . . . 37

3 Schur functions 42
- 3.1 Alternating polynomials . . . 42
- 3.2 Schur polynomials and Schur functions . . . 46
- 3.3 The Hall inner product . . . 51
- 3.4 Skew Schur functions and semistandard tableaux . . . 56
- 3.5 Transition matrices . . . 64
- 3.6 The Littlewood-Richardson rule . . . 71
- 3.7 The Murnaghan–Nakayama rule . . . 78

4 Adjoint operators and Hopf algebras 81
- 4.1 Tensor products . . . 81
- 4.2 Adjoint operators . . . 85
- 4.3 Hopf algebras . . . 91
- 4.4 Hopf algebra structure on the ring of symmetric functions . . . 95

5 The boson-fermion correspondence 102
- 5.1 The Heisenberg algebra . . . 102
- 5.2 Bosonic Fock space . . . 106
- 5.3 Fermionic Fock space . . . 108

5.4 Bosonization 114
5.5 Fermionization 117

6 Applications to representation theory **126**
6.1 Characters of finite groups 126
6.2 Induction and restriction 130
6.3 Characters of symmetric groups 134
6.4 Specht modules 142
6.5 Representations of general linear groups 147

Index of notation **150**

Preface

These are notes for the topics course *Symmetric Functions* (MAT 4995/5327) at the University of Ottawa, which took place in Winter 2022. We first study the basics of symmetric functions, including the important topic of Schur functions. We essentially cover the material found in the first chapter of [Mac15], but we often give more detail and assume less background from the reader. We then cover some more advanced topics, including the Hopf algebra structure on the ring of symmetric functions, the boson-fermion correspondence, and applications to the representation theory of the symmetric groups and general linear groups.

We assume that the reader has a basic knowledge of the theory of rings and modules as covered, for example, in Rings and Modules (MAT 3143), or its French equivalent Anneaux et Modules (MAT 3543), at the University of Ottawa.

I would like to thank the students of the course for making it an enjoyable experience. I also thank Darij Grinberg for very useful comments and corrections.

Alistair Savage

Course website: https://alistairsavage.ca/symfunc/

Chapter 1

Preliminaries

In this first chapter we cover some background that will be used throughout the course. In particular, we will discuss the symmetric group, partitions, compositions, graded rings, and formal power series. Our discussion will be brief. Throughout these notes we let

$$\mathbb{N} := \{0, 1, 2, \dots\}$$

denote the set of natural numbers, including zero.

1.1 The symmetric group

We begin recalling a few important facts about the symmetric group. For $n \in \mathbb{N}$, let $\mathfrak{S}_n$ denote the *symmetric group* on n letters. Thus $\mathfrak{S}_n$ is the group of permutations of

$$\{1, 2, \dots, n\}.$$

By convention $\mathfrak{S}_0$ is the trivial group. The group $\mathfrak{S}_n$ has order $n!$.

If $\pi \in \mathfrak{S}_n$, we can write it in *one-line notation as*

$$\pi = \pi(1)\,\pi(2)\,\cdots\,\pi(n).$$

Alternatively, we may write it in *cycle notation*, which describes the effect of repeatedly applying the permutation to the elements of $\{1, 2, \dots, n\}$.

Example 1.1.1. Consider the element $\pi \in \mathfrak{S}_6$ given by

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

Then in one-line notation we have

$$\pi = 6\,1\,4\,3\,5\,2$$

and in cycle notation we have

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

We sometimes omit cycles of length one in cycle notation, writing

$$\pi = (1\,6\,2)(3\,4).$$

The inverse of an element can be given easily in cycle notation by reversing the order of the cycles. For example,

$$\pi^{-1} = (2\,6\,1)(4\,3). \qquad \triangle$$

Note that we have a natural inclusion $\{1, 2, \dots, n\} \subseteq \{1, 2, \dots, n+1\}$. Thus, we can view $\mathfrak{S}_n$ as the groups of $\mathfrak{S}_{n+1}$ that fixes $n+1$. In this way we have a chain of groups

$$\mathfrak{S}_0 \subseteq \mathfrak{S}_1 \subseteq \mathfrak{S}_2 \subseteq \mathfrak{S}_3 \subseteq \cdots.$$

Define

$$\mathfrak{S}_\infty := \bigcup_{n \in \mathbb{N}} \mathfrak{S}_n.$$

Thus, $\mathfrak{S}_\infty$ is the set of all bijections $\pi\colon \{1, 2, 3, \dots\} \to \{1, 2, 3, \dots\}$ such that there exists $n \in \mathbb{N}$ with $\pi(m) = m$ for $m \geq n$.

Exercises.

1.1.1. Give an example of a bijection $\pi\colon \{1, 2, 3, \dots\} \to \{1, 2, 3, \dots\}$ such that $\pi \notin \mathfrak{S}_\infty$.

1.1.2. Suppose $\pi \in \mathfrak{S}_k$ and that, for $i \in \mathbb{N}$, π has m_i cycles of length i. Prove that the number of permutations $\sigma \in \mathfrak{S}_k$ that commute with π (that is, such that $\sigma\pi = \pi\sigma$) is

$$\prod_{i \geq 1} i^{m_i} m_i!.$$

Hint: Write π in cycle notation. What are the cycles of $\sigma\pi\sigma^{-1}$?

1.2 Partitions and compositions

In the theory of symmetric functions, many mathematical objects are labelled by partitions and compositions. In this section we define these concepts and prove some basic results concerning them that will be used in the sequel.

Definition 1.2.1 (Partition)**.** A *partition* λ of a nonnegative integer $k \in \mathbb{N}$ is a weakly decreasing sequence $(\lambda_j)_{j=1}^\infty$ of nonnegative integers such that $\sum_{j=1}^\infty \lambda_j = k$. If λ is a partition of k, then we often write $\lambda \vdash k$ and $|\lambda| = k$, and we sometimes call k the *size* of λ. We let $\mathrm{Par}(k)$ denote the set of partitions of k and let $\mathrm{Par} := \bigcup_{k \in \mathbb{N}} \mathrm{Par}(k)$ denote the set of all partitions. We denote the unique partition $(0, 0, \dots)$ of 0 by $\varnothing$.

A partition has finitely many nonzero terms, and we typically omit zero terms when writing λ. For example

$$(5, 2, 2, 1), \qquad (5, 2, 2, 1, 0), \qquad (5, 2, 2, 1, 0, 0), \qquad \text{and} \quad (5, 2, 2, 1, 0, 0, \dots)$$

are all the same partition of 10, which we will usually write as $(5, 2, 2, 1)$. The nonzero terms in a partition λ are called the *parts* of λ. The *length* of λ, denoted $\ell(\lambda)$, is the number of parts of λ. For $n \in \mathbb{N}$, we let

$$\mathrm{Par}_n(k) := \{\lambda \in \mathrm{Par}(k) : \ell(\lambda) \leq n\} \qquad \text{and} \qquad \mathrm{Par}_n := \bigcup_{k \in \mathbb{N}} \mathrm{Par}_n(k).$$

We will often view elements λ of Par_n as weakly decreasing sequences $(\lambda_1, \ldots, \lambda_n)$ of length n, since we have $\lambda_j = 0$ for $j > n$ by definition. It follows immediately from the definitions that

$$\mathrm{Par}_n(k) \subseteq \mathrm{Par}_m(k) \qquad \text{and} \qquad \mathrm{Par}_n \subseteq \mathrm{Par}_m \qquad \text{for} \quad n \leq m$$

and that

$$\mathrm{Par}(k) = \bigcup_{n \in \mathbb{N}} \mathrm{Par}_n(k) \qquad \text{and} \qquad \mathrm{Par} = \bigcup_{n \in \mathbb{N}} \mathrm{Par}_n.$$

For any partition λ and $i \geq 1$, the *multiplicity of i in λ*, denoted $m_i(\lambda)$, is the number of parts of λ equal to i. In other words,

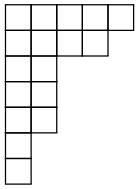
$$m_i(\lambda) = |\{j \geq 1 : \lambda_j = i\}|,$$

where $|X|$ denotes the cardinality of a finite set X. We will sometimes write a partition by writing each of its distinct parts once, with an exponent indicating the multiplicity of that part. For example we might write

$$(8, 8, 7, 5, 5, 5, 5, 3, 2, 2, 1, 1, 1) \qquad \text{as} \qquad (8^2, 7^1, 5^4, 3^1, 2^2, 1^3).$$

We will often think about partitions geometrically as follows. The *Young diagram*, also known as the *Ferrers diagram*, of a partition λ is a stack of boxes in which each row is left-justified and the j-th row has λ_j boxes. We will use the English convention, in which the first row is the top row. Thus, the Young diagram of the partition $(5, 4, 2, 2, 2, 1, 1)$ is

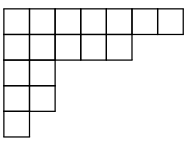
(1.1)



.

Some references use the French convention, where the first row is the bottom row. (There is also a Russian convention!) We will often identify a partition with its Young diagram.

The *conjugate* of a partition λ is the partition λ' whose Young diagram is the transpose of the Young diagram of λ, that is, the Young diagram obtained by reflection in the main diagonal. For example if $\lambda = (5, 4, 2, 2, 2, 1, 1)$, whose Young diagram is given in (1.1), then the conjugate partition $\lambda' = (7, 5, 2, 2, 1)$ has Young diagram



.

It follows that λ'_j is the number of boxes in the j-th column of λ, or equivalently

$$\lambda'_j = |\{i : \lambda_i \geq j\}|. \tag{1.2}$$

In particular, we have $\lambda'_1 = \ell(\lambda)$ and $\lambda_1 = \ell(\lambda')$. We also have $\lambda'' = \lambda$.

For $k \in \mathbb{N}$, we define the *dominance partial ordering* (sometimes called the *natural partial ordering*) $\leq$ on $\mathrm{Par}(k)$ as follows: for $\lambda, \mu \in \mathrm{Par}$,

$$\lambda \leq \mu \iff \lambda_1 + \cdots + \lambda_i \leq \mu_1 + \cdots + \mu_i \qquad \text{for all } i \geq 1. \tag{1.3}$$

We write $\lambda < \mu$ if $\lambda \leq \mu$ and $\lambda \neq \mu$.

Example 1.2.2. If $k = 6$, we have

$$(1^6) < (2, 1^4) \qquad \text{and} \qquad (2, 2, 2) < (4, 2).$$

However, $(3, 1^3)$ and (2^3) are not comparable. △

Definition 1.2.3 (Weak composition)**.** A *weak composition* α of $k \in \mathbb{N}$ is a sequence $(\alpha_j)_{j=1}^{\infty}$ of nonnegative integers such that $\sum_{j=1}^{\infty} \alpha_j = k$. (By contrast, a *composition* of $k \in \mathbb{N}$ is a finite sequence $(\alpha_j)_{j=1}^{n}$ of *positive* integers such that $\sum_{j=1}^{n} \alpha_j = k$.) We let $\mathrm{WComp}(k)$ denote the set of weak compositions of k and let $\mathrm{WComp} = \bigcup_{k \in \mathbb{N}} \mathrm{WComp}(k)$ denote the set of all weak compositions. We denote the unique composition $(0, 0, \dots)$ of 0 by $\varnothing$. For $\alpha \in \mathrm{WComp}$, we define

$$|\alpha| = \sum_{j=1}^{\infty} \alpha_j.$$

For $k, n \in \mathbb{N}$, we define

$$\mathrm{WComp}_n(k) := \{\alpha \in \mathrm{WComp}(k) : \alpha_j = 0 \text{ for } j > n\} \qquad \text{and}$$
$$\mathrm{WComp}_n := \bigcup_{k \in \mathbb{N}} \mathrm{WComp}_n(k).$$

As for partitions, we will often view elements α of WComp_n as sequences $(\alpha_1, \dots, \alpha_n)$ of length n. We have

$$\mathrm{WComp}_m(k) \subseteq \mathrm{WComp}_n(k) \qquad \text{and} \qquad \mathrm{WComp}_m \subseteq \mathrm{WComp}_n \qquad \text{for} \quad m \leq n$$

and

$$\mathrm{WComp}(k) = \bigcup_{n \in \mathbb{N}} \mathrm{WComp}_n(k) \qquad \text{and} \qquad \mathrm{WComp} = \bigcup_{n \in \mathbb{N}} \mathrm{WComp}_n.$$

Note that every partition is a weak composition but not the other way around since we require partitions to be *weakly decreasing* sequences. For example $(2, 0, 1, 1)$ is a weak composition of 4, but is not a partition of 4.

The symmetric group $\mathfrak{S}_n$ acts on the set $\mathrm{WComp}_n(k)$ as follows. For $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\pi \in \mathfrak{S}_n$, define

$$\alpha^{\pi} := (\alpha_{\pi(1)}, \alpha_{\pi(2)}, \dots, \alpha_{\pi(n)}). \tag{1.4}$$

This defines a right action of $\mathfrak{S}_n$ on $\mathrm{WComp}_n(k)$; see Exercise 1.2.1. For $\alpha \in \mathrm{WComp}_n(k)$, we define

$$\alpha\mathfrak{S}_n := \{\alpha^\pi : \pi \in \mathfrak{S}_n\}. \tag{1.5}$$

(We use the notation $\alpha\mathfrak{S}_n$ instead of $\mathfrak{S}_n\alpha$ since we have a *right* action of $\mathfrak{S}_n$.) For every $\alpha \in \mathrm{WComp}_n(k)$, there exists a permutation $\pi \in \mathfrak{S}_n$ such that $\alpha^\pi \in \mathrm{Par}_n(k)$. (This permutation π is unique if and only if α has no repeated terms.) It follows that

$$\bigcup_{\alpha\in\mathrm{Par}_n(k)} \alpha\mathfrak{S}_n = \{\alpha^\pi : \alpha \in \mathrm{Par}_n(k),\ \pi \in \mathfrak{S}_n\} = \mathrm{WComp}_n(k).$$

Similarly, the set $\mathfrak{S}_\infty$ acts on the set $\mathrm{WComp}(k)$. Precisely, for $\alpha = (\alpha_1, \alpha_2, \dots) \in \mathrm{WComp}(k)$ and $\pi \in \mathfrak{S}_\infty$, we define

$$\alpha^\pi := (\alpha_{\pi(1)}, \alpha_{\pi(2)}, \dots) \qquad \text{and} \qquad \alpha\mathfrak{S}_\infty = \{\alpha^\pi : \pi \in \mathfrak{S}_\infty\}.$$

For every $\alpha \in \mathrm{WComp}(k)$, there exists a (not unique!) permutation $\pi \in \mathfrak{S}_\infty$ such that $\alpha^\pi \in \mathrm{Par}(k)$ and we have a disjoint union

$$\bigsqcup_{\lambda\in\mathrm{Par}(k)} \lambda\mathfrak{S}_\infty = \{\lambda^\pi : \lambda \in \mathrm{Par}(k),\ \pi \in \mathfrak{S}_\infty\} = \mathrm{WComp}(k). \tag{1.6}$$

The dominance partial ordering (1.3) can be naturally extended to a partial order on weak compositions. (Note that we will still reserve the term *dominance partial ordering* to refer to partitions only.) Namely, for $\alpha, \beta \in \mathrm{WComp}$, we define

$$\alpha \le \beta \iff \alpha_1 + \cdots + \alpha_i \le \beta_1 + \cdots + \beta_i \qquad \text{for all} \quad i \in \mathbb{N}. \tag{1.7}$$

Lemma 1.2.4. *Suppose $\alpha \in \mathrm{WComp}$. Then*

$$\alpha \in \mathrm{Par} \iff \alpha^\pi \le \alpha \quad \textit{for all} \quad \pi \in \mathfrak{S}_\infty.$$

Proof. Suppose that $\alpha \in \mathrm{Par}$. Then $\alpha_1 \ge \alpha_2 \ge \cdots$. Suppose $\beta = \alpha^\pi$ for some $\pi \in \mathfrak{S}_\infty$. For $i \in \mathbb{N}$, $\alpha_1 + \cdots + \alpha_i$ is the sum of the i greatest elements of α, which are the same as the i greatest elements of β. Thus

$$\beta_1 + \cdots + \beta_i \le \alpha_1 + \cdots + \alpha_i.$$

Hence $\alpha^\pi = \beta \le \alpha$.

Conversely, suppose $\alpha^\pi \le \alpha$ for all $\pi \in \mathfrak{S}_\infty$. Then, for $i \ge 1$, we have

$$(\alpha_1, \alpha_2, \dots, \alpha_{i-1}, \alpha_{i+1}, \alpha_i, \alpha_{i+2}, \alpha_{i+3}, \dots) \le \alpha.$$

It follows that, for $i \ge 1$,

$$\alpha_1 + \cdots + \alpha_{i-1} + \alpha_{i+1} \le \alpha_1 + \cdots + \alpha_{i-1} + \alpha_i$$

and so $\alpha_{i+1} \le \alpha_i$. Hence $\alpha \in \mathrm{Par}$. □

Exercises.

1.2.1. If G is a group, let G^{op} denote the opposite group. Thus,

$$G^{\mathrm{op}} = \{g^{\mathrm{op}} : g \in G\}, \quad g^{\mathrm{op}} h^{\mathrm{op}} = (hg)^{\mathrm{op}}, \quad g, h \in G.$$

A *right action* of a group G on a set X is a group homomorphism $G^{\mathrm{op}} \to \mathfrak{S}_X$, where $\mathfrak{S}_X$ denotes the group of permutations of X. By contrast, a *left action* of G on X is a group homomorphism $G \to \mathfrak{S}_X$.

(a) Show that (1.4) defines a right action of $\mathfrak{S}_n$ on $\mathrm{WComp}_n(k)$. In other words, show that

$$\alpha^{\pi_1 \pi_2} = (\alpha^{\pi_1})^{\pi_2}$$

for all $\alpha \in \mathrm{WComp}_n(k)$ and $\pi_1, \pi_2 \in \mathfrak{S}_n$.

(b) Show that, for $n \geq 3$, (1.4) does *not* define a left action of $\mathfrak{S}_n$ on $\mathrm{WComp}_n(k)$.

(c) How could you modify (1.4) to give a natural left action of $\mathfrak{S}_n$ on $\mathrm{WComp}_n(k)$?

1.2.2. Prove that, for $k \leq 5$, the dominance partial ordering on $\mathrm{Par}(k)$ is a *total* ordering.

1.2.3. A partition λ is *self-conjugate* if $\lambda' = \lambda$. Prove that, for $k \in \mathbb{N}$, the number of self-conjugate partitions of k is equal to the number of partitions of k into distinct, odd parts. *Hint*: Think about Young diagrams.

1.2.4. For $k \in \mathbb{N}$, define the *lexicographic ordering* $<_{\mathrm{lex}}$ on $\mathrm{Par}(k)$ as follows: For $\lambda, \mu \in \mathrm{Par}(k)$, we write $\lambda <_{\mathrm{lex}} \mu$ if and only if there exists some $i \in \mathbb{N}$ such that

$$\lambda_j = \mu_j \text{ for all } j < i, \quad \text{and } \lambda_i < \mu_i.$$

This is a total order on $\mathrm{Par}(k)$. Prove that, for $\lambda, \mu \in \mathrm{Par}(k)$,

$$\lambda < \mu \implies \lambda <_{\mathrm{lex}} \mu.$$

Thus, the lexicographic order is a refinement of the dominance ordering.

1.3 Graded rings

In these notes we will assume all rings are unital (i.e. they posses a multiplicative identity). In this section we discuss the notion of a graded ring, which plays a crucial role in theory of symmetric functions.

Definition 1.3.1 (Graded ring)**.** A *graded ring* is a ring R with a direct sum decomposition

$$R = \bigoplus_{k \in \mathbb{N}} R_k \tag{1.8}$$

such that each R_k is an additive subgroup of R, and such that

$$R_k R_l \subseteq R_{k+l} \quad \text{for all } k, l \in \mathbb{N}.$$

A nonzero element of R_k is said to be *homogeneous of degree* k. If R is any ring, we say that the decomposition (1.8) is a *gradation* on R if this decomposition makes it a graded ring.

Remarks 1.3.2. One can consider rings graded by any abelian monoid. Since our main example of interest is grading by $\mathbb{N}$, we stick to that case here. Later we will use superscripts to denote the gradation, since our rings will already have subscripts. △

Examples 1.3.3. (a) Any ring R can be given a gradation by letting $R_0 = R$ and $R_i = 0$ for $i > 0$. This is called the *trivial gradation* on R.

(b) The polynomial ring $\mathbb{Z}[x]$ is graded by degree. More precisely, we have a gradation $\mathbb{Z}[x] = \bigoplus_{k\in\mathbb{N}} \mathbb{Z}[x]_k$, where $\mathbb{Z}[x]_k$ consists of the homogeneous polynomials of degree k, together with the zero polynomial.

(c) The polynomial ring $\mathbb{Z}[x, y]$ is also graded by total degree. △

If R is a graded ring, then by the definition of direct sum, every element $r \in R$ can be written *uniquely* as a sum

$$r = \sum_{i\in\mathbb{N}} r_i, \qquad r_i \in R_i,$$

where only *finitely* many of the r_i are nonzero. The r_i are called the *homogeneous components of* r.

Definition 1.3.4 (Graded subring)**.** Suppose R is a graded ring. A *graded subring* of R is a subring S of R such that

$$S = \bigoplus_{k\in\mathbb{N}} S_k, \qquad S_k = S \cap R_k. \tag{1.9}$$

Equivalently, S is a graded subring of R if it is a subring of R and the homogeneous components of all elements of S are again elements of S. If S is a graded subring of R, then it is itself a graded ring with gradation (1.9).

Example 1.3.5. Recall that, if R is a ring and $X \subseteq R$, then the subring of R generated by X is the smallest subring of R containing X. Consider the case $R = \mathbb{Z}[x, y]$, graded by total degree. Then the subring generated by $x + y$, which consists of all polynomials in $x + y$ with integer coefficients, is a graded subring of R, while the subring generated by $x + y^2$ is *not* a graded subring of R. See Exercise 1.3.1. △

Definition 1.3.6. Suppose R is a graded ring. A *graded ideal* of R is an ideal I of R such that

$$I = \bigoplus_{k\in\mathbb{N}} I_k, \qquad I_k = I \cap R_k.$$

Equivalently, I is a graded ideal of R if it is an ideal of R and the homogeneous components of all elements of I are again elements of I.

Example 1.3.7. Consider the ring $\mathbb{Z}[x, y]$, graded by total degree. Then the ideal $I = x\mathbb{Z}[x, y]$ generated by x is a graded ideal of $\mathbb{Z}[x, y]$. Precisely, we have

$$I = \bigoplus_{k\in\mathbb{N}} I_k, \qquad I_k = I \cap \mathbb{Z}[x, y]_k = x\mathbb{Z}[x, y]_{k-1}.$$

However, the ideal $(x + 1)\mathbb{Z}[x, y]$ is *not* a graded ideal. See Exercise 1.3.2. △

Definition 1.3.8 (Homomorphism of graded rings)**.** Suppose R and S are graded rings. A *homomorphism of graded rings* from R to S is a ring homomorphism $\sigma\colon R \to S$ such that

$$\sigma(R_k) \subseteq S_k \quad \text{for all } k \in \mathbb{N}.$$

If, in addition, σ is an isomorphism, then we say it is an *isomorphism of graded rings*. A homomorphism of graded rings $\sigma\colon R \to R$ is an *endomorphism of graded rings*; it is an *automorphism of graded rings* if it is also an isomorphism.

If $\sigma\colon R \to S$ is a homomorphism of graded rings, then we have its *homogeneous components*

$$\sigma_k\colon R_k \to S_k, \qquad k \in \mathbb{N},$$

obtained from the restriction of σ to R_k. These are homomorphisms of additive groups. We often write $\sigma = \sum_{k\in\mathbb{N}} \sigma_k$ since

$$\sigma(r) = \sum_{k\in\mathbb{N}} \sigma_k(r_k), \qquad r \in R.$$

(Here $r = \sum_{k\in\mathbb{N}} r_k$ is the decomposition of r into its homogeneous components.) This sum is always well defined since only finitely many terms are nonzero.

Lemma 1.3.9. *If $\sigma\colon R \to S$ is a homomorphism of graded rings, then its kernel* $\ker\sigma$ *is a graded ideal of R.*

Proof. We know from ring theory that $\ker\sigma$ is an ideal of R. It remains to show that the homogeneous components of $\ker\sigma$ are again elements of $\ker\sigma$. Suppose $r = \sum_{k\in\mathbb{N}} r_k \in \ker\sigma$. Then we have

$$0 = \sigma(r) = \sum_{k\in\mathbb{N}} \sigma(r_k).$$

Since σ is a homomorphism of graded rings, we have $\sigma(r_k) \in S_k$ for all $k \in \mathbb{N}$. Since the sum $S = \bigoplus_{k\in\mathbb{N}} S_k$ is direct, this implies that $\sigma(r_k) = 0$ for all $k \in \mathbb{N}$. In other words $r_k \in \ker\sigma$ for all $k \in \mathbb{N}$, as desired. □

Example 1.3.10. Suppose R is a commutative ring and consider the polynomial ring $R[x]$, graded by degree. For any $a \in R$, we have the evaluation homomorphism

$$R[x] \to R, \quad f(x) \mapsto f(a).$$

Since x has degree 1 and a has degree zero if $a \neq 0$, the evaluation homomorphism is graded if and only if $a = 0$. △

Proposition 1.3.11. *If R is a graded ring and X is a set of graded ring automorphisms of R, then*

$$R^X := \{r \in R : \sigma(r) = r \text{ for all } \sigma \in X\}$$

is a graded subring of R.

Proof. We leave it as an exercise (Exercise 1.3.3) to verify that R^X is a subring of R. It remains to show that the homogenous components of elements of R^X are again elements of R^X. Suppose $r \in R^X$. Then, for $\sigma \in X$,

$$\sum_{k\in\mathbb{N}} r_k = r = \sigma(r) = \sum_{k\in\mathbb{N}} \sigma(r_k).$$

Since $\sigma(r_k) \in R_k$, it follows from the uniqueness of the decomposition of an element of R as a sum of homogeneous components that $\sigma(r_k) = r_k$. Hence $r_k \in R^X$ for all $k \in \mathbb{N}$, as desired. □

Many of the important results on rings can be generalized to the setting of graded rings. For example, the quotient of a graded ring by a graded ideal is naturally a graded ring (see Exercise 1.3.7) and one has graded versions of the isomorphism theorems for rings.

We conclude this section with a brief discussion of associative algebras, which are more general than rings. We fix a commutative ring $\Bbbk$, which is often called the *ground ring*. A *unital associative $\Bbbk$-algebra* R is a ring that is also a $\Bbbk$-module in such a way that the two additions (the ring addition and the module addition) are the same operation and the scalar multiplication satisfies

$$a(xy) = (ax)y = x(ay) \qquad \text{for all} \quad a \in \Bbbk,\ x, y \in R.$$

In these notes, the term *$\Bbbk$-algebra* (or simply *algebra*, if the ground ring $\Bbbk$ is clear from the context) will always mean unital associative $\Bbbk$-algebra. Since a $\mathbb{Z}$-module is the same as an abelian group, a $\mathbb{Z}$-algebra is simply a ring. This is the sense in which associative algebras generalize rings.

Examples 1.3.12. (a) The ring of polynomials $\Bbbk[x]$ with coefficients in $\Bbbk$ is a $\Bbbk$-algebra.

(b) For $n \geq 1$, the ring of matrices $\mathrm{Mat}_n(\Bbbk)$ with entries in $\Bbbk$ is a $\Bbbk$-algebra.

(c) If G is a finite group, the group ring $\Bbbk G$ is a $\Bbbk$-algebra. △

A *graded $\Bbbk$-algebra* is a $\Bbbk$-algebra R with a direct sum decomposition (1.8) such that each R_k is a $\Bbbk$-submodule of R, and such that $R_k R_l \subseteq R_{k+l}$ for all $k, l \in \mathbb{N}$. All of the results of this section on graded rings have natural analogues for $\Bbbk$-algebras.

Examples 1.3.13. We can repeat Example 1.3.3 with $\Bbbk$-algebras in place of rings.

(a) Any $\Bbbk$-algebra R can be given a gradation by letting $R_0 = R$ and $R_i = 0$ for $i > 0$. This is called the *trivial gradation* on R.

(b) The $\Bbbk$-algebras $\Bbbk[x]$ and $\Bbbk[x, y]$ are graded by degree. △

Exercises.

1.3.1. Consider the ring $\mathbb{Z}[x, y]$ graded by total degree.

(a) Show that the subring generated by $x + y$ is a graded subring of R.

(b) Show that the subring generated by $x + y^2$ is not a graded subring of R.

1.3.2. Prove that $(x + 1)\mathbb{Z}[x, y]$ is not a graded ideal of $\mathbb{Z}[x, y]$, where the gradation is given by total degree.

1.3.3. Suppose R is a ring and X is a set of ring automorphisms. Show that

$$R^X := \{r \in R : \sigma(r) = r \text{ for all } \sigma \in X\}$$

is a subring of R.

1.3.4. Prove that the multiplicative unit in a graded ring is homogeneous of degree zero.

1.3.5. Suppose u is a homogeneous invertible element in a graded ring R. Show that u has degree zero.

1.3.6. Suppose u is a homogeneous invertible element in a graded ring R. Show that the map

$$\rho_u \colon R \to R, \qquad \rho_u(r) = uru^{-1}, \quad r \in R,$$

is a graded ring automorphism of R.

1.3.7. Suppose that I is a graded ideal of a graded ring R. Show that $\bigoplus_{k\in\mathbb{N}} R_k/I_k$ has a natural induced structure of a graded ring.

1.4 Formal power series

In this section we briefly introduce the concept of formal power series. We will use this formalism in the sequel to concisely encode relations between various sequences of symmetric functions. Throughout this section R denotes a commutative ring.

The *ring of formal power series with coefficients in* R is

$$R[\![t]\!] := \left\{ \sum_{n=0}^{\infty} a_n t^n : a_n \in R \text{ for all } n \in \mathbb{N} \right\}.$$

More precisely, elements of $R[\![t]\!]$ are functions $f \colon \mathbb{N} \to R$, where we write such a function as $\sum_{n=0}^{\infty} f(n)t^n$. The unit element of R is 1. (Here we are writing 1 to denote the formal

power series $\sum a_n t^n$, where $a_0 = 1$, the multiplicative identity of R, and $a_n = 0$ for $r > 0$.) Addition and multiplication are given by

$$\begin{aligned} \sum_{n=0}^{\infty} a_n t^n + \sum_{n=0}^{\infty} b_n t^n &= \sum_{n=0}^{\infty} (a_n + b_n) t^n, \\ \left(\sum_{n=0}^{\infty} a_n t^n \right) \left(\sum_{n=0}^{\infty} b_n t^n \right) &= \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} a_i b_{n-i} \right) t^n. \end{aligned} \tag{1.10}$$

Note that, for the multiplication of two formal power series, the coefficient of each t^n involves a *finite* sum and hence is well defined.

The use of the word *formal* in our terminology comes from the fact that we do not worry ourselves with issues of convergence. The formal power series $\sum_{n=0}^{\infty} a_n t^n$ is sometimes called the *generating function* of the sequence $(a_n)_{n=0}^{\infty}$. Note that we have a natural inclusion of rings

$$R[t] \subseteq R[\![t]\!].$$

We denote by $R^\times$ the group of units of R:

$$R^\times = \{a \in R : ab = 1 \text{ for some } b \in R\}.$$

The following proposition describes the units of $R[\![t]\!]$.

Proposition 1.4.1. *We have*

$$R[\![t]\!]^\times = \left\{ \sum_{n=0}^{\infty} a_n t^n \in R[\![t]\!] : a_0 \in R^\times \right\}.$$

Proof. Let $a(t) = \sum_{n=0}^{\infty} a_n t^n \in R[\![t]\!]$. If $a(t) \in R[\![t]\!]^\times$, then there exists $\sum_{n=0}^{\infty} b_n t^n \in R[\![t]\!]$ such that

$$1 = \left(\sum_{n=0}^{\infty} a_n t^n \right) \left(\sum_{n=0}^{\infty} b_n t^n \right) = a_0 b_0 + \sum_{n=1}^{\infty} \left(\sum_{i=0}^{n} a_i b_{n-i} \right) t^n.$$

It follows that $a_0 b_0 = 1$, and hence $a_0 \in R^\times$.

Conversely, suppose that $a_0 \in R^\times$. Then we define a sequence $(b_n)_{n=0}^{\infty}$ recursively by

$$b_0 = a_0^{-1}, \qquad b_n = -a_0^{-1} \sum_{i=1}^{n} a_i b_{n-i}, \quad n \geq 1.$$

It is then straightforward to verify that

$$\left(\sum_{n=0}^{\infty} a_n t^n \right) \left(\sum_{n=0}^{\infty} b_n t^n \right) = 1. \qquad \square$$

Example 1.4.2. Suppose $a \in R$. Then, by Proposition 1.4.1, $1 - at \in R[\![t]\!]^\times$. We leave it as an exercise (Exercise 1.4.3) to show that

$$(1 - at)^{-1} = \sum_{n=0}^{\infty} a^n t^n. \tag{1.11}$$ $\triangle$

For $n \in \mathbb{N}$, let $p(n)$ denote the number of partitions of n. No simple explicit formula for $p(n)$ is known. However, the sequence $(p(n))_{n=0}^{\infty}$ satisfies some recurrence relations that can be used to efficiently compute $p(n)$. As the following result shows, these recurrence relations can be encoded as a product expression for the generating function of the sequence.

Proposition 1.4.3. *If $p(n;k)$ denotes the number of partitions $\lambda \vdash n$ with largest part $\lambda_1 \leq k$, then*

$$\sum_{n=0}^{\infty} p(n;k)t^n = \prod_{j=1}^{k} \frac{1}{1-t^j}. \tag{1.12}$$

Proof. We prove the result by induction on k. When $k = 0$, both sides of (1.12) are equal to 1. Now, for $k \geq 1$, we have

$$p(n;k) = p(n;k-1) + p(n-k;k-1) + p(n-2k;k-1) + \cdots,$$

where $p(n-lk;k-1)$ counts the number of partitions $\lambda \vdash n$ with $m_k(\lambda) = l$. Thus

$$p(n;k)t^n = p(n;k-1)t^n + t^k p(n-k;k-1)t^{n-k} + t^{2k} p(n-2k;k-1)t^{n-2k} + \cdots.$$

Summing over n gives

$$\sum_{n=0}^{\infty} p(n;k)t^n = \left(1 + t^k + t^{2k} + \cdots\right) \sum_{n=0}^{\infty} p(n;k-1)t^n = \prod_{j=1}^{k} \frac{1}{1-t^j},$$

where we used the induction hypothesis in the last step. □

Taking the limit at $k \to \infty$ then gives a generating function for the sequence $(p(n))_{n=0}^{\infty}$.

Corollary 1.4.4. *We have*

$$\sum_{n=0}^{\infty} p(n)t^n = \prod_{j=1}^{\infty} \frac{1}{1-t^j}. \tag{1.13}$$

Proof. First note that the right-hand side of (1.13) a well-defined formal power series, even though it involves an infinite product. This is because, to compute the coefficient of t^k in $\prod_{j=1}^{\infty} \frac{1}{1-t^j}$, only the factors $\frac{1}{1-t^j}$ for $j \leq k$ contribute. Then the result follows from (1.12) and the fact that $p(n;k) = p(n)$ for $k \geq n$. □

Our discussion of formal power series can be generalized to any number of indeterminates, even infinitely many. Since the case of countably many indeterminates $x_1, x_2, \ldots$ will be important for us, let us discuss it here. It is useful to use *multi-index notation* for monomials. For $\alpha = (\alpha_1, \alpha_2, \ldots) \in \mathrm{WComp}$, we define the monomial

$$x^{\alpha} := \prod_{i=1}^{\infty} x_i^{\alpha_i}. \tag{1.14}$$

Note that since all but finitely many of the exponents α_i are equal to zero, this is actually a finite product. The degree of this monomial is $|\alpha|$. For example

$$x^{(5,3,0,4)} = x_1^5 x_2^3 x_4^4 \quad \text{has degree} \quad |(5,3,0,4)| = 12.$$

Define

$$R[\![x_1, x_2, \dots]\!] := \left\{ \sum_{\alpha \in \mathrm{WComp}} a_\alpha x^\alpha : a_\alpha \in R \text{ for all } \alpha \in \mathrm{WComp} \right\}. \tag{1.15}$$

This is a ring with multiplication given by

$$\left(\sum_{\alpha \in \mathrm{WComp}} a_\alpha x^\alpha \right) \left(\sum_{\beta \in \mathrm{WComp}} b_\beta x^\beta \right) = \sum_{\alpha, \beta \in \mathrm{WComp}} a_\alpha b_\beta x^{\alpha+\beta}, \tag{1.16}$$

where $\alpha + \beta = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots)$. The infinite sum on the right-hand side of (1.16) is well defined since, for any $\gamma \in \mathrm{WComp}$, there exists only finitely many pairs (α, β) such that $\alpha + \beta = \gamma$, and so the coefficient of x^γ is finite.

For $k \in \mathbb{N}$, define

$$R[\![x_1, x_2, \dots]\!]^k := \left\{ \sum_{\alpha \in \mathrm{WComp}(k)} a_\alpha x^\alpha : a_\alpha \in R \text{ for all } \alpha \in \mathrm{WComp}(k) \right\}. \tag{1.17}$$

This is the R-submodule of $R[\![x_1, x_2, \dots]\!]$ consisting of zero and formal power series of degree k. Note that

$$R[\![x_1, x_2, \dots]\!] \neq \bigoplus_{k \in \mathbb{N}} R[\![x_1, x_2, \dots]\!]^k, \tag{1.18}$$

since the right-hand side consists of formal power series of bounded total degree. For example,

$$\sum_{n=0}^{\infty} x_1^n \in R[\![x_1, x_2, \dots]\!] \qquad \text{but} \qquad \sum_{n=0}^{\infty} x_1^n \notin \bigoplus_{k \in \mathbb{N}} R[\![x_1, x_2, \dots]\!]^k.$$

Exercises.

Throughout these exercises, R denotes a commutative ring.

1.4.1. The ring of *formal Laurent series with coefficients in* R is

$$R(\!(t)\!) = \left\{ \sum_{n=N}^{\infty} a_n t^n : N \in \mathbb{Z},\ a_n \in R \text{ for all } n \in \mathbb{N} \right\}.$$

Note that the exponents n appearing in any given element of $R(\!(t)\!)$ are bounded below (by some $N \in \mathbb{Z}$), but we allow arbitrarily low bounds.

(a) Show that $R(\!(t)\!)$ is a ring with the usual addition and multiplication (that is, as in (1.10), but with lower bound N in the sums).

(b) Show that

$$\left\{ \sum_{n=-\infty}^{\infty} a_n t^n : a_n \in R \text{ for all } n \in \mathbb{Z} \right\}$$

is *not* a ring under the usual addition and multiplication.

1.4.2. Is the ring $R[\![t]\!]$ graded by degree?

1.4.3. Prove (1.11).

1.4.4. Recall that the *Fibonacci sequence* $(f_n)_{n=0}^{\infty}$ is defined recursively by

$$f_0 = f_1 = 1, \qquad f_n = f_{n-1} + f_{n-2} \quad \text{for } n \geq 2.$$

(a) Use the recursion relation to find a generating function for the Fibonacci sequence.

(b) Show that, for $n \in \mathbb{N}$,

$$f_n = \frac{a^{n+1} - b^{n+1}}{\sqrt{5}}, \qquad \text{where} \quad a = \frac{1+\sqrt{5}}{2},\ b = \frac{1-\sqrt{5}}{2}.$$

The next four exercises aim to justify some manipulations we will do with power series in the sequel. For the remainder of the exercises we work over a commutative $\mathbb{Q}$-algebra $\Bbbk$ (for example, a field of characteristic zero, $\mathbb{Q}[x_1, \ldots, x_n]$, or $\mathrm{Sym}_{\mathbb{Q}}$) and fix a formal power series $A(t) = \sum_{n \geq 0} a_n t^n \in \Bbbk[\![t]\!]$.

1.4.5. If $a_0 = 1$, show that

$$A(t)^{-1} = \sum_{n=0}^{\infty} (1 - A(t))^n.$$

Hint: Write $A(t) = 1 - (1 - A(t))$ and compare to (1.11).

1.4.6. Recall the Taylor series for the natural logarithm: $\log(1+x) = \sum_{n \geq 1} (-1)^{n-1} \frac{x^n}{n}$. If $a_0 = 1$ (so that the constant term of $A(t)$ is 1), we define

$$\log A(t) := \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(A(t)-1)^n}{n}.$$

Prove that this is indeed a well-defined formal power series.

1.4.7. Define the formal derivative of $A(t)$ to be

$$A'(t) = \frac{\mathrm{d}}{\mathrm{d}t} A(t) := \sum_{n \geq 1} n a_n t^{n-1}.$$

If $a_0 = 1$, prove that

$$\frac{\mathrm{d}}{\mathrm{d}t} \log A(t) = A'(t)/A(t).$$

1.4.8. (a) Suppose $a_0 = 0$. Prove that

$$\exp A(t) := \sum_{n\geq 0} \frac{1}{n!} A(t)^n$$

is a well-defined formal power series and that its constant term is 1. Furthermore, show that

$$\exp \sum_{n\geq 1} a_n t^n = \prod_{n\geq 1} \exp\left(a_n t^n\right).$$

(b) Show that, if $a_0 = 0$, then

$$\log \exp A(t) = A(t).$$

(c) Show that if $a_0 = 1$, then

$$\exp \log A(t) = A(t).$$

Chapter 2

The ring of symmetric functions

In this chapter we introduce the main object of study in the course: the ring of symmetric functions. We will give two descriptions of this ring, one as a certain type of limit and another more direct definition. Both definitions will be useful in what follows. After defining the ring of symmetric functions, we will introduce several important bases and generating sets for this ring.

2.1 Symmetric polynomials

Let

$$\mathrm{Pol}_n := \mathbb{Z}[x_1, \ldots, x_n]$$

denote the ring of polynomials in n indeterminates with integer coefficients. This is a unique factorization domain. It is also a graded ring with gradation

$$\mathrm{Pol}_n = \bigoplus_{k \geq 0} \mathrm{Pol}_n^k,$$

where Pol_n^k consists of homogeneous polynomials of degree k, together with the zero polynomial. (Note that we use superscripts to indicate the grading, instead of subscripts as in Section 1.3.)

It is useful to use *multi-index notation* for monomials. For $\alpha \in \mathrm{WComp}_n$, we define the monomial

$$x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}.$$

The degree of this monomial is

$$|\alpha| := \sum_{i=1}^{n} \alpha_i.$$

The monomials x^α, $\alpha \in \mathrm{WComp}_n$, form a $\mathbb{Z}$-basis for Pol_n. Similarly, the monomials x^α, $\alpha \in \mathrm{WComp}_n(k)$, form a basis for Pol_n^k.

For $\pi \in \mathfrak{S}_n$, the map

$$f \mapsto \pi f := f(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}), \qquad f \in \mathrm{Pol}_n, \tag{2.1}$$

is a ring automorphism of Pol_n. This defines a left action of $\mathfrak{S}_n$ on Pol_n; see Exercise 2.1.1. We say that a polynomial $f \in \mathrm{Pol}_n$ is *symmetric* if it is invariant under the action of $\mathfrak{S}_n$. In other words, f is symmetric if

$$\pi f = f \quad \text{for all } \pi \in \mathfrak{S}_n.$$

The *ring of symmetric polynomials* is the subring of Pol_n consisting of symmetric polynomials:

$$\mathrm{Sym}_n := \mathrm{Pol}_n^{\mathfrak{S}_n} := \{f \in \mathrm{Pol}_n \text{ such that } \pi f = f \text{ for all } \pi \in \mathfrak{S}_n\}.$$

By Proposition 1.3.11, Sym_n is a graded ring:

$$\mathrm{Sym}_n := \bigoplus_{k\in\mathbb{N}} \mathrm{Sym}_n^k, \qquad \mathrm{Sym}_n^k := (\mathrm{Pol}_n^k)^{\mathfrak{S}_n}. \tag{2.2}$$

Let $\lambda \in \mathrm{Par}_n$ and recall the definition of $\lambda\mathfrak{S}_n$ from (1.5). The corresponding *monomial symmetric polynomial* is

$$m_\lambda(x_1, \ldots, x_n) := \sum_{\alpha\in\lambda\mathfrak{S}_n} x^\alpha. \tag{2.3}$$

Remark 2.1.1. Note that (2.3) is *not* equal to $\sum_{\pi\in\mathfrak{S}_n} x^{\alpha^\pi}$ in general since there may be $\pi_1, \pi_2 \in \mathfrak{S}_n$ with $\alpha^{\pi_1} = \alpha^{\pi_2}$, and we only want to count these terms once. For example

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

but

$$\sum_{\pi\in\mathfrak{S}_4} x^{(3,2,2)^\pi} = 2m_{(3,2,2)}(x_1, x_2, x_3, x_4).$$

△

We have

$$m_\lambda(x_1, \ldots, x_n) \in \mathrm{Sym}_n^k \qquad \text{for} \quad \lambda \in \mathrm{Par}_n(k).$$

Define

$$m_\lambda(x_1, \ldots, x_n) = 0 \qquad \text{for} \quad \ell(\lambda) > n.$$

Proposition 2.1.2. *The monomial symmetric polynomials*

$$m_\lambda(x_1, \ldots, x_n), \qquad \lambda \in \mathrm{Par}_n(k), \tag{2.4}$$

form a $\mathbb{Z}$-basis of Sym_n^k.

Proof. We must show that the elements (2.4) are linearly independent and span Sym_n^k. To see that they are linearly independent, note that, for any monomial x^α, there is only one partition λ that is a permutation of α, and x^α is a term in $m_\lambda(x_1, \ldots, x_n)$. It follows that,

for partitions $\lambda \neq \mu$, the polynomials $m_\lambda(x_1, \ldots, x_n)$ and $m_\mu(x_1, \ldots, x_n)$ have no terms in common. It follows that

$$\sum_{\lambda \in \mathrm{Par}_n(k)} a_\lambda m_\lambda(x_1, \ldots, x_n) = 0 \iff (a_\lambda = 0 \text{ for all } \lambda).$$

Hence the elements (2.4) are linearly independent.

Next we prove that the elements (2.4) span Sym_n^k. Suppose $f \in \mathrm{Sym}_n^k$. We prove that f lies in the span of (2.4) by induction on the number of nonzero terms in f. The base case $f = 0$ is clear, so we suppose $f \neq 0$. Choose a term ax^α, $a \in \mathbb{Z} \setminus \{0\}$, of f. Since f is symmetric, the terms ax^β also appear in f for all permutations β of α. Let λ be the unique permutation of α that is a partition. Then

$$f - am_\lambda \in \mathrm{Sym}_n^k$$

has fewer terms than f. By the induction hypothesis, $f - am_\lambda$ can be written as a $\mathbb{Z}$-linear combination of the elements (2.4); hence the same is true for f. □

Corollary 2.1.3. (a) *The monomial symmetric polynomials*

$$m_\lambda(x_1, \ldots, x_n), \qquad \lambda \in \mathrm{Par}_n,$$

form a $\mathbb{Z}$-basis of Sym_n.

(b) *If $n \geq k$, then monomial symmetric polynomials*

$$m_\lambda(x_1, \ldots, x_n), \qquad \lambda \in \mathrm{Par}(k),$$

form a $\mathbb{Z}$-basis of Sym_n^k. *In particular,* Sym_n^k *is a free $\mathbb{Z}$-module of rank $p(k)$, the number of partitions of k.*

Proof. Part (a) follows from Proposition 2.1.2 and (2.2). Part (b) follows from Proposition 2.1.2 and the fact that, when $n \geq k$, we have $\mathrm{Par}_n(k) = \mathrm{Par}(k)$ since $\ell(\lambda) \leq n$ for all partitions λ of k. □

Exercises.

2.1.1. For a group G, recall the definition of oppose group G^{op} from Exercise 1.2.1. A *right action* of a group G on a ring R is a group homomorphism $G^{\mathrm{op}} \to \mathrm{Aut}\, R$, where $\mathrm{Aut}\, R$ denotes the group of ring automorphisms of R. By contrast, a *left action* of G on R is a group homomorphism $G \to \mathrm{Aut}\, R$.

(a) Show that (2.1) defines a left action of $\mathfrak{S}_n$ on Pol_n. In other words, show that

$$\pi_1(\pi_2 f) = (\pi_1 \pi_2) f$$

for all $f \in \mathrm{Pol}_n$ and $\pi_1, \pi_2 \in \mathfrak{S}_n$.

(b) Show that, for $n \geq 3$, (2.1) does *not* define a right action of $\mathfrak{S}_n$ on Pol_n.

(c) How could you modify (2.1) to give a natural right action of $\mathfrak{S}_n$ on Pol_n?

(Compare this exercise to Exercise 1.2.1.)

2.1.2. Show that, for $\alpha \in \mathrm{WComp}_n$ and $\pi \in \mathfrak{S}_n$, we have

$$\pi\left(x^{\alpha}\right) = x^{\alpha^{\pi^{-1}}}. \tag{2.5}$$

The next three problems use the following notation. For a subset $X \subseteq \mathfrak{S}_n$, define

$$\mathrm{Pol}_n^X = \{f \in \mathrm{Pol}_n : \pi f = f \text{ for all } \pi \in X\}.$$

For a subset $P \subseteq \mathrm{Pol}_n$, define

$$G(P) = \{\pi \in \mathfrak{S}_n : \pi f = f \text{ for all } f \in P\}.$$

2.1.3. Prove that $G(P)$ is a subgroup of $\mathfrak{S}_n$ for all subsets $P \subseteq \mathrm{Pol}_n$.

2.1.4. Is it true that $\mathrm{Pol}_n^{G(P)} = P$ for all subrings $P \subseteq \mathrm{Pol}_n$, $n \in \mathbb{N}$? Prove or give a counterexample.

2.1.5. It is true that $G(\mathrm{Pol}_n^H) = H$ for all subgroups $H \subseteq \mathfrak{S}_n$, $n \in \mathbb{N}$? Prove or give a counterexample.

2.2 Symmetric functions

It turns out that, in many ways, the symmetric polynomials behaves better in the limit where we have an infinite number of indeterminates $x_1, x_2, \ldots$. This limit, which we must be careful to define precisely, leads to the concept of a symmetric function.

For $m \geq n \geq 0$, consider the ring homomorphism

$$\mathrm{Pol}_m \to \mathrm{Pol}_n, \quad f(x_1, \ldots, x_m) \mapsto f(x_1, \ldots, x_n, 0, \ldots, 0).$$

Restriction to Sym_m gives a ring homomorphism

$$\rho_{m,n} \colon \mathrm{Sym}_m \to \mathrm{Sym}_n. \tag{2.6}$$

Note that, for $l \geq m \geq n \geq 0$, we have

$$\rho_{m,n} \circ \rho_{l,m} = \rho_{l,n}.$$

Proposition 2.2.1. *For $m \geq n \geq 0$ and $\lambda \in \mathrm{Par}$, we have*

$$\rho_{m,n}\left(m_\lambda(x_1, \ldots, x_m)\right) = m_\lambda(x_1, \ldots, x_n).$$

In particular,

$$\rho_{m,n}\left(m_\lambda(x_1, \ldots, x_m)\right) = 0 \quad \text{if} \quad \ell(\lambda) > n.$$

Proof. Let A be the (possibly empty) set of permutations $\alpha = (\alpha_1, \ldots, \alpha_m)$ of $\lambda = (\lambda_1, \ldots, \lambda_m)$ such that $\alpha_i = 0$ for all $n < i \leq m$. Let B be the set of permutations $\alpha = (\alpha_1, \ldots, \alpha_m)$ of $\lambda = (\lambda_1, \ldots, \lambda_m)$ such that $\alpha_i \neq 0$ for some $n < i \leq m$. Then

$$m_\lambda(x_1, \ldots, x_m) = \sum_{\alpha \in A} x^\alpha + \sum_{\alpha \in B} x^\alpha,$$

and so

$$\rho_{m,n}\left(m_\lambda(x_1, \ldots, x_m)\right) = \sum_{\alpha \in A} \rho_{m,n}(x^\alpha) + \sum_{\alpha \in B} \rho_{m,n}(x^\alpha) = \sum_{\alpha \in A} x^\alpha = m_\lambda(x_1, \ldots, x_n).$$

Note that, if $\ell(\lambda) > n$, then $A = \varnothing$, and so $\rho_{m,n}\left(m_\lambda(x_1, \ldots, x_m)\right) = 0$. □

Corollary 2.2.2. *For all $m \geq n \geq 0$, the ring homomorphism $\rho_{m,n}$ is surjective.*

Proof. This follows from Propositions 2.1.2 and 2.2.1. □

Restricting $\rho_{m,n}$ to Sym_n^k, we obtain its homogeneous components (see Section 1.3)

$$\rho_{m,n}^k \colon \mathrm{Sym}_m^k \twoheadrightarrow \mathrm{Sym}_n^k, \qquad k \geq 0,\ m \geq n.$$

Lemma 2.2.3. *The map $\rho_{m,n}^k$ is bijective if $m \geq n \geq k$.*

Proof. The proof of this lemma is left as Exercise 2.2.1. □

We now take the inverse limit

$$\mathrm{Sym}^k = \varprojlim_n \mathrm{Sym}_n^k \tag{2.7}$$

of the $\mathbb{Z}$-modules Sym_n^k with respect to the homomorphisms $\rho_{m,n}^k$. By definition,

$$\mathrm{Sym}^k := \left\{ f = (f_n)_{n=0}^\infty \in \prod_{n=0}^\infty \mathrm{Sym}_n^k : \rho_{m,n}^k(f_m) = f_n \text{ for all } m \geq n \right\}. \tag{2.8}$$

For $k, n \geq 0$, consider the projection

$$\rho_n^k \colon \mathrm{Sym}^k \to \mathrm{Sym}_n^k, \quad \rho_n^k(f) = f_n.$$

Proposition 2.2.4. *The projection ρ_n^k is an isomorphism for all $n \geq k$.*

Proof. By Lemma 2.2.3, any element $f \in \mathrm{Sym}^k$ is uniquely determined by f_k since we must have

$$f_n = \rho_{k,n}^k(f_k) \text{ for } n \leq k \qquad \text{and} \qquad f_n = (\rho_{n,k}^k)^{-1}(f_k) \text{ for } n > k.$$ □

By Proposition 2.2.1, for any $\lambda \in \mathrm{Par}(k)$, the sequence of monomial symmetric polynomials $(m_\lambda(x_1, \dots, x_n))_{n=0}^\infty$ is an element of Sym^k, which we call the *monomial symmetric function corresponding to* λ and denote by m_λ. We think of it as a formal sum

$$m_\lambda = \sum_{\alpha \in \lambda \mathfrak{S}_\infty} x^\alpha \in \mathrm{Sym}^k. \tag{2.9}$$

Note that, in contrast with (2.3), the above sum is infinite as long as $\lambda \neq \varnothing$. This is because we are considering permutations of the infinite sequence $(\lambda_1, \lambda_2, \dots)$ rather than the finite sequence $(\lambda_1, \dots, \lambda_n)$. For example,

$$m_{(2,1)} = \sum_{i \neq j} x_i^2 x_j = x_1^2 x_2 + x_1 x_2^2 + x_1 x_3^2 + x_1^2 x_3 + x_2 x_3^2 + x_2^2 x_3 + \cdots.$$

However, when we apply ρ_n^k for some $n \in \mathbb{N}$, only finitely many terms survive. More precisely, for $\lambda \in \mathrm{Par}(k)$,

$$\rho_n^k(m_\lambda) = m_\lambda(x_1, \dots, x_n) \qquad \text{for all } n \geq 0.$$

Now define

$$\mathrm{Sym} := \bigoplus_{k=0}^\infty \mathrm{Sym}^k. \tag{2.10}$$

Thus, Sym is the free $\mathbb{Z}$-module generated by the monomial symmetric functions m_λ for *all* partitions λ. We have surjective $\mathbb{Z}$-module homomorphisms

$$\rho_n := \bigoplus_{k=0}^\infty \rho_n^k \colon \mathrm{Sym} \to \mathrm{Sym}_n, \qquad n \geq 0. \tag{2.11}$$

If we define, for $n, k \geq 0$,

$$\mathrm{Sym}^{\leq k} := \bigoplus_{l=0}^k \mathrm{Sym}^l \qquad \text{and} \qquad \mathrm{Sym}_n^{\leq k} := \bigoplus_{l=0}^k \mathrm{Sym}_n^l, \tag{2.12}$$

then the restriction of ρ_n to $\mathrm{Sym}^{\leq k}$ yields a homomorphism of $\mathbb{Z}$-modules

$$\rho_n^{\leq k} \colon \mathrm{Sym}^{\leq k} \to \mathrm{Sym}_n^{\leq k}$$

which is an isomorphism when $k \leq n$.

Proposition 2.2.5. (a) *The monomial symmetric functions m_λ, $\lambda \in \mathrm{Par}(k)$, form a basis of Sym^k. In particular, Sym^k is a free $\mathbb{Z}$-module of rank $p(k)$, the number of partitions of k.*

(b) *The monomial symmetric functions m_λ, $\lambda \in \mathrm{Par}$, form a basis of* Sym.

Proof. Part (a) follows from Proposition 2.2.4 and Corollary 2.1.3(a). Then part (b) follows from (2.10). □

So far we have only described Sym as a $\mathbb{Z}$-module. However, as the next result shows, it is naturally a ring.

Proposition 2.2.6. *There is a unique structure of a graded ring on* Sym *such that the* ρ_n *are ring homomorphisms.*

Proof. We first prove uniqueness. Suppose we have defined a multiplication on Sym such that the ρ_n are ring homomorphisms. Let $f, g \in \mathrm{Sym}$. Then there exists some $k \geq 0$ such that $f, g \in \mathrm{Sym}^{\leq k}$ and we must have

$$\rho_{2k}(fg) := \rho_{2k}(f)\rho_{2k}(g) \in \mathrm{Sym}^{\leq 2k}.$$

Now, since ρ_{2k} is an isomorphism when restricted to $\mathrm{Sym}^{\leq 2k}$, we must have

$$fg = \left(\rho_{2k}^{\leq 2k}\right)^{-1} (\rho_{2k}(f)\rho_{2k}(g)). \tag{2.13}$$

This completes the proof of uniqueness. Existence is then seen by taking (2.13) as the *definition* of the multiplication in Sym. We leave it as exercise Exercise 2.2.2 to verify that the structure we have defined satisfies the axioms of a ring. □

The graded ring Sym is called the *ring of symmetric functions*. The term 'function' is used here in a formal sense, to distinguish this ring from the ring of symmetric *polynomials*. We do not view elements of Sym as actual functions.

It was important that we first took the inverse limit in (2.7) of the graded pieces Sym_n^k and then took the direct sum in (2.10). This construction ensures that elements of Sym are infinite sums of monomials with bounded degree. What we are doing is taking the inverse limit of the Sym_n in the category of graded rings. (General inverse limits in a category are defined in terms of a universal property.) If we had instead taken the inverse limit of the Sym_n directly (the inverse limit in the category of rings) our ring would include elements such as the infinite product $\prod_{i=1}^{\infty}(1 + x_i)$.

To illustrate this difference further, let us discuss another approach to the definition of the ring of symmetric functions. Recall the definitions of $\mathbb{Z}[\![x_1, x_2, \ldots]\!]$ and $\mathbb{Z}[\![x_1, x_2, \ldots]\!]^k$ from (1.15) and (1.17). The group $\mathfrak{S}_\infty$ acts on each $\mathbb{Z}[\![x_1, x_2, \ldots]\!]^k$ in the usual way:

$$\pi f(x_1, x_2, \ldots) := f(x_{\pi(1)}, x_{\pi(2)}, \ldots), \qquad f(x_1, x_2, \ldots) \in \mathbb{Z}[\![x_1, x_2, \ldots]\!]^k, \quad \pi \in \mathfrak{S}_\infty.$$

Then

$$\mathrm{Sym}^k = (\mathbb{Z}[\![x_1, x_2, \ldots]\!]^k)^{\mathfrak{S}_\infty}, \qquad \mathrm{Sym} = \bigoplus_{k \in \mathbb{N}} (\mathbb{Z}[\![x_1, x_2, \ldots]\!]^k)^{\mathfrak{S}_\infty}.$$

Note that

$$\mathrm{Sym} \neq \mathbb{Z}[\![x_1, x_2, \ldots]\!]^{\mathfrak{S}_\infty}. \tag{2.14}$$

(Compare to (1.18).) For example,

$$\prod_{i=1}^{\infty}(1 + x_i) \in \mathbb{Z}[\![x_1, x_2, \ldots]\!]^{\mathfrak{S}_\infty} \qquad \text{but} \qquad \prod_{i=1}^{\infty}(1 + x_i) \notin \mathrm{Sym}.$$

Warning 2.2.7. The definition of the ring of symmetric functions given in [Egg19, Def. 1.19] seems to be incorrect. The ring defined there is $\mathbb{Z}[\![x_1, x_2, \ldots]\!]^{\mathfrak{S}_\infty}$. See Warning 2.4.8.

Remark 2.2.8. We have worked in this section over the coefficient ring $\mathbb{Z}$. We could have instead worked over any other commutative ring. In particular, we will denote by $\mathrm{Sym}_{\mathbb{Q}}$ the ring of symmetric functions over $\mathbb{Q}$, which will be important for us later on. We can also obtain $\mathrm{Sym}_{\mathbb{Q}}$ by an extension of scalars: $\mathrm{Sym}_{\mathbb{Q}} \cong \mathbb{Q} \otimes_{\mathbb{Z}} \mathrm{Sym}$. (Students not familiar with tensor products can find the basics in [Sav, Ch. 3].) △

Exercises.

2.2.1. Prove Lemma 2.2.3. Furthermore, prove that $\rho^k_{m,n}$ is *not* bijective if $m > n$ and $k > n$.

2.2.2. Show that the multiplication on Sym defined in Proposition 2.2.6 satisfies the defining axioms for a ring.

2.2.3. For $n \in \mathbb{N}$, let

$$R_n = \mathbb{Z}[x]/x^n\mathbb{Z}[x]$$

be the quotient of $\mathbb{Z}[x]$ by the ideal generated by x^n. Then, for all $m \geq n \geq 0$, we have a projection

$$\varphi_{m,n} \colon R_m \to R_n, \qquad \varphi_{m,n}(f + x^m\mathbb{Z}[x]) = f + x^n\mathbb{Z}[x], \quad f \in \mathbb{Z}[x].$$

It is clear that $\varphi_{m,n} \circ \varphi_{l,m} = \varphi_{l,n}$ for all $l \geq m \geq n \geq 0$. Identify the inverse limit

$$\varprojlim_n R_n := \left\{ f = (f_n)_{n=0}^\infty \in \prod_{n=0}^\infty R_n : \varphi_{m,n}(f_m) = f_n \text{ for all } m \geq n \right\}$$

with a ring we studied in Chapter 1.

2.2.4. Suppose $k, l \in \mathbb{N}$ with $k > l$. Write $m_{(k)}m_{(l)}$ as a linear combination of m_λ, $\lambda \in \mathrm{Par}$.

2.2.5. Write $m_{(4)}m_{(3)}m_{(2)}m_{(1)}$ as a linear combination of m_λ, $\lambda \in \mathrm{Par}$.

2.3 Tableaux

Before we continue our discussion of symmetric functions, we introduce a combinatorial tool that will prove useful in our discussions.

For a partition λ, a *tableau* (plural *tableaux*) of *shape* λ is a filling of a Young diagram of λ with positive integers. (These are also called *Young tableaux*.) For example

(2.15) $T_1 =$ $\begin{array}{|c|c|c|c|}\hline 5 & 1 & 3 & 2 \\ \hline 1 & 1 & 2 \\ \cline{1-3} 3 \\ \cline{1-1}\end{array}$, $T_2 =$ $\begin{array}{|c|c|c|c|}\hline 1 & 3 & 4 & 5 \\ \hline 1 & 2 & 8 \\ \cline{1-3} 6 \\ \cline{1-1}\end{array}$, and $T_3 =$ $\begin{array}{|c|c|c|c|}\hline 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \cline{1-3} 1 \\ \cline{1-1}\end{array}$

are tableaux of shape $(4,3,1)$. We let $\mathrm{shape}(T)$ denote the shape of a tableau T. The *weight* $\mathrm{wt}(T)$ of a tableau T (also called its *content*) is the composition $\mu = (\mu_1, \ldots, \mu_r)$ where μ_i is the number of entries of T equal to i. For example, for the tableaux in (2.15), we have

$$\mathrm{wt}(T_1) = (3,2,2,0,1), \qquad \mathrm{wt}(T_2) = (2,1,1,1,1,1,0,1), \qquad \mathrm{wt}(T_3) = (8).$$

Note that $|\mathrm{wt}(T)| = |\mathrm{shape}(T)|$.

We say that a tableau is *row strict* if its entries are strictly increasing along each row. For $\lambda \in \mathrm{Par}$, we let $\mathrm{RowStrict}(\lambda)$ denote the set of row-strict tableaux of shape λ, and we let $\mathrm{RowStrict}(\lambda, n)$ denote the set of row-strict tableaux of shape λ with entries chosen from the set $\{1, 2, \ldots, n\}$. For example, for the tableaux of (2.15) and $\lambda = (4,3,1)$, we have

$$T_1 \notin \mathrm{RowStrict}(\lambda), \qquad T_2 \in \mathrm{RowStrict}(\lambda, 8), \qquad T_2 \in \mathrm{RowStrict}(\lambda),$$
$$T_2 \notin \mathrm{RowStrict}(\lambda, 5), \qquad T_3 \notin \mathrm{RowStrict}(\lambda).$$

We say that a tableau is *row nondecreasing* if its entries are weakly increasing along each row. For $\lambda \in \mathrm{Par}$, let $\mathrm{RowWeak}(\lambda)$ denote the set of row-nondecreasing tableaux of shape λ. We let $\mathrm{RowWeak}(\lambda, n)$ denote the set of such tableaux with entries chosen from the set $\{1, 2, \ldots, n\}$. For example, for the tableaux of (2.15) and $\lambda = (4,3,1)$, we have

$$T_1 \notin \mathrm{RowWeak}(\lambda), \qquad T_2 \in \mathrm{RowWeak}(\lambda, 8), \qquad T_3 \in \mathrm{RowWeak}(\lambda).$$

We say that a tableau is *semistandard* if it is row nondecreasing and its entries are strictly increasing down each column. For example

(2.16)
$$T_4 = \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 2 & 4 & 7 \\ \hline 2 & 3 & 3 \\ \cline{1-3} 6 \\ \cline{1-1} \end{array} \qquad \text{and} \qquad T_5 = \begin{array}{|c|c|c|c|c|} \hline 1 & 3 & 4 & 5 & 7 \\ \hline 2 & 6 & 9 \\ \cline{1-3} 8 \\ \cline{1-1} \end{array}$$

are semistandard tableaux. For $\lambda \in \mathrm{Par}$, we let $\mathrm{SST}(\lambda)$ denote the set of semistandard tableau of shape λ and we let $\mathrm{SST}(\lambda, n)$ denote the set of such tableaux with entries chosen from the set $\{1, 2, \ldots, n\}$.

We say that a tableau is *standard* if it is row strict, its entries are strictly increasing down each column, and it contains each integer $1, \ldots, n$ exactly once, where n is the number of boxes in T. Thus, a standard tableau has weight $\mathrm{wt}(T) = (1^n)$. For example, for the tableaux in (2.16), T_5 is standard, but T_4 is not. For $\lambda \in \mathrm{Par}$, we let $\mathrm{ST}(\lambda)$ denote the set of standard tableaux of shape λ.

If T is a tableau, we define the monomial

$$x^T := \prod_{j \in T} x_j = x^{\mathrm{wt}(T)} \tag{2.17}$$

where the first product is over the entries in T. For example, for the tableaux of (2.15), we have

$$x^{T_1} = x_1^3 x_2^2 x_3^2 x_5, \qquad x^{T_2} = x_1^2 x_2 x_3 x_4 x_5 x_6 x_8, \qquad \text{and} \qquad x^{T_3} = x_1^8.$$

We will use this convention to give explicit descriptions of important symmetric functions as sums of monomials x^T over certain tableaux.

Exercises.

2.3.1. Find all elements of $\mathrm{SST}((2,2,1),3)$. Which are standard?

2.3.2. Find all elements of $\mathrm{SST}((2,1),3)$. Which are standard?

2.4 Elementary symmetric functions

In this section we introduce our first important generating set for the ring of symmetric functions: the elementary symmetric functions. The fact that these give an algebraically independent set of generators for Sym is the fundamental theorem of symmetric functions (Theorem 2.4.6).

For $r \in \mathbb{N}$, the r-th *elementary symmetric function* e_r is the sum of all products of r distinct indeterminates x_i. Thus,

$$e_0 = 1, \qquad e_r := \sum_{1 \le i_1 < i_2 < \cdots < i_r} x_{i_1} x_{i_2} \cdots x_{i_r} = m_{(1^r)}, \quad r \ge 1. \tag{2.18}$$

We adopt the convention that

$$e_r = 0 \qquad \text{for} \qquad r < 0.$$

For $r, n \in \mathbb{N}$, the r-th *elementary symmetric polynomial in n indeterminates* is

$$e_r(x_1, \ldots, x_n) = \rho_n(e_r) = \sum_{1 \le i_1 < i_2 < \cdots < i_r \le n} x_{i_1} x_{i_2} \cdots x_{i_r} = m_{(1^r)}(x_1, \ldots, x_n). \tag{2.19}$$

It follows immediately from the definition that

$$e_r(x_1, \ldots, x_n) = 0 \qquad \text{for } r > n. \tag{2.20}$$

Convention 2.4.1. From now on, whenever we define a symmetric function $f \in \mathrm{Sym}$, we adopt the convention that

$$f(x_1, \ldots, x_n) = \rho_n(f).$$

Lemma 2.4.2. *The generating function for the elementary symmetric polynomials in n indeterminates is*

$$\sum_{r=0}^{n} e_r(x_1, \ldots, x_n) t^r = \prod_{i=1}^{n} (1 + x_i t).$$

Proof. We prove the result by induction on n. When $n = 0$, the result is clear when we recall the convention that the product $\prod_{i=1}^{n}(1 + x_i t) = 1$ when $n = 0$. Now suppose the result holds for some $n \in \mathbb{N}$. Then we have

$$\prod_{i=1}^{n+1} (1 + x_i t) = \left(\sum_{r=0}^{n} e_r(x_1, \ldots, x_n) t^r \right) (1 + x_{n+1} t)$$

$$= 1 + \sum_{r=1}^{n+1} \left(e_r(x_1, \dots, x_n) + e_{r-1}(x_1, \dots, x_n)x_{n+1}\right) t^r$$
$$= \sum_{r=0}^{n+1} e_r(x_1, \dots, x_{n+1})t^r,$$

where we have use the fact that, for $r \geq 1$,

$$e_{r-1}(x_1, \dots, x_n)x_{n+1} + e_r(x_1, \dots, x_n)$$
$$= \sum_{1 \leq i_1 < i_2 < \cdots < i_r = n+1} x_{i_1}x_{i_2}\cdots x_{i_r} + \sum_{1 \leq i_1 < i_2 < \cdots < i_r \leq n} x_{i_1}x_{i_2}\cdots x_{i_r}$$
$$= e_r(x_1, \dots, x_{n+1}).$$ □

It follows from Lemma 2.4.2 that the generating function for the elementary symmetric functions is

$$E(t) := \sum_{r=0}^{\infty} e_r t^r = \prod_{i=1}^{\infty}(1 + x_i t). \tag{2.21}$$

Remark 2.4.3. Let us take a moment to discuss how to interpret the infinite product occurring on the right-hand side of (2.21). We want to interpret it as an element of $\mathrm{Sym}[\![t]\!]$. So it should be a formal power series with coefficients in Sym. Elements of Sym are sequences $f = (f_n)_{n=0}^{\infty}$ such that $\rho_{m,n}(f_m) = f_n$ for all $m \geq n \geq 0$. For each $r \in \mathbb{N}$, the coefficient of t^r in $\prod_{i=1}^{\infty}(1 + x_i t)$ is the symmetric function $f = (f_n)_{n=0}^{\infty}$ where f_n is the coefficient of t^r in the *finite* product $\prod_{i=1}^{n}(1 + x_i t)$. By Lemma 2.4.2, we have $f_n = e_r(x_1, \dots, x_n)$, and so $f = e_r$. This justifies the equality (2.21). △

For a partition $\lambda = (\lambda_1, \lambda_2, \dots)$, define

$$e_\lambda := e_{\lambda_1} e_{\lambda_2} \cdots . \tag{2.22}$$

Note that this is a finite product since λ has finitely many nonzero terms. Our next goal is to show that the e_λ, $\lambda \in \mathrm{Par}$, form a basis of Sym. To do this, we work out how to express the e_λ in terms of the monomial symmetric functions, which we know form a basis of Sym by Proposition 2.2.5. Recall the definition of the conjugate partition (1.2) and the dominance ordering (1.3).

Proposition 2.4.4. *Let $\lambda \in \mathrm{Par}(k)$ and let λ' be the conjugate partition. Then*

$$e_\lambda = m_{\lambda'} + \sum_{\mu < \lambda'} M_{\lambda,\mu}(e, m) m_\mu. \tag{2.23}$$

where, for $\mu \in \mathrm{Par}(k)$, $\mu < \lambda'$, $M_{\lambda,\mu}(e, m)$ is the number of row-strict Young tableaux of shape λ and weight μ.

Proof. Let $\lambda = (\lambda_1, \dots, \lambda_k)$. We have

$$e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_k}, \qquad e_{\lambda_j} = \sum_{i_{j,1} < i_{j,2} < \cdots < i_{j,\lambda_j}} x_{i_{j,1}} x_{i_{j,2}} \cdots x_{i_{j,\lambda_j}}, \quad 1 \leq j \leq k.$$

Fix a particular choice of the indices

$$i_{j,1} < i_{j,2} < \cdots < i_{j,\lambda_j}, \qquad 1 \le j \le k,$$

and consider the corresponding term

$$x^\alpha = \prod_{j=1}^{k} x_{i_{j,1}} x_{i_{j,2}} \cdots x_{i_{j,\lambda_j}} \tag{2.24}$$

in the product $e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_k}$. For each $j \in \{1, \ldots, k\}$, write the indices $i_{j,1} < \cdots < i_{j,\lambda_j}$ in the j-th row of the Young diagram of λ:

(2.25)

$i_{1,1}$	$i_{1,2}$	$\cdots$	$\cdots$	$\cdots$	i_{1,λ_1}
$i_{2,1}$	$i_{2,2}$	$\cdots$	$\cdots$	$\cdots$	i_{2,λ_2}
$\vdots$	$\vdots$				
$i_{j,1}$	$\cdots$	i_{j,λ_j}			
$\vdots$					
$\vdots$	$i_{\lambda'_2,2}$				
$\vdots$					
$i_{\lambda'_1,1}$					

Since the entries are strictly increasing along each row, all entries $\le r$ must occur in the leftmost r columns of the diagram. Hence the total number of entries $\le r$ is less than or equal to the total number of boxes in the first r columns. Since α_r is the number of entries equal to r, this implies that

$$\alpha_1 + \cdots + \alpha_r \le \lambda'_1 + \cdots + \lambda'_r.$$

Thus we have $\alpha \le \lambda'$.

Now, since $\alpha \le \lambda'$ for all terms (2.24) and e_λ is $\mathfrak{S}_\infty$ invariant, it follows from Lemma 1.2.4 that only m_μ for $\mu \le \lambda'$ appear when we write e_λ in the basis of monomial symmetric functions. Thus

$$e_\lambda = \sum_{\mu \le \lambda'} M_{\lambda,\mu}(e, m) m_\mu,$$

where $M_{\lambda,\mu}(e, m)$ is the number of row-strict Young tableaux of shape λ and weight μ. If $\mu = \lambda'$, then the only such filling is the one in which each entry in (2.25) is equal to its column number, that is, $i_{j,l} = l$ for all j, l. Thus $M_{\lambda,\lambda'}(e, m) = 1$. □

The proof of Proposition 2.4.4 gives us an explicit combinatorial description of the elementary symmetric functions in terms of Young tableaux.

Proposition 2.4.5. *For any partition λ and $n \geq 1$, we have*

$$e_\lambda = \sum_{T \in \mathrm{RowStrict}(\lambda)} x^T \qquad \text{and} \qquad e_\lambda(x_1, \ldots, x_n) = \sum_{T \in \mathrm{RowStrict}(\lambda, n)} x^T. \tag{2.26}$$

Proof. The first equation follows from the proof of Proposition 2.4.4. Since $e_\lambda(x_1, \ldots, x_n) = \rho_n(e_\lambda)$, the second equation then follows from the fact that $\rho_n(x^T) = 0$ if T has any entries greater than n. □

Recall that elements of a commutative ring R are *algebraically independent* over a subring $S \subseteq R$ if the elements do not satisfy any polynomial equation with coefficients in S.

Theorem 2.4.6 (Fundamental theorem of symmetric functions)**.** *The elementary symmetric functions e_λ, $\lambda \in \mathrm{Par}$, form a basis of* Sym*. Furthermore,*

$$\mathrm{Sym} = \mathbb{Z}[e_1, e_2, \ldots]$$

and the e_r, $r > 0$, are algebraically independent over $\mathbb{Z}$.

Proof. For the first statement of the theorem, it suffices to prove that, for $k \in \mathbb{N}$, the e_λ, $\lambda \in \mathrm{Par}(k)$, form a basis of Sym^k. Fix a total order on the basis m_λ, $\lambda \in \mathrm{Par}(k)$, of Sym^k that refines the dominance ordering on Sym^k. (For instance, we can take the lexicographic ordering; see Exercise 1.2.4.) Then Proposition 2.4.4 implies that matrix expressing the $e_{\lambda'}$, $\lambda \in \mathrm{Par}(k)$, in terms of this ordered basis is triangular with ones on the diagonal, and hence invertible. It follows every monomial symmetric function can be written as a $\mathbb{Z}$-linear combination of elementary symmetric functions. Thus the elementary symmetric functions e_λ, $\lambda \in \mathrm{Par}$, form a basis of Sym. This implies that every element of Sym is uniquely expressible as a polynomial in the e_r, proving the second statement of the theorem. □

Corollary 2.4.7 (Fundamental theorem of symmetric polynomials)**.** *For $n \in \mathbb{N}$, the $e_\lambda(x_1, \ldots, x_n)$, $\lambda \in \mathrm{Par}_n$, form a basis of Sym_n. Furthermore,*

$$\mathrm{Sym}_n = \mathbb{Z}[e_1(x_1, \ldots, x_n), \ldots, e_n(x_1, \ldots, x_n)]$$

and the elementary symmetric polynomials $e_1(x_1, \ldots, x_n), \ldots, e_n(x_1, \ldots, x_n)$ are algebraically independent over $\mathbb{Z}$.

Proof. It follows from Theorem 2.4.6 and (2.20) that Sym_n is generated by the elementary symmetric polynomials $e_1(x_1, \ldots, x_n), \ldots, e_n(x_1, \ldots, x_n)$. By Proposition 2.1.2, the rank of Sym_n^k as a $\mathbb{Z}$-module is the cardinality of $\mathrm{Par}_n(k)$. Thus, the $e_\lambda(x_1, \ldots, x_n)$, $\lambda' \in \mathrm{Par}_n(k)$, must be linearly independent. Hence $e_1(x_1, \ldots, x_n), \ldots, e_n(x_1, \ldots, x_n)$ are algebraically independent over $\mathbb{Z}$. □

Theorem 2.4.6 implies that the ring Sym of symmetric functions is isomorphic, as a ring, to a polynomial ring in infinitely many indeterminates. You might be tempted to think that this implies Sym is not so interesting! However, it is the *relationship between various elements* of Sym (for example, different bases) that makes the theory of symmetric functions so rich. For example, in the next section will introduce another set of generators for Sym, and then study their relation to the elementary symmetric functions.

Warning 2.4.8. It is crucial for Theorem 2.4.6 that we defined symmetric functions as we did, instead of defining them to be elements of $\mathbb{Z}[\![x_1, x_2, \ldots]\!]^{\mathfrak{S}_\infty}$. The elementary symmetric functions $e_1, e_2, \ldots$ do *not* generate $\mathbb{Z}[\![x_1, x_2, \ldots]\!]^{\mathfrak{S}_\infty}$ as a ring; see Exercise 2.4.1. Thus, the statement at the top of [Egg19, p. 38] is incorrect (since the definition of symmetric functions in that reference is incorrect); see Warning 2.2.7.

Exercises.

2.4.1. Prove that the elementary symmetric functions $e_1, e_2, \ldots$ do *not* generate $\mathbb{Z}[\![x_1, x_2, \ldots]\!]^{\mathfrak{S}_n}$ as a ring. *Hint*: Find an element of $\mathbb{Z}[\![x_1, x_2, \ldots]\!]^{\mathfrak{S}_n}$ that cannot be written as a polynomial in the $e_1, e_2, \ldots$.

2.4.2. Write $e_{(2,2)}$ and $e_{(2,1,1)}$ as linear combinations of monomial symmetric functions.

2.4.3. Prove that, for all $0 \leq k \leq n$, we have

$$e_k(\underbrace{1, \ldots, 1}_{n \text{ entries}}) = \binom{n}{k}.$$

Here the notation $e_k(1, \ldots, 1)$ means evaluating the elementary symmetric polynomial $e_k(x_1, \ldots, x_n)$ at $(x_1, \ldots, x_n) = (1, \ldots, 1)$.

2.4.4. Find and prove a formula for $M_{\lambda,(n)}(e, m)$, where $\lambda \vdash n$.

2.4.5. Find and prove a formula for $M_{\lambda,(1^n)}(e, m)$, where $\lambda \vdash n$.

2.4.6. Find and prove a formula for $M_{\lambda,(n-1,1)}(e, m)$, where $\lambda \vdash n$.

2.5 Complete homogeneous symmetric functions

For $r \in \mathbb{N}$, the *r-th complete homogeneous symmetric function* h_r is the sum of all monomials of total degree r in the indeterminates $x_1, x_2, \ldots$, so that

$$h_0 = 1, \qquad h_r := \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_r} x_{i_1} x_{i_2} \cdots x_{i_r} = \sum_{\lambda \in \operatorname{Par}(r)} m_\lambda, \quad r \geq 1. \tag{2.27}$$

Note that

$$h_1 = e_1.$$

We adopt the convention that

$$h_r = 0 \qquad \text{for} \qquad r < 0.$$

As usual, for $n \in \mathbb{N}$, we then define the *r-th complete homogeneous symmetric polynomial in n indeterminates* to be

$$h_r(x_1, \dots, x_n) := \rho_n(h_r) = \sum_{1 \le i_1 \le i_2 \le \cdots \le i_r \le n} x_{i_1} x_{i_2} \cdots x_{i_r}.$$

Proposition 2.5.1. *The generating function for the complete homogeneous symmetric functions is*

$$H(t) := \sum_{r \ge 0} h_r t^r = \prod_{i=1}^{\infty} \frac{1}{1 - x_i t}. \tag{2.28}$$

Proof. The proof of this result is left as Exercise 2.5.1. □

It follows from (2.21) and (2.28) that

$$H(t)E(-t) = 1. \tag{2.29}$$

Comparing coefficients of powers of t, this implies that

$$\sum_{r=0}^{n} (-1)^r e_r h_{n-r} = 0 \qquad \text{for all} \quad n \ge 1. \tag{2.30}$$

(We see here the power of the approach of generating functions!)

By Theorem 2.4.6, we may define a homomorphism of graded rings

$$\omega \colon \mathrm{Sym} \to \mathrm{Sym}, \qquad \omega(e_r) = h_r, \quad r \in \mathbb{N}. \tag{2.31}$$

Recall that an endomorphism is an *involution* if its square is the identity map (equivalently, it is an automorphism and equal to its own inverse.)

Proposition 2.5.2. *We have*

$$\omega(h_r) = e_r, \qquad \textit{for all} \quad r \in \mathbb{N}. \tag{2.32}$$

In particular, the map ω is an involution.

Proof. Once we prove the first statement, we have $\omega^2(e_n) = e_n$ for all $n \in \mathbb{N}$, and hence the second follows statement follows from Theorem 2.4.6.

We prove that $\omega(h_n) = e_n$ for all $n \in \mathbb{N}$ by induction on n. Since ω is a ring homomorphism, we have

$$\omega(h_0) = \omega(1) = 1 = e_0,$$

and so the result holds for $n = 0$. Now suppose that $n \ge 1$ and $\omega(h_r) = e_r$ for $r < n$. By (2.30), we have

$$h_n = -\sum_{r=1}^{n} (-1)^r e_r h_{n-r} \qquad \text{and} \qquad e_n = (-1)^{n+1} \sum_{r=0}^{n-1} (-1)^r e_r h_{n-r}.$$

Thus

$$\omega(h_n) = -\sum_{r=1}^{n}(-1)^r\omega(e_r)\omega(h_{n-r}) = -\sum_{r=1}^{n}(-1)^r h_r e_{n-r}$$
$$= -\sum_{r=0}^{n-1}(-1)^{n-r}h_{n-r}e_r = (-1)^{n+1}\sum_{r=0}^{n-1}(-1)^r e_r h_{n-r} = e_n,$$

where, in the third equality, we replaced r by $n-r$. □

Theorem 2.5.3. *We have*

$$\mathrm{Sym} = \mathbb{Z}[h_1, h_2, \dots]$$

and the complete homogeneous symmetric functions h_r, $r > 0$, are algebraically independent over $\mathbb{Z}$.

Proof. By Proposition 2.5.2, ω is an automorphism. Thus the result follows from Theorem 2.4.6. □

Remark 2.5.4. If we work with symmetric polynomials in n indeterminates, we still have a mapping

$$\omega\colon \mathrm{Sym}_n \to \mathrm{Sym}_n, \qquad \omega(e_r) = h_r, \quad 1 \le r \le n.$$

Applying ρ_n to (2.30) shows that

$$\sum_{r=0}^{n}(-1)^r e_r(x_1, \dots, x_n)h_{n-r}(x_1, \dots, x_n) = 0.$$

Thus, one can repeat the argument in the proofs of Proposition 2.5.2 and Theorem 2.5.3 to see that

$$\mathrm{Sym}_n = \mathbb{Z}[h_1(x_1, \dots, x_n), \dots, h_n(x_1, \dots, x_n)],$$

with $h_1, \dots, h_n$ algebraically independent. However, even though $e_r(x_1, \dots, x_n) = 0$ for $r > n$, the complete homogeneous symmetric polynomials $h_{n+1}(x_1, \dots, x_n), h_{n+2}(x_1, \dots, x_n), \dots$ are *nonzero*. For example,

$$h_3(x_1, x_2) = x_1^3 + x_1^2x_2 + x_1x_2^2 + x_3^2 = 2h_1(x_1, x_2)h_2(x_1, x_2) - h_1(x_1, x_2)^3.$$

In particular, we do *not* have $\omega(e_r) = h_r$ for $r > n$. This is example of how symmetric functions are often simpler than symmetric polynomials. △

For $\lambda = (\lambda_1, \lambda_2, \dots) \in \mathrm{Par}$, we define

$$h_\lambda = h_{\lambda_1}h_{\lambda_2}\cdots. \tag{2.33}$$

By Theorem 2.5.3, the h_λ, $\lambda \in \mathrm{Par}$, form a $\mathbb{Z}$-basis of Sym. Define the *forgotten symmetric functions*

$$f_\lambda = \omega(m_\lambda), \qquad \lambda \in \mathrm{Par}. \tag{2.34}$$

The name "forgotten" comes from the fact that these symmetric functions have no simple direct description and thus tend not to appear much in applications. We now have four bases of Sym:

$$\{e_\lambda : \lambda \in \mathrm{Par}\}, \quad \{h_\lambda : \lambda \in \mathrm{Par}\}, \quad \{m_\lambda : \lambda \in \mathrm{Par}\}, \quad \{f_\lambda : \lambda \in \mathrm{Par}\}.$$

The automorphism ω interchanges the first two bases and the last two bases.

It follows from Proposition 2.2.5 that we can write each h_λ in terms of the monomial symmetric functions. The next result gives an explicit formula.

Proposition 2.5.5. *For $\lambda \in \mathrm{Par}(k)$, we have*

$$h_\lambda = \sum_{\mu \vdash |\lambda|} M_{\lambda,\mu}(h, m) m_\mu, \tag{2.35}$$

where, for $\mu \vdash |\lambda|$, $M_{\lambda,\mu}(h, m)$ is the number of row-nondecreasing Young tableaux of shape λ and weight μ.

Proof. The proof of this proposition is left as Exercise 2.5.2. □

We can now give an explicit combinatorial description of the complete homogeneous symmetric functions in terms of Young tableaux. (Compare to Proposition 2.4.5.)

Proposition 2.5.6. *For any partition λ and $n \geq 1$, we have*

$$h_\lambda = \sum_{T \in \mathrm{RowWeak}(\lambda)} x^T \qquad \text{and} \qquad h_\lambda(x_1, \ldots, x_n) = \sum_{T \in \mathrm{RowWeak}(\lambda, n)} x^T. \tag{2.36}$$

Proof. The proof of this proposition is analogous to that of Proposition 2.4.5. In particular, to prove the first equality, one alters the proof of Proposition 2.4.4 to work with the h_λ instead of the e_λ. The first equality also follows from Proposition 2.5.5. Then applying ρ_n gives the second equality. □

Exercises.

2.5.1. Prove (2.28) by following the method of the proof of Lemma 2.4.2 and (2.21).

2.5.2. Prove Proposition 2.5.5.

2.5.3. Write $h_{(2,2)}$ and $h_{(2,1,1)}$ as linear combinations of monomial symmetric functions.

2.5.4. Prove that, for all $0 \leq k \leq n$, we have

$$h_k(\underbrace{1, \ldots, 1}_{n \text{ entries}}) = \binom{n+k-1}{k}.$$

Here the notation $h_k(1, \ldots, 1)$ means evaluating the elementary symmetric polynomial $h_k(x_1, \ldots, x_n)$ at $(x_1, \ldots, x_n) = (1, \ldots, 1)$.

2.5.5. Find and prove a formula for $M_{\lambda,(n)}(h,m)$, where $\lambda \vdash n$.

2.5.6. Find and prove a formula for $M_{\lambda,(1^n)}(h,m)$, where $\lambda \vdash n$.

2.5.7. Find and prove a formula for $M_{\lambda,(n-1,1)}(h,m)$, where $\lambda \vdash n$.

2.5.8. Prove that, for all $n \geq 1$,

$$e_n = \det(h_{1-j+k})_{1\leq j,k\leq n},$$

where $(h_{1-j+k})_{1\leq j,k\leq n}$ is the matrix with (j,k) entry equal to h_{1-j+k}.

2.5.9. Prove that, for all $n \geq 1$,

$$h_n = \det(e_{1-j+k})_{1\leq j,k\leq n},$$

where $(e_{1-j+k})_{1\leq j,k\leq n}$ is the matrix with (j,k) entry equal to e_{1-j+k}.

2.6 Power sums

In this section we introduce another important set of symmetric functions. As we will see, these new symmetric functions are most well-behaved when we work over the rational numbers instead of the integers. We will make use in this section of the results of Exercises 1.4.5 to 1.4.8 on formal power series.

For $r \geq 1$, the *r-th power sum* is

$$p_r := \sum_{i=1}^{\infty} x_i^r = m_{(r)}. \tag{2.37}$$

As usual, for $n \in \mathbb{N}$, we define

$$p_r(x_1,\ldots,x_n) := \rho_n(p_r) = x_1^r + x_2^r + \cdots + x_n^r.$$

It turns out that many formulas are more natural when we consider the *scaled power sums* $\frac{p_r}{r}$.

Proposition 2.6.1. *The generating function for the sequence of scaled powers sums in n indeterminates is*

$$\sum_{r=1}^{\infty} \frac{p_r(x_1,\ldots,x_n)}{r} t^r = \log\left(\prod_{j=1}^{n} \frac{1}{1-x_j t}\right).$$

Proof. Recall that $\log\left(\frac{1}{1-x}\right) = -\log(1-x) = \sum_{r=1}^{\infty} \frac{1}{r} x^r$. Thus we have

$$\sum_{r=1}^{\infty} \frac{p_r(x_1,\ldots,x_n)}{r} t^r = \sum_{r=1}^{\infty} \sum_{j=1}^{n} \frac{1}{r} (x_j t)^r = \sum_{j=1}^{n} \log\left(\frac{1}{1-x_j t}\right) = \log\left(\prod_{j=1}^{n} \frac{1}{1-x_j t}\right). \qquad \square$$

Now, taking the limit as $n \to \infty$, we obtain the generating function for the power sums p_n:

$$P(t) := \sum_{r=1}^{\infty} \frac{p_r}{r} t^r = \log\left(\prod_{j=1}^{\infty} \frac{1}{1-x_j t}\right) \overset{(2.28)}{=} \log H(t). \tag{2.38}$$

(See Remark 2.4.3 for a discussion on how we interpret this type of infinite product of generating functions.)

Remark 2.6.2. Note that, in [Mac15, p. 23], $P(t)$ denotes the generating function $\sum_{r=1}^{\infty} p_r t^{r-1}$, which is the formal derivative of our generating function (2.38). △

Note that

$$P'(t) = \frac{\mathrm{d}}{\mathrm{d}t} \log H(t) = H'(t)/H(t). \tag{2.39}$$

Also, by (2.29), $E(t) = H(-t)^{-1}$, and so

$$E'(t) = H(-t)^{-2} H'(-t)$$

Thus

$$E'(t)/E(t) = H'(-t)/H(-t) = P'(-t). \tag{2.40}$$

Proposition 2.6.3. *For $n \geq 1$, we have*

$$nh_n = \sum_{r=1}^{n} p_r h_{n-r}, \tag{2.41}$$

$$ne_n = \sum_{r=1}^{n} (-1)^{r-1} p_r e_{n-r}. \tag{2.42}$$

(The equations (2.41) and (2.42) are known as Newton's identities*.)*

Proof. From (2.39), we have $H'(t) = P'(t)H(t)$. Thus

$$\sum_{n=1}^{\infty} n h_n t^{n-1} = \left(\sum_{n=1}^{\infty} p_n t^{n-1}\right)\left(\sum_{m=0}^{\infty} h_m t^m\right) = \sum_{n=1}^{\infty}\sum_{r=1}^{n} p_r h_{n-r} t^{n-1}.$$

Equating coefficients yields (2.41). The proof of (2.42) is analogous. □

For a partition $\lambda = (\lambda_1, \lambda_2, \dots) \in \mathrm{Par}$, we define

$$p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots. \tag{2.43}$$

Recall, from Remark 2.2.8, that $\mathrm{Sym}_{\mathbb{Q}}$ denotes the symmetric functions with coefficients in $\mathbb{Q}$.

Theorem 2.6.4. *We have*

$$\mathrm{Sym}_{\mathbb{Q}} = \mathbb{Q}[p_1, p_2, \dots]$$

and the power sums $p_1, p_2, \dots$ are algebraically independent over $\mathbb{Q}$. Equivalently, the p_λ, $\lambda \in \mathrm{Par}$, form a basis for $\mathrm{Sym}_{\mathbb{Q}}$.

Proof. The identity (2.41) allows us to solve for the power sums recursively in terms of the complete homogeneous symmetric functions and vice versa:

$$p_n = nh_n - \sum_{r=1}^{n-1} p_r h_{n-r} \qquad \text{and} \qquad h_n = \frac{1}{n}\sum_{r=1}^{n} p_r h_{n-r}. \tag{2.44}$$

It follows that

$$\mathrm{Sym}_{\mathbb{Q}} = \mathbb{Q}[h_1, h_2, \dots] = \mathbb{Q}[p_1, p_2, \dots],$$

where the first equality is Theorem 2.5.3. Furthermore, if the p_r satisfied a polynomial equation over $\mathbb{Q}$ then we can replace them by their expressions in terms of the h_r to see that the h_r satisfy a nontrivial polynomial equation over $\mathbb{Q}$. (See Exercise 2.6.1.) Then, clearing denominators, we would have that the h_r are algebraically dependent over $\mathbb{Z}$. But this contradicts Theorem 2.5.3. Hence the p_r are algebraically independent over $\mathbb{Q}$. □

Remark 2.6.5. It is crucial in Theorem 2.6.4 that we work over $\mathbb{Q}$ instead of $\mathbb{Z}$. For example,

$$h_2 = \frac{1}{2}\left(p_1^2 + p_2\right)$$

does not have integral coefficients when expressed in terms of the p_λ. For this reason, when we really want to work over the integers, the powers sums can be problematic. △

Proposition 2.6.6. *For $n \in \mathbb{N}$ and $\lambda \in \mathrm{Par}$, we have*

$$\omega(p_n) = (-1)^{n-1} p_n \qquad \text{and} \qquad \omega(p_\lambda) = (-1)^{|\lambda|-\ell(\lambda)} p_\lambda, \tag{2.45}$$

where ω is the automorphism of Sym *defined in (2.31).*

Proof. We prove the first equality by induction on n. Since

$$\omega(p_1) = \omega(e_1) = h_1 = p_1,$$

the result holds for $n = 1$. Then, for $n \geq 2$, we have, using (2.44),

$$\omega(p_n) \overset{(2.44)}{=} n\omega(h_n) - \sum_{r=1}^{n-1} \omega(p_r)\omega(h_{n-r}) = ne_n - \sum_{r=1}^{n-1} (-1)^{r-1} p_r e_{n-r} \overset{(2.42)}{=} (-1)^{n-1} p_n,$$

completing the proof of the induction step. Then, for $\lambda \in \mathrm{Par}$, we have

$$\omega(p_\lambda) = \omega(p_{\lambda_1}) \cdots \omega(p_{\lambda_{\ell(\lambda)}}) = (-1)^{|\lambda|-\ell(\lambda)} p_\lambda.$$ □

It follows from Proposition 2.2.5 that we can write each p_λ in terms of the monomial symmetric functions. The next result gives an explicit formula.

Proposition 2.6.7. *For $k \geq 1$ and $\lambda \in \mathrm{Par}(k)$, we have*

$$p_\lambda = \sum_{\mu \vdash k} M_{\lambda,\mu}(p, m) m_\mu, \tag{2.46}$$

where $M_{\lambda,\mu}(p,m)$ is the number of Young tableaux of shape λ and weight μ in which the entries in each row are constant.

Proof. The proof of this result is left as Exercise 2.6.2. □

We conclude this section by deducing explicit formulas for the h_n and e_n as linear combinations of the p_λ. Recall that, for $\lambda \in \mathrm{Par}$, $m_i(\lambda)$ denotes the multiplicity of i in λ. Define

$$z_\lambda := \prod_{i \geq 1} i^{m_i(\lambda)} m_i(\lambda)! \tag{2.47}$$

(See Exercise 1.1.2 for a combinatorial interpretation of z_λ.)

Proposition 2.6.8. *We have*

$$H(t) = \sum_{\lambda \in \mathrm{Par}} \frac{p_\lambda}{z_\lambda} t^{|\lambda|} \qquad \text{and} \qquad E(t) = \sum_{\lambda \in \mathrm{Par}} (-1)^{|\lambda| - \ell(\lambda)} \frac{p_\lambda}{z_\lambda} t^{|\lambda|}. \tag{2.48}$$

Proof. By (2.38), we have

$$H(t) = \exp P(t) = \exp \sum_{r \geq 1} \frac{p_r}{r} t^r = \prod_{r=1}^{\infty} \exp\left(\frac{p_r}{r} t^r\right) = \prod_{r=1}^{\infty} \sum_{m_r=0}^{\infty} \frac{(p_r t^r)^{m_r}}{m_r! r^{m_r}} = \sum_{\lambda \in \mathrm{Par}} \frac{p_\lambda}{z_\lambda} t^{|\lambda|}.$$

We now obtain the second equation in (2.48) by applying ω and using (2.45). □

Corollary 2.6.9. *For $n \in \mathbb{N}$, we have*

$$h_n = \sum_{\lambda \in \mathrm{Par}(n)} \frac{p_\lambda}{z_\lambda} \qquad \text{and} \qquad e_n = (-1)^n \sum_{\lambda \in \mathrm{Par}(n)} (-1)^{\ell(\lambda)} \frac{p_\lambda}{z_\lambda}. \tag{2.49}$$

Proof. This follows from equating coefficients in (2.48). □

Exercises.

2.6.1. Prove that if f is a nonzero polynomial in the p_r, then, after using (2.44) to replace all p_r by their expressions in terms of the h_r, we obtain a nonzero polynomial. *Hint*: Use the fact that p_r is equal to rh_r plus a polynomial in $h_1, \ldots, h_{r-1}$.

2.6.2. Prove Proposition 2.6.7.

2.6.3. Write $p_{(2,2,1)}$ and $p_{(2,2,2)}$ as linear combinations of monomial symmetric functions.

2.6.4. Find and prove a formula for $M_{\lambda,(n)}(p,m)$, where $\lambda \vdash n$.

2.6.5. Find and prove a formula for $M_{\lambda,(1^n)}(p,m)$, where $\lambda \vdash n$.

2.6.6. Find and prove a formula for $M_{\lambda,(n-1,1)}(p,m)$, where $\lambda \vdash n$.

2.6.7. Using Newton's identity (2.42), prove that

$$p_n = \begin{vmatrix} e_1 & 1 & 0 & 0 & \cdots & 0 \\ 2e_2 & e_1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 & 0 \\ \vdots & \vdots & & & e_1 & 1 \\ ne_n & e_{n-1} & e_{n-2} & \cdots & \cdots & e_1 \end{vmatrix} \quad \text{and} \quad n!e_n = \begin{vmatrix} p_1 & 1 & 0 & 0 & \cdots & 0 \\ p_2 & p_1 & 2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ p_{n-1} & p_{n-2} & \cdots & \cdots & p_1 & n-1 \\ p_n & p_{n-1} & \cdots & \cdots & p_2 & p_1 \end{vmatrix}.$$

Chapter 3

Schur functions

In this chapter we introduce Schur functions, which form what is probably the most important basis of the ring of symmetric functions. In most applications of symmetric functions to other areas of mathematics, such as geometry and representation theory, the Schur functions play a crucial role. We begin with a discussion of alternating polynomials, which we then use these to define Schur functions. Next we define a natural bilinear form on the space of symmetric functions and examine the behaviour of the Schur functions with respect to this bilinear form. We then give a description of Schur functions in terms of semistandard tableaux similar to the tableaux descriptions of the other bases of the ring of symmetric functions we studied in Chapter 2. We also give a complete description of the transition matrices between all the bases we have seen so far. Finally, we prove the important Littlewood–Richardson rule, which gives a method of computing the product of Schur functions.

3.1 Alternating polynomials

An *inversion* for $\pi \in \mathfrak{S}_n$ is a pair (i, j) with $1 \leq i < j \leq n$ such that $\pi(i) > \pi(j)$. Let $\mathrm{inv}(\pi)$ denote the number of inversions of π. The *sign* of π is defined to be

$$\mathrm{sgn}(\pi) := (-1)^{\mathrm{inv}(\pi)}.$$

Alternatively, if we write π as a product of m transpositions (recall that a transposition is a cycle of length two) then $\mathrm{sgn}(\pi) = (-1)^m$. (Note that m is not unique, but its parity is.) We have

$$\mathrm{sgn}(\pi\sigma) = \mathrm{sgn}(\pi)\,\mathrm{sgn}(\sigma) \qquad \text{for all} \quad \pi, \sigma \in \mathfrak{S}_n.$$

Example 3.1.1. If $\pi = (1\,3\,5) \in \mathfrak{S}_5$, then the set of inversions is

$$\{(1,2), (1,5), (2,5), (3,4), (3,5), (4,5)\}.$$

Thus

$$\mathrm{inv}(\pi) = 6 \qquad \text{and} \qquad \mathrm{sgn}(\pi) = (-1)^6 = 1.$$

We can write

$$\pi = (1\,3)(3\,5) = (2\,4)(1\,3)(3\,5)(2\,4),$$

and both expressions involve an even number of permutations. △

We say a polynomial $f \in \mathrm{Pol}_n$ is *alternating*, or *skew-symmetric*, if

$$\pi f = \mathrm{sgn}(\pi) f \qquad \text{for all} \quad \pi \in \mathfrak{S}_n.$$

Let

$$\mathrm{Alt}_n := \{f \in \mathrm{Pol}_n : \pi f = \mathrm{sgn}(\pi) f \text{ for all } \pi \in \mathfrak{S}_n\}$$

denote the $\mathbb{Z}$-module of all alternating polynomials in n indeterminates. Note that Alt_n is *not* a subring of Pol_n. In fact, if $f, g \in \mathrm{Alt}_n$, then

$$\pi(fg) = (\pi f)(\pi g) = \mathrm{sgn}(\pi) f \,\mathrm{sgn}(\pi) g = fg \qquad \text{for all} \quad \pi \in \mathfrak{S}_n,$$

and so $fg \in \mathrm{Sym}_n$. On the other hand, if $f \in \mathrm{Sym}_n$ and $g \in \mathrm{Alt}_n$, then

$$\pi(fg) = (\pi f)(\pi g) = \mathrm{sgn}(\pi) fg \qquad \text{for all} \quad \pi \in \mathfrak{S}_n,$$

and so $fg \in \mathrm{Alt}_n$. It follows that Alt_n is a Sym_n-module.

We can construct alternating polynomials by antisymmetrizing monomials as follows. If $\alpha \in \mathrm{WComp}_n$, define

$$a_\alpha = a_\alpha(x_1, \ldots, x_n) := \sum_{\pi \in \mathfrak{S}_n} \mathrm{sgn}(\pi) x^{\alpha^\pi}. \tag{3.1}$$

Then a_α is alternating (Exercise 3.1.4).

Example 3.1.2. If $n = 3$ and $\alpha = (5, 3, 2)$, then

$$a_\alpha = \sum_{\pi \in \mathfrak{S}_3} \mathrm{sgn}(\pi) x^{\alpha^\pi} = x_1^5 x_2^3 x_3^2 - x_1^5 x_2^2 x_3^3 - x_1^3 x_2^5 x_3^2 - x_1^2 x_2^3 x_3^5 + x_1^2 x_2^5 x_3^3 + x_1^3 x_2^2 x_3^5. \qquad \triangle$$

Lemma 3.1.3. *For $\alpha \in \mathrm{WComp}_n$, we have*

$$a_\alpha = \det(x_i^{\alpha_j})_{1 \le i,j \le n} = \det \begin{pmatrix} x_1^{\alpha_1} & x_1^{\alpha_2} & \cdots & x_1^{\alpha_n} \\ x_2^{\alpha_1} & x_2^{\alpha_2} & \cdots & x_2^{\alpha_n} \\ \vdots & \vdots & \vdots & \vdots \\ x_n^{\alpha_1} & x_n^{\alpha_2} & \cdots & x_n^{\alpha_n} \end{pmatrix}.$$

Proof. Recall that the determinant of a matrix $B = (b_{i,j})_{1 \le i,j \le n}$ can be written as

$$\det B = \sum_{\pi \in \mathfrak{S}_n} \left(\mathrm{sgn}(\pi) \prod_{i=1}^n b_{i,\pi(i)} \right). \tag{3.2}$$

Thus

$$\det(x_i^{\alpha_j})_{1 \le i,j \le n} = \sum_{\pi \in \mathfrak{S}_n} \left(\mathrm{sgn}(\pi) \prod_{i=1}^n x_i^{\alpha_{\pi(i)}} \right) = \sum_{\pi \in \mathfrak{S}_n} \mathrm{sgn}(\pi) x^{\alpha^\pi} = a_\alpha. \qquad \square$$

Remark 3.1.4. Note that if two parts of α are equal, say $\alpha_i = \alpha_j$ for some $1 \le i < j \le n$, then, for $\pi = (i\,j)$, we have $\operatorname{sgn}(\pi) = -1$, and so

$$a_\alpha = -a_{\alpha^\pi} = -a_\alpha.$$

Hence $a_\alpha = 0$. (This can also be seen using Lemma 3.1.3 since, if $\alpha_i = \alpha_j$, then two columns of the matrix there are equal; hence the determinant is zero.) It follows that there are no repeated terms in the sum (3.1) unless $a_\alpha = 0$. (Compare to Remark 2.1.1.) It also follows that, for any $\alpha \in \mathrm{WComp}_n$ with $a_\alpha \neq 0$, there exists a *unique* $\pi \in \mathfrak{S}_n$ such that $\alpha^\pi \in \mathrm{Par}_n$. Then $a_\alpha = \operatorname{sgn}(\pi) a_{\alpha^\pi}$, and so we can restrict our attention to a_λ where

$$\lambda_1 > \lambda_2 > \cdots > \lambda_n. \qquad \triangle$$

Let

$$\delta^n = (n-1, n-2, \ldots, 1, 0) \in \mathrm{Par}_n. \tag{3.3}$$

For $\alpha, \beta \in \mathrm{WComp}_n$, define

$$\alpha + \beta = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n) \in \mathrm{WComp}_n.$$

Lemma 3.1.5. *We have a bijection of sets*

$$\mathrm{Par}_n \to \{\lambda \in \mathrm{Par}_n : \lambda_1 > \lambda_2 > \cdots > \lambda_n\}, \qquad \lambda \mapsto \lambda + \delta^n.$$

Proof. The proof of this lemma is left as Exercise 3.1.5. □

Proposition 3.1.6. *The elements*

$$a_{\lambda+\delta^n}(x_1, \ldots, x_n), \qquad \lambda \in \mathrm{Par}_n, \tag{3.4}$$

form a $\mathbb{Z}$*-basis of* Alt_n.

Proof. We must show that the elements (3.4) are linearly independent and span Alt_n. Our proof will be similar to the proof of Proposition 2.1.2. First, it follows from Remark 3.1.4 that, for $\lambda, \mu \in \mathrm{Par}_n$ with $\lambda \neq \mu$, the polynomials $a_{\lambda+\delta^n}$ and $a_{\mu+\delta^n}$ have no terms in common. Hence these elements are linearly independent.

Next we prove that the elements (3.4) span Alt_n. Suppose $f \in \mathrm{Alt}_n$. We prove that f lies in the span of (3.4) by induction on the number of nonzero terms in f. The base case $f = 0$ is clear, so we suppose $f \neq 0$. Choose a nonzero term cx^α, $c \in \mathbb{Z} \setminus \{0\}$, of f. Since f is alternating, the terms $\operatorname{sgn}(\pi) c x^{\alpha^\pi}$ also appear in f for all $\pi \in \mathfrak{S}_n$. By Remark 3.1.4 and Lemma 3.1.5, we have

$$\alpha^\pi = \lambda + \delta^n \qquad \text{for some unique} \quad \lambda \in \mathrm{Par}_n,\ \pi \in \mathfrak{S}_n.$$

Then $c a_{\lambda+\delta^n}(x_1, \ldots, x_n)$ contains the term $\operatorname{sgn}(\pi) c x^\alpha$, and so

$$f - \operatorname{sgn}(\pi) c a_{\lambda+\delta^n}(x_1, \ldots, x_n) \in \mathrm{Alt}_n$$

has fewer terms than f. By the induction hypothesis, $f - \operatorname{sgn}(\pi) c a_{\lambda+\delta^n}(x_1, \ldots, x_n)$ can be written as a $\mathbb{Z}$-linear combination of the elements (3.4); hence the same is true for f. □

Lemma 3.1.7. *Every element of* Alt_n *is divisible by* $\prod_{1\le i<j\le n}(x_i - x_j)$.

Proof. Suppose $f \in \mathrm{Alt}_n$. For $1 \le i < j \le n$, interchanging x_i and x_j multiplies f by -1. It follows that substituting x_i for x_j in f gives zero:

$$f(x_1, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_n) = 0.$$

Thus, by the factor theorem for polynomials, f is divisible by $x_i - x_j$. Now, since $x_i - x_j$ is degree one, it is an irreducible polynomial. Therefore, the $x_i - x_j$, $1 \le i < j \le n$, are all coprime. Hence, since Pol_n is a unique factorization domain, f is divisible by their product $\prod_{1\le i<j\le n}(x_i - x_j)$. □

Corollary 3.1.8. *We have*

$$a_{\delta^n}(x_1, \dots, x_n) = \det \begin{pmatrix} x_1^{n-1} & x_1^{n-2} & \cdots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \cdots & x_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \\ x_n^{n-1} & x_n^{n-2} & \cdots & x_n & 1 \end{pmatrix} = \prod_{1\le i<j\le n} (x_i - x_j). \tag{3.5}$$

(The polynomial (3.5) is called a Vandermonde determinant*.)*

Proof. The first equality is Lemma 3.1.3. By Lemma 3.1.7, we have

$$a_{\delta^n}(x_1, \dots, x_n) = c \prod_{1\le i<j\le n} (x_i - x_j)$$

for some $c \in \mathrm{Pol}_n$. Since both sides are homogeneous of degree $n(n-1)/2$, we see that $c \in \mathbb{Z}$. Now, in the expression (3.2) for the determinant of the matrix in (3.5), the term where π is the identity element of $\mathfrak{S}_n$ gives that the product $x_1^{n-1}x_2^{n-2}\cdots x_n$ of the diagonal entries appears with coefficient 1. On the other hand, when expanding $\prod_{1\le i<j\le n}(x_i - x_j)$, this term arises from choosing the first term x_i in each factor, and hence also appears with coefficient 1. Hence $c = 1$, completing the proof. □

As noted above, Alt_n is a Sym_n-module. The ring Sym_n is also naturally a Sym_n-module. (Any ring is a left module over itself, with action given by left multiplication.)

Proposition 3.1.9. *The map*

$$\mathrm{Sym}_n \to \mathrm{Alt}_n, \qquad f \mapsto f a_{\delta^n}(x_1, \dots, x_n), \tag{3.6}$$

is an isomorphism of Sym_n*-modules. In particular,* Alt_n *is a free* Sym_n *module of rank* 1 *generated by* $a_{\delta^n}(x_1, \dots, x_n)$.

Proof. We already noted earlier in this section that the product of a symmetric polynomial and an alternating polynomial is an alternating polynomial. Thus the map (3.6) is indeed a map from Sym_n to Alt_n. It is straightforward to verify that it is a homomorphism of Sym_n-modules. Then by Lemma 3.1.7 and (3.5), we have an inverse map

$$\mathrm{Alt}_n \to \mathrm{Sym}_n, \qquad f \mapsto f / a_{\delta^n}(x_1, \dots, x_n).$$ □

Exercises.

3.1.1. Prove that, for $\alpha \in \mathrm{WComp}_n$,

$$a_\alpha = \sum_{\pi \in \mathfrak{S}_n} \mathrm{sgn}(\pi)\pi\left(x^\alpha\right).$$

3.1.2. Prove that $f \in \mathrm{Pol}_n$ is alternating if and only if $\pi f = -f$ for all transpositions $\pi \in \mathfrak{S}_n$.

3.1.3. Prove that $\mathrm{Alt}_n \cap \mathrm{Sym}_n = \{0\}$ for $n \geq 2$.

3.1.4. Show that, for all $\alpha \in \mathrm{WComp}_n$, the polynomial a_α defined in (3.1) is alternating.

3.1.5. Prove Lemma 3.1.5.

3.1.6. For $k, n \in \mathbb{N}$, let Alt_n^k denote the $\mathbb{Z}$-module of alternating polynomials in n indeterminates that are homogeneous of degree k (and the zero polynomial). Find $\operatorname{rank} \mathrm{Alt}_n^k$.

3.2 Schur polynomials and Schur functions

We are now ready to define Schur polynomials and Schur functions. We do this by using Proposition 3.1.9 to pull back a basis of the space of alternating polynomials to space of symmetric polynomials. We then prove the important Jacobi–Trudi identities, which give the Schur functions in terms of the elementary and complete homogeneous symmetric functions.

For $\lambda \in \mathrm{Par}_n$, it follows from Lemma 3.1.7 that $a_{\lambda+\delta^n}(x_1, \ldots, x_n)$ is divisible by $a_{\delta^n}(x_1, \ldots, x_n)$. Furthermore, for all $\pi \in \mathfrak{S}_n$, we have

$$\pi\left(\frac{a_{\lambda+\delta^n}(x_1,\ldots,x_n)}{a_{\delta^n}(x_1,\ldots,x_n)}\right) = \frac{\pi a_{\lambda+\delta^n}(x_1,\ldots,x_n)}{\pi a_{\delta^n}(x_1,\ldots,x_n)} = \frac{\mathrm{sgn}(\pi) a_{\lambda+\delta^n}(x_1,\ldots,x_n)}{\mathrm{sgn}(\pi) a_{\delta^n}(x_1,\ldots,x_n)} = \frac{a_{\lambda+\delta^n}(x_1,\ldots,x_n)}{a_{\delta^n}(x_1,\ldots,x_n)},$$

and so $\frac{a_{\lambda+\delta^n}(x_1,\ldots,x_n)}{a_{\delta^n}(x_1,\ldots,x_n)} \in \mathrm{Sym}_n$. The degree of this polynomial is $|\lambda + \delta^n| - |\delta^n| = |\lambda|$.

Definition 3.2.1 (Schur polynomial)**.** For $\lambda \in \mathrm{Par}_n$, define the *Schur polynomial in n indeterminates*

$$s_\lambda(x_1, \ldots, x_n) = \frac{a_{\lambda+\delta^n}(x_1,\ldots,x_n)}{a_{\delta^n}(x_1,\ldots,x_n)} \in \mathrm{Sym}_n^{|\lambda|}. \tag{3.7}$$

Lemma 3.2.2. *If $\lambda \in \mathrm{Par}_n$, then*

$$a_{\lambda+\delta^{n+1}}(x_1, \ldots, x_n, 0) = \begin{cases} a_{\lambda+\delta^n}(x_1,\ldots,x_n) e_n(x_1,\ldots,x_n) & \text{if } \ell(\lambda) \leq n, \\ 0 & \text{if } \ell(\lambda) > n. \end{cases}$$

Proof. The proof of this result is left as Exercise 3.2.1. □

Remark 3.2.3. It is stated in [Mac15, p. 41] that $a_\alpha(x_1,\dots,x_n,0) = a_\alpha(x_1,\dots,x_n)$ for $\alpha \in \mathrm{WComp}_n$. However, this is false. See Exercise 3.2.2. △

It follows from Lemma 3.2.2 that, for $\ell(\lambda) \le n$,

$$\rho_{n+1,n}\left(s_\lambda(x_1,\dots,x_{n+1})\right) = \frac{a_{\lambda+\delta^n}(x_1,\dots,x_n)e_n(x_1,\dots,x_n)}{a_{\delta^n}(x_1,\dots,x_n)e_n(x_1,\dots,x_n)} = s_\lambda(x_1,\dots,x_n)$$

in the notation of (2.6). Therefore, if we define

$$s_\lambda(x_1,\dots,x_n) = 0 \qquad \text{for} \quad \ell(\lambda) > n, \tag{3.8}$$

we have

$$\rho_{n+1,n}\left(s_\lambda(x_1,\dots,x_{n+1})\right) = s_\lambda(x_1,\dots,x_n) \qquad \text{for all} \quad \lambda \in \mathrm{Par}.$$

Thus, the sequence $\big(s_\lambda(x_1,\dots,x_n)\big)_{n=0}^\infty$ defines an element of $\mathrm{Sym}^{|\lambda|}$; see (2.8). For $\lambda \in \mathrm{Par}$, the element

$$s_\lambda := \big(s_\lambda(x_1,\dots,x_n)\big)_{n=0}^\infty \tag{3.9}$$

is called the *Schur function* corresponding to the partition λ.

Theorem 3.2.4. (a) *The Schur polynomials $s_\lambda(x_1,\dots,x_n)$, $\lambda \in \mathrm{Par}_n(k)$, form a $\mathbb{Z}$-basis of Sym_n^k.*

(b) *The Schur polynomials $s_\lambda(x_1,\dots,x_n)$, $\lambda \in \mathrm{Par}_n$, form a $\mathbb{Z}$-basis of Sym_n.*

(c) *The Schur functions s_λ, $\lambda \in \mathrm{Par}(k)$, form a $\mathbb{Z}$-basis of Sym^k.*

(d) *The Schur functions s_λ, $\lambda \in \mathrm{Par}$, form a $\mathbb{Z}$-basis of Sym.*

Proof. By definition, the alternating function $a_{\lambda+\delta^n}(x_1,\dots,x_n)$ is the image under the map (3.6) of the Schur polynomial $s_\lambda(x_1,\dots,x_n)$. Then part (b) follows from Propositions 3.1.6 and 3.1.9. Since s_λ is homogeneous of degree $|\lambda|$, part (a) follows immediately. Finally, (c) and (d) follow from Proposition 2.2.4 since $\mathrm{Par}_n^k = \mathrm{Par}_n$ when $n \ge k$. □

It follows from Theorems 2.4.6 and 2.5.3 that the Schur functions can be expressed as polynomials in the elementary and complete homogeneous symmetric functions. The following theorem gives these expressions explicitly.

Theorem 3.2.5 (Jacobi–Trudi identities)**.** *For $\lambda \in \mathrm{Par}$, we have*

$$s_\lambda = \det(h_{\lambda_i-i+j})_{1\le i,j\le n} = \det\begin{pmatrix} h_{\lambda_1} & h_{\lambda_1+1} & \cdots & h_{\lambda_1+n-1} \\ h_{\lambda_2-1} & h_{\lambda_2} & \cdots & h_{\lambda_2+n-2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{\lambda_n-n+1} & h_{\lambda_n-n+2} & \cdots & h_{\lambda_n} \end{pmatrix} \quad \text{for } n \ge \ell(\lambda) \tag{3.10}$$

and

$$s_\lambda = \det(e_{\lambda'_i-i+j})_{1\le i,j\le m} = \det\begin{pmatrix} e_{\lambda'_1} & e_{\lambda'_1+1} & \cdots & e_{\lambda'_1+m-1} \\ e_{\lambda'_2-1} & e_{\lambda'_2} & \cdots & e_{\lambda'_2+m-2} \\ \vdots & \vdots & \ddots & \vdots \\ e_{\lambda'_m-m+1} & e_{\lambda'_m-m+2} & \cdots & e_{\lambda'_m} \end{pmatrix} \quad \text{for } m \ge \ell(\lambda'). \tag{3.11}$$

(Recall that $e_r = h_r = 0$ for $r < 0$.)

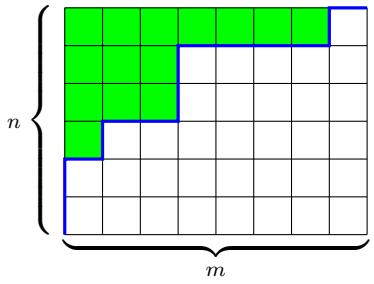
We split the proof of Theorem 3.2.5 into pieces. We begin with a technical lemma. Then we prove that the determinants appearing in (3.10) and (3.11) are equal. Finally, we prove (3.10).

Lemma 3.2.6. *Suppose λ is partition, and let $n \geq \ell(\lambda)$, $m \geq \ell(\lambda')$. Then the $n+m$ numbers*

$$\lambda_i + n - i,\ 1 \leq i \leq n, \qquad n - 1 + j - \lambda'_j,\ 1 \leq j \leq m,$$

are a permutation of $\{0, 1, 2, \ldots, n+m-1\}$.

Proof. The Young diagram of λ (coloured green in the diagram below) is contained in the Young diagram of the partition (m^n), which is a $n \times m$ rectangle.



Number the segments of the boundary (coloured blue in the diagram above) between λ and its complement in (m^n) with the numbers $0, 1, \ldots, n+m-1$, starting at the bottom. The vertical segments are then numbered $\lambda_i + n - i$, $1 \leq i \leq n$, and (considering the transposed diagram) the horizontal segments are numbered

$$(n+m-1) - (\lambda'_j + m - j) = n - 1 + j - \lambda'_j, \qquad 1 \leq j \leq m. \qquad \square$$

Next we prove that the determinants appearing in (3.10) and (3.11) are equal. In fact, for future use, we prove a slightly more general result.

Lemma 3.2.7. *For $\lambda, \mu \in \mathrm{Par}$ with $\ell(\lambda), \ell(\mu) \leq n$ and $\ell(\lambda'), \ell(\mu') \leq m$, we have*

$$\det(h_{\lambda_i - \mu_j - i + j})_{1 \leq i,j \leq n} = \det(e_{\lambda'_i - \mu'_j - i + j})_{1 \leq i,j \leq m}. \tag{3.12}$$

Proof. Let N be a positive integer and consider the $(N+1) \times (N+1)$ matrices

$$H = (h_{i-j})_{0 \leq i,j \leq N}, \qquad E = \left((-1)^{i-j} e_{i-j}\right)_{0 \leq i,j \leq N}.$$

Both H and E are lower triangular with 1's on the diagonal. Thus $\det(H) = \det(E) = 1$. Furthermore, (2.30) implies that H and E are mutually inverse.

Let λ be a partition of length $\ell(\lambda) \leq n$ such that $\ell(\lambda') \leq m$, with $n + m = N + 1$. Let H_λ be the $n \times n$ submatrix of H formed by entries whose row index is $\lambda_i + n - i$, $1 \leq i \leq n$, and column index is $\mu_j + n - j$, $1 \leq j \leq n$. The, by Lemma 3.2.6, the $m \times m$ submatrix of E formed by *deleting* these rows and column is the submatrix of E formed by entries whose row index is $n - 1 + i - \lambda'_i$, $1 \leq i \leq m$, and column index is $n - 1 + j - \mu'_j$, $1 \leq j \leq m$. Then, by a generalization of Cramer's rule, we have

$$\det\left(h_{\lambda_i-\mu_j-i+j}\right)_{1\le i,j\le n} = \det(H_\lambda) \tag{3.13}$$
$$= (-1)^{|\lambda|+|\mu|}\det\left((-1)^{\lambda'_i-\mu'_j-i+j}e_{\lambda'_i-\mu'_j-i+j}\right)_{1\le i,j\le m}.$$

(See [GAE02] for a proof of this generalization of Cramer's rule. The matrices A, X, and Y in [GAE02] are our H, E and the identity matrix, respectively. Then $\det(A) = 1$ in [GAE02, (2)], and we expand the determinant of A_Y appearing there along the columns coming from the identity matrix, then take the transpose.) Using the fact that multiplying a row or column of a matrix by -1 multiplies the determinant by -1, we see that the minus signs on the right-hand side of (3.13) cancel, giving (3.12). □

Note that taking $\mu = \varnothing$ in (3.12) shows that the determinants appearing in (3.10) and (3.11) are equal. We can now give the proof of the Jacobi–Trudi identities.

Proof of Theorem 3.2.5. In light of (3.12), it suffices to prove (3.10). Assume $n \ge \ell(\lambda)$. It suffices to work with n indeterminates $x_1, \dots, x_n$. (See Exercise 3.2.3.) So we will write h_r for $h_r(x_1, \dots, x_n)$ and e_r for $e_r(x_1, \dots, x_n)$.

For $1 \le k \le n$, let

$$e_r^{(k)} = e_r(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n)$$

and define the matrix

$$M = \left((-1)^{n-i}e_{n-i}^{(j)}\right)_{1\le i,j\le n}.$$

We first prove that, for $\alpha \in \mathrm{WComp}_n$, we have

$$A_\alpha = H_\alpha M \tag{3.14}$$

where

$$A_\alpha = (x_j^{\alpha_i})_{1\le i,j\le n}, \qquad H_\alpha = (h_{\alpha_i-n+j})_{1\le i,j\le n}.$$

Let

$$E^{(k)}(t) = \sum_{r=0}^{n-1} e_r^{(k)}t^r = \prod_{\substack{1\le i\le n\\ i\ne k}} (1 + x_i t).$$

Then, using (2.28), we have

$$H(t)E^{(k)}(-t) = \frac{1}{1 - x_k t}.$$

Comparing coefficients of t^{α_i} gives

$$\sum_{j=1}^{n} h_{\alpha_i-n+j} \cdot (-1)^{n-j}e_{n-j}^{(k)} = x_k^{\alpha_i},$$

which implies (3.14).

Now, taking determinants of both sides of (3.14) and using Lemma 3.1.3 gives

$$a_\alpha = \det(A_\alpha) = \det(H_\alpha)\det(M) \qquad \text{for all} \quad \alpha \in \mathrm{WComp}_n.$$

Note that H_{δ^n} is triangular with 1's on the diagonal, and hence $\det(H_{\delta^n}) = 1$. Therefore $\det(M) = a_{\delta^n}$, and so

$$a_\alpha = a_{\delta^n} \det(H_\alpha). \tag{3.15}$$

Taking $\alpha = \lambda + \delta^n$ and noting that $H_{\lambda+\delta^n} = (h_{\lambda_i - i+j})_{1\le i,j\le n}$ then gives (3.10). □

Corollary 3.2.8. *For all $n \in \mathbb{N}$, we have*

$$s_{(n)} = h_n \qquad \text{and} \qquad s_{(1^n)} = e_n. \tag{3.16}$$

Proof. The first equality follows from taking $\lambda = (n)$ in (3.10) and the second equality follows from taking $\lambda = (1^n)$ in (3.11). □

Corollary 3.2.9. *We have*

$$\omega(s_\lambda) = s_{\lambda'} \qquad \text{for all} \quad \lambda \in \mathrm{Par}, \tag{3.17}$$

where ω is the involution of Sym *defined in (2.31).*

Proof. Since the determinant of a matrix is a polynomial in its entries, we have

$$\omega(s_\lambda) = \det\left(\omega(h_{\lambda_i - i+j})\right)_{1\le i,j\le n} = \det\left(e_{\lambda_i - i+j}\right)_{1\le i,j\le n} = s_{\lambda'},$$

where we have used the fact that $\lambda'' = \lambda$. □

Exercises.

3.2.1. Prove Lemma 3.2.2.

3.2.2. Show that, for all $n \ge 1$, there exists $\alpha \in \mathrm{WComp}_n$ such that

$$a_\alpha(x_1, \dots, x_n, 0) \ne a_\alpha(x_1, \dots, x_n).$$

Hint: Find an example where the left-hand side is zero, but the right-hand side is not.

3.2.3. Note that the Jacobi–Trudi identity (3.10) holds for *any* $n \ge \ell(\lambda)$. If $m \ge \ell(\lambda)$, explain the relationship between the matrix appearing in (3.10) when $n = m$ and $n = m+1$. Why do these values of n give the same formula for s_λ? Now do the same analysis for the identity (3.11). Note that this analysis justifies working in n variables in the proof of Theorem 3.2.5 since, when $n \ge |\lambda|$, we can apply Proposition 2.2.4.

3.2.4. Write the Schur function $s_{(2,1)}$ as a linear combination of the complete homogeneous symmetric functions, the elementary symmetric functions, and the monomial symmetric functions.

3.2.5. Write the Schur function $s_{(3,1)}$ as a linear combination of the complete homogeneous symmetric functions, the elementary symmetric functions, and the monomial symmetric functions.

3.3 The Hall inner product

In this section we introduce a natural bilinear form on the space of symmetric functions and explore how the various bases of Sym that we have found behave with respect to this bilinear form. We will follow the presentation of [Mac15, §I.4] but include more detail than is found there.

Definition 3.3.1 (Hall inner product)**.** We define a $\mathbb{Z}$-valued bilinear form $\langle\ ,\ \rangle$ on Sym, called the *Hall inner product*, by requiring that the bases $(h_\lambda)_{\lambda\in\mathrm{Par}}$ and $(m_\lambda)_{\lambda\in\mathrm{Par}}$ are dual to each other:

$$\langle h_\lambda, m_\mu\rangle = \delta_{\lambda\mu} := \begin{cases} 1 & \text{if } \lambda = \mu, \\ 0 & \text{if } \lambda \neq \mu. \end{cases} \tag{3.18}$$

The inner product of two arbitrary elements can then by computed by writing them in these bases and using $\mathbb{Z}$-bilinearity. We also use the notation $\langle\ ,\ \rangle$ to denote the corresponding $\mathbb{Q}$-valued bilinear form on $\mathrm{Sym}_{\mathbb{Q}}$.

Remark 3.3.2. We will see later a motivation for the definition of the Hall inner product coming from the representation theory of the symmetric group; see Remark 6.3.5. △

Let $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ be two sequences of independent indeterminates. We will use the notation $f(x)$ and $f(y)$ to denote symmetric functions in the x_i and y_i, respectively.

For the proofs that follow, it will be useful to introduce the following notation. For $\alpha \in \mathbb{Z}^n$, define

$$h_\alpha(x) := h_{\alpha_1}(x)h_{\alpha_2}(x)\cdots. \tag{3.19}$$

Since $h_r(x) = 0$ for $r < 0$, we see that $h_\alpha(x) = 0$ unless $\alpha \in \mathrm{WComp}$. Furthermore, for $\alpha \in \mathrm{WComp}$, $h_\alpha(x)$ is equal to $h_\lambda(x)$, where λ is the unique partition that is a permutation of α.

Proposition 3.3.3. *We have*

$$\prod_{i,j=1}^{\infty}(1 - x_iy_j)^{-1} = \sum_{\lambda\in\mathrm{Par}} h_\lambda(x)m_\lambda(y) = \sum_{\lambda\in\mathrm{Par}} m_\lambda(x)h_\lambda(y). \tag{3.20}$$

Proof. Taking $t = y_j$ in (2.28), and then taking the product over $j \geq 1$ gives

$$\prod_{i,j=1}^{\infty}(1 - x_iy_j)^{-1} = \prod_{j=1}^{\infty} H(y_j) = \prod_{j=1}^{\infty}\left(\sum_{r=0}^{\infty} h_r(x)y_j^r\right) = \sum_{\alpha\in\mathrm{WComp}} h_\alpha(x)y^\alpha. \tag{3.21}$$

Now, by (1.6), we have

$$\mathrm{WComp} = \bigsqcup_{\lambda\in\mathrm{Par}} \lambda\mathfrak{S}_\infty. \tag{3.22}$$

Thus

$$\sum_{\alpha\in\mathrm{WComp}} h_\alpha(x)y^\alpha = \sum_{\lambda\in\mathrm{Par}}\sum_{\alpha\in\lambda\mathfrak{S}_\infty} h_\alpha(x)y^\alpha = \sum_{\lambda\in\mathrm{Par}} h_\lambda(x)\sum_{\alpha\in\lambda\mathfrak{S}_\infty} y^\alpha = \sum_{\lambda\in\mathrm{Par}} h_\lambda(x)m_\lambda(y).$$

This completes the proof of the first equality in (3.20). The second equality follows by interchanging the x_i and y_i. □

Note that the expressions for $\prod_{i,j=1}^\infty (1-x_iy_j)^{-1}$ in (3.20) involve the dual bases $(h_\lambda)_{\lambda\in\mathrm{Par}}$ and $(m_\lambda)_{\lambda\in\mathrm{Par}}$. The next result makes this observation more precise.

Proposition 3.3.4. *For each $k \geq 0$, let $(u_\lambda)_{\lambda\in\mathrm{Par}(k)}$ and $(v_\lambda)_{\lambda\in\mathrm{Par}(k)}$ be $\mathbb{Q}$-bases of $\mathrm{Sym}^k_{\mathbb{Q}}$. Then the following conditions are equivalent:*

(a) $\langle u_\lambda, v_\mu\rangle = \delta_{\lambda\mu}$ *for all* λ, μ;

(b) $\sum_{\lambda\in\mathrm{Par}} u_\lambda(x)v_\lambda(y) = \prod_{i,j=1}^\infty (1-x_iy_j)^{-1}$.

Proof. Since the $(h_\lambda)_{\lambda\in\mathrm{Par}(k)}$ and $(m_\lambda)_{\lambda\in\mathrm{Par}(k)}$ form $\mathbb{Q}$-bases of Sym^k, we have, for $\lambda \in \mathrm{Par}(k)$,

$$u_\lambda = \sum_{\rho\in\mathrm{Par}(k)} c_{\lambda\rho}h_\rho \qquad \text{and} \qquad v_\mu = \sum_{\sigma\in\mathrm{Par}(k)} d_{\mu\sigma}m_\sigma \qquad \text{for some} \quad c_{\lambda\rho}, d_{\mu\sigma} \in \mathbb{Q}. \tag{3.23}$$

Then

$$\langle u_\lambda, v_\mu\rangle = \sum_{\rho\in\mathrm{Par}(k)} c_{\lambda\rho}d_{\mu\rho}.$$

Thus

$$\text{(a)} \iff \sum_{\rho\in\mathrm{Par}(k)} c_{\lambda\rho}d_{\mu\rho} = \delta_{\lambda\mu} \quad \text{for all } \lambda, \mu \in \mathrm{Par}(k),\ k \in \mathbb{N}. \tag{3.24}$$

On the other hand, by (3.20), condition (b) is equivalent to

$$\sum_{\lambda\in\mathrm{Par}(k)} u_\lambda(x)v_\lambda(y) = \sum_{\rho\in\mathrm{Par}(k)} h_\rho(x)m_\rho(y).$$

Using (3.23), we have

$$\sum_{\lambda\in\mathrm{Par}(k)} u_\lambda(x)v_\lambda(y) = \sum_{\lambda,\rho,\sigma\in\mathrm{Par}(k)} c_{\lambda\rho}d_{\lambda\sigma}h_\rho(x)m_\sigma(y).$$

Thus

$$\begin{aligned} \text{(b)} &\iff \sum_{\rho\in\mathrm{Par}(k)} h_\rho(x)m_\rho(y) = \sum_{\lambda,\rho,\sigma\in\mathrm{Par}(k)} c_{\lambda\rho}d_{\lambda\sigma}h_\rho(x)m_\sigma(y) \\ &\iff \sum_{\lambda\in\mathrm{Par}(k)} c_{\lambda\rho}d_{\lambda\sigma} = \delta_{\rho\sigma} \quad \text{for all } \rho, \sigma \in \mathrm{Par}(k),\ k \in \mathbb{N}. \end{aligned} \tag{3.25}$$

If we define the matrices

$$C_k = (c_{\lambda\mu})_{\lambda,\mu\in\mathrm{Par}(k)} \qquad \text{and} \qquad D_k = (d_{\lambda\mu})_{\lambda,\mu\in\mathrm{Par}(k)}\,.$$

Then, letting M^{T} denote the transpose of a matrix M, we have

$$\text{(a)} \overset{(3.24)}{\iff} C_k D_k^{\mathsf{T}} = I \iff D_k C_k^{\mathsf{T}} = I \iff C_k^{\mathsf{T}} D_k = I \overset{(3.25)}{\iff} \text{(b)},$$

where, in the second bi-implication, we took transposes of both sides. □

Proposition 3.3.4 tells us that, to find other dual bases, we can look for other expressions for the infinite product $\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1}$. This is our next goal.

Proposition 3.3.5. *We have*

$$\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1} = \sum_{\lambda\in\mathrm{Par}} z_\lambda^{-1} p_\lambda(x)p_\lambda(y). \tag{3.26}$$

Proof. For $r \in \mathbb{N}$, the r-th power sum in the variables x_iy_j, $i,j \geq 1$, is

$$p_r(xy) := \sum_{i,j=1}^{\infty}(x_iy_j)^r = \left(\sum_{i=1}^{\infty} x_i^r\right)\left(\sum_{i=1}^{\infty} y_j^r\right) = p_r(x)p_r(y).$$

It follows that, for $\lambda \in \mathrm{Par}$, we have

$$p_\lambda(xy) = p_{\lambda_1}(xy)p_{\lambda_2}(xy)\cdots = p_{\lambda_1}(x)p_{\lambda_1}(y)p_{\lambda_2}(x)p_{\lambda_2}(y)\cdots = p_\lambda(x)p_\lambda(y).$$

Now, by (2.28) and (2.48), we have

$$\prod_{i=1}^{\infty}(1-x_it)^{-1} = \sum_{\lambda\in\mathrm{Par}} z_\lambda^{-1} p_\lambda(x)t^{|\lambda|}. \tag{3.27}$$

Using this equality, but with the set of variables x_iy_j, $i,j \geq 1$, instead of the set of variables x_i, $i \geq 1$, gives

$$\prod_{i,j=1}^{\infty}(1-x_iy_jt)^{-1} = \sum_{\lambda\in\mathrm{Par}} z_\lambda^{-1} p_\lambda(xy)t^{|\lambda|} = \sum_{\lambda\in\mathrm{Par}} z_\lambda^{-1} p_\lambda(x)p_\lambda(y)t^{|\lambda|}.$$

Taking $t = 1$ gives (3.26). □

Proposition 3.3.6 (Cauchy's formula)**.** *We have*

$$\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1} = \sum_{\lambda\in\mathrm{Par}} s_\lambda(x)s_\lambda(y). \tag{3.28}$$

Proof. As usual, it suffices to prove the result over finite sets of variables $x = (x_1,\ldots,x_n)$ and $y = (y_1,\ldots,y_n)$. Recall the definition of the alternating symmetric functions a_α from Section 3.1 and the partition $\delta^n = (n-1, n-2, \ldots, 1, 0)$. Taking $x_k = y_k = 0$ for $k > n$ in (3.21), then multiplying both sides by $a_{\delta^n}(x)a_{\delta^n}(y)$, we have

$$a_{\delta^n}(x)a_{\delta^n}(y)\prod_{i,j=1}^{n}(1-x_iy_j)^{-1}$$

$$= a_{\delta^n}(x) \sum_{\substack{\alpha \in \mathrm{WComp}_n \\ \pi \in \mathfrak{S}_n}} h_\alpha(x) \operatorname{sgn}(\pi) y^{\alpha + (\delta^n)^\pi}$$
$$= a_{\delta^n}(x) \sum_{\substack{\beta \in \mathrm{WComp}_n \\ \pi \in \mathfrak{S}_n}} \operatorname{sgn}(\pi) h_{\beta - (\delta^n)^\pi}(x) y^\beta,$$

where, in the second equality, we changed to a sum over $\beta = \alpha + (\delta^n)^\pi$. Now, by (3.15), we have

$$a_\beta(x) = a_{\delta^n}(x) \det\left(h_{\beta_i - n + j}(x)\right)_{1 \le i,j \le n} = a_{\delta^n}(x) \det\left(h_{\beta_i - \delta^n_j}(x)\right)_{1 \le i,j \le n}$$
$$\overset{(3.2)}{=} a_{\delta^n}(x) \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn}(\pi) \prod_{i=1}^n h_{\beta_i - \delta^n_{\pi(i)}}(x) = a_{\delta^n}(x) \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn}(\pi) h_{\beta - (\delta^n)^\pi}(x).$$

Combining this with the above, we have

$$a_{\delta^n}(x) a_{\delta^n}(y) \prod_{i,j=1}^n (1 - x_i y_j)^{-1} = \sum_{\beta \in \mathrm{WComp}_n} a_\beta(x) y^\beta \overset{(3.22)}{=} \sum_{\mu \in \mathrm{Par}} \sum_{\pi \in \mathfrak{S}_n} a_{\mu^\pi}(x) y^{\mu^\pi}$$
$$= \sum_{\mu \in \mathrm{Par}} \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn}(\pi) a_\mu(x) y^{\mu^\pi} = \sum_{\mu \in \mathrm{Par}} a_\mu(x) a_\mu(y) = \sum_{\lambda \in \mathrm{Par}} a_{\lambda + \delta^n}(x) a_{\lambda + \delta^n}(y),$$

where, in the last equality, we used the fact that $a_\mu = 0$ unless $\mu = \lambda + \delta^n$ for some $\lambda \in \mathrm{Par}$. Finally, we divide both sides by $a_{\delta^n}(x) a_{\delta^n}(y)$ and use the definition (3.7) of the Schur polynomials. □

We can now draw several conclusions about the bilinear form.

Theorem 3.3.7. (a) *We have*

$$\langle p_\lambda, p_\mu \rangle = \delta_{\lambda\mu} z_\lambda \qquad \text{for all} \quad \lambda, \mu \in \mathrm{Sym}. \tag{3.29}$$

In particular, the p_λ, $\lambda \in \mathrm{Par}$, *form an orthogonal basis of* $\mathrm{Sym}_{\mathbb{Q}}$.

(b) *We have*

$$\langle s_\lambda, s_\mu \rangle = \delta_{\lambda\mu}, \qquad \text{for all} \quad \lambda, \mu \in \mathrm{Sym}. \tag{3.30}$$

In particular, the Schur functions form an orthonormal basis of Sym.

(c) *The bilinear form* $\langle\ ,\ \rangle$ *is symmetric and positive definite.*

(d) *The involution* ω *of (2.31) is an isometry. In other words,*

$$\langle \omega(u), \omega(v) \rangle = \langle u, v \rangle \qquad \text{for all} \quad u, v \in \mathrm{Sym}_{\mathbb{Q}}. \tag{3.31}$$

(e) *We have*

$$\langle e_\lambda, f_\mu \rangle = \delta_{\lambda\mu}, \qquad \text{for all} \quad \lambda, \mu \in \mathrm{Par}, \tag{3.32}$$

where $f_\mu = \omega(m_\mu)$, *as in (2.34). In other words, the bases* $(e_\lambda)_{\lambda \in \mathrm{Par}}$ *and* $(f_\lambda)_{\lambda \in \mathrm{Par}}$ *are dual.*

Proof. (a) It follows from Proposition 3.3.4 and (3.26) that

$$\langle z_\lambda^{-1} p_\lambda, p_\mu \rangle = \delta_{\lambda\mu} \qquad \text{for all} \quad \lambda, \mu \in \mathrm{Par}.$$

Then we multiply both sides by z_λ and use bilinearity of the form.

(b) This follows immediately from Proposition 3.3.4 and (3.28).

(c) The fact that the form is symmetric follows from the symmetry of (3.30). Furthermore, since $\langle\ ,\ \rangle$ admits an orthonormal basis, it is positive definite.

(d) Using (3.17), we have, for all $\lambda, \mu \in \mathrm{Par}$,

$$\langle \omega(s_\lambda), \omega(s_\mu) \rangle = \langle s_{\lambda'}, s_{\mu'} \rangle = \delta_{\lambda'\mu'} = \delta_{\lambda\mu} = \langle s_\lambda, s_\mu \rangle.$$

Since the s_λ, $\lambda \in \mathrm{Par}$, form a basis of Sym, (3.31) follows.

(e) For all $\lambda, \mu \in \mathrm{Par}$, we have

$$\langle e_\lambda, f_\mu \rangle = \langle \omega(h_\lambda), \omega(m_\mu) \rangle = \langle h_\lambda, m_\mu \rangle = \delta_{\lambda\mu}. \qquad \square$$

We can now deduce an important characterizations of the Schur functions.

Proposition 3.3.8. *The Schur functions s_λ, $\lambda \in \mathrm{Par}(k)$, are the unique orthonormal basis of Sym^k up to sign and reordering.*

Proof. Suppose $(u_\lambda)_{\lambda \in \mathrm{Par}(k)}$ is another orthonormal basis for Sym^k. Then, for all $\lambda \in \mathrm{Par}(k)$, we have

$$u_\lambda = \sum_{\mu \in \mathrm{Par}(k)} c_{\lambda\mu} s_\mu \qquad \text{for some} \quad c_{\lambda\mu} \in \mathbb{Z}.$$

So $C = (c_{\lambda\mu})_{\lambda,\mu \in \mathrm{Par}(k)}$ is an integer matrix. (We view this as a matrix by choosing some ordering of $\mathrm{Par}(k)$.) Furthermore C is orthogonal, since it is the change of basis matrix between two orthonormal bases. However, the only orthogonal integer matrices are signed permutation matrices (that is, matrices that have exactly one entry ± 1 in each row and column, with all other entries being zero); see Exercise 3.3.1. Thus the u_λ are a permutation of the s_λ up to sign. $\square$

Exercises.

3.3.1. Prove that the only orthogonal integer matrices are signed permutation matrices (that is, matrices that have exactly one entry ± 1 in each row an column, with all other entries being zero).

3.3.2. Recall the forgotten symmetric functions f_λ from (2.34). Express $\sum_{\lambda \in \mathrm{Par}} e_\lambda(x) f_\lambda(y)$ as an infinite product.

3.3.3. Compute $\langle e_n, h_n \rangle$ for $n \in \mathbb{N}$.

3.3.4. Compute $\langle h_n, p_n \rangle$ for $n \in \mathbb{N}$.

3.3.5. Compute $\langle e_n, p_n \rangle$ for $n \in \mathbb{N}$.

3.3.6. Give a combinatorial formula for $\langle h_\lambda, h_\mu \rangle$, $\lambda, \mu \in \text{Par}$. More precisely, show that $\langle h_\lambda, h_\mu \rangle$ is equal to the number of Young tableaux satisfying certain conditions. *Hint*: Use Proposition 2.5.5.

3.4 Skew Schur functions and semistandard tableaux

In Propositions 2.4.5 and 2.5.6 we gave explicit expressions for the elementary and complete homogeneous symmetric functions in terms of tableaux. These expressions were closely related to the expressions of these symmetric functions as linear combinations of monomial symmetric functions given in Propositions 2.4.4 and 2.5.5. In this section, our goal is to develop such expressions for Schur functions. In turns out that this is most easily done if we slightly generalize our definition of Schur functions, to consider *skew* Schur functions. We follow the presentation of [Mac15, §I.5] but add some additional details.

Since the Schur functions form an orthonormal basis for Sym, every symmetric function is uniquely determined by its scalar product with the s_λ. In particular, we have

$$f = \sum_{\lambda \in \text{Par}} \langle f, s_\lambda \rangle s_\lambda \qquad \text{for all} \quad f \in \text{Sym}. \tag{3.33}$$

Definition 3.4.1 (Skew Schur functions)**.** For $\lambda, \mu \in \text{Par}$, define the *skew Schur function*

$$s_{\lambda/\mu} := \sum_{\nu \in \text{Par}} \langle s_\lambda, s_\mu s_\nu \rangle s_\nu \in \text{Sym}. \tag{3.34}$$

Remark 3.4.2. The equation (3.34) implies that $s_{\lambda/\mu}$ is obtained by acting on s_λ with the operator adjoint to multiplication by s_μ. More precisely, for $f \in \text{Sym}$ and $\nu, \lambda \in \text{Par}$, we have

$$\langle s_{\nu/\lambda}, f \rangle \overset{(3.34)}{=} \sum_{\rho \in \text{Par}} \langle s_\nu, s_\lambda s_\rho \rangle \langle s_\rho, f \rangle = \left\langle s_\nu, s_\lambda \sum_{\rho \in \text{Par}} \langle s_\rho, f \rangle s_\rho \right\rangle = \langle s_\nu, s_\lambda f \rangle. \tag{3.35}$$

We will discuss these adjoint operators in more detail in Section 4.2. △

Define integers $c^\lambda_{\mu\nu}$ for $\lambda, \mu, \nu \in \text{Par}$ by

$$s_\mu s_\nu = \sum_{\lambda \in \text{Par}} c^\lambda_{\mu\nu} s_\lambda. \tag{3.36}$$

Since s_λ is homogeneous of degree $|\lambda|$, we have

$$c^\lambda_{\mu\nu} = 0 \quad \text{if} \quad |\lambda| \neq |\mu| + |\nu|. \tag{3.37}$$

The integers $c^\lambda_{\mu\nu}$ are called *Littlewood–Richardson* coefficients, and we will explore them further in Section 3.6.

Lemma 3.4.3. *Suppose $\lambda, \mu \in \mathrm{Par}$. Recall that $\varnothing = (0,0,0,\dots)$ denotes the zero partition.*

(a) *We have $s_{\lambda/\varnothing} = s_\lambda$.*

(b) *We have*

$$s_{\lambda/\mu} = \sum_{\nu \in \mathrm{Par}} c^\lambda_{\mu\nu} s_\nu. \tag{3.38}$$

(c) *The skew Schur function $s_{\lambda/\mu}$ is homogeneous of degree $|\lambda| - |\mu|$. In particular, it is zero if $|\lambda| < |\mu|$.*

Proof. (a) Since $s_\varnothing = 1$, we have

$$s_{\lambda/\varnothing} \overset{(3.34)}{=} \sum_{\nu \in \mathrm{Par}} \langle s_\lambda, s_\nu \rangle s_\nu \overset{(3.33)}{=} s_\lambda.$$

(b) We have

$$s_{\lambda/\mu} \overset{(3.34)}{=} \sum_{\nu \in \mathrm{Par}} \langle s_\lambda, s_\mu s_\nu \rangle s_\nu \overset{(3.36)}{=} \sum_{\nu,\sigma \in \mathrm{Par}} c^\sigma_{\mu\nu} \langle s_\lambda, s_\sigma \rangle s_\nu \overset{(3.30)}{=} \sum_{\nu \in \mathrm{Par}} c^\lambda_{\mu\nu} s_\nu.$$

(c) This follows from (3.37) and (3.38). □

By Theorems 2.4.6 and 2.5.3, the skew Schur functions can be written as polynomials in the h_r or the e_r. The following result, which is a generalization of the Jacobi–Trudi identities (Theorem 3.2.5), gives these expression explicitly.

Proposition 3.4.4. *For $\lambda, \mu \in \mathrm{Par}$, we have*

$$s_{\lambda/\mu} = \det\left(h_{\lambda_i - \mu_j - i + j}\right)_{1 \le i,j \le n} \qquad \text{for } n \ge \ell(\lambda), \ell(\mu) \tag{3.39}$$

and

$$s_{\lambda/\mu} = \det\left(e_{\lambda'_i - \mu'_j - i + j}\right)_{1 \le i,j \le m} \qquad \text{for } m \ge \ell(\lambda'), \ell(\mu'). \tag{3.40}$$

Proof. By (3.12), the determinants appearing in (3.39) and (3.40) are equal, and so it suffices to prove (3.39).

Let $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ be two sets of variables. Then

$$\sum_{\lambda \in \mathrm{Par}} s_{\lambda/\mu}(x) s_\lambda(y) \overset{(3.38)}{=} \sum_{\lambda,\nu \in \mathrm{Par}} c^\lambda_{\mu\nu} s_\nu(x) s_\lambda(y) \overset{(3.36)}{=} \sum_{\nu \in \mathrm{Par}} s_\nu(x) s_\mu(y) s_\nu(y).$$

Now, by (3.20) and (3.28), we have

$$\sum_{\nu \in \mathrm{Par}} s_\nu(x) s_\nu(y) = \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\nu \in \mathrm{Par}} h_\nu(x) m_\nu(y).$$

Thus we have

$$\sum_{\lambda\in\mathrm{Par}} s_{\lambda/\mu}(x)s_\lambda(y) = \sum_{\nu\in\mathrm{Par}} h_\nu(x)m_\nu(y)s_\mu(y). \tag{3.41}$$

Now work over n indeterminates $y = (y_1, \ldots, y_n)$, so that sums above are only over partitions of length $\leq n$ (recall (3.8)). Then, using the definition (3.7) of the Schur polynomials and multiplying both sides of (3.41) by a_{δ^n}, we have

$$\begin{aligned}
\sum_{\lambda\in\mathrm{Par}_n} s_{\lambda/\mu}(x)a_{\lambda+\delta^n}(y) &= \sum_{\nu\in\mathrm{Par}_n} h_\nu(x)m_\nu(y)a_{\mu+\delta^n}(y) \\
&\overset{(3.1)}{=} \sum_{\nu\in\mathrm{Par}_n} h_\nu(x)m_\nu(y) \sum_{\pi\in\mathfrak{S}_n} \mathrm{sgn}(\pi)y^{(\mu+\delta^n)^\pi} \\
&\overset{(2.3)}{=} \sum_{\alpha\in\mathrm{WComp}_n} h_\alpha(x)y^\alpha \sum_{\pi\in\mathfrak{S}_n} \mathrm{sgn}(\pi)y^{(\mu+\delta^n)^\pi} \\
&= \sum_{\alpha\in\mathrm{WComp}_n} \sum_{\pi\in\mathfrak{S}_n} h_\alpha(x)\,\mathrm{sgn}(\pi)y^{\alpha+(\mu+\delta^n)^\pi},
\end{aligned}$$

where we used the definition (3.19) of h_α for $\alpha \in \mathbb{Z}^n$. Equating coefficients of $y^{\lambda+\delta^n}$, we have

$$\begin{aligned}
s_{\lambda/\mu}(x) &= \sum_{\pi\in\mathfrak{S}_n} \mathrm{sgn}(\pi)h_{\lambda+\delta^n-(\mu+\delta^n)^\pi}(x) \\
&= \sum_{\pi\in\mathfrak{S}_n} \mathrm{sgn}(\pi) \prod_{i=1}^{n} h_{(\lambda+\delta^n)_i-(\mu+\delta^n)_{\pi(i)}} \\
&\overset{(3.2)}{=} \det\left(h_{\lambda_i-\mu_j-i+j}\right)_{1\leq i,j\leq n}.
\end{aligned}$$

□

Corollary 3.4.5. *For all $\lambda, \mu \in \mathrm{Par}$, we have*

$$\omega(s_{\lambda/\mu}) = s_{\lambda'/\mu'}. \tag{3.42}$$

Proof. Choose $n \geq \ell(\lambda), \ell(\mu)$. Then

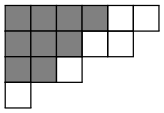
$$\omega(s_{\lambda/\mu}) \overset{(3.39)}{=} \det\left(\omega(h_{\lambda_i-\mu_j-i+j})\right)_{1\leq i,j\leq n} = \det\left(e_{\lambda_i-\mu_j-i+j}\right)_{1\leq i,j\leq n} \overset{(3.40)}{=} s_{\lambda'/\mu'}.$$

□

Define a relation $\supseteq$ on Par as follows: for $\lambda, \mu \in \mathrm{Par}$, we define

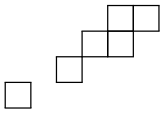
$$\lambda \supseteq \mu \iff \lambda_i \geq \mu_i \text{ for all } i. \tag{3.43}$$

Equivalently, $\lambda \supseteq \mu$ if and only if the Young diagram of λ contains the Young diagram of μ. If $\lambda \supseteq \mu$, then we identify λ/μ with a *skew Young diagram* (sometimes just called a *skew diagram*), which is the set theoretic difference of the Young diagrams of λ and μ. For example, if $\lambda = (6, 5, 3, 1)$ and $\mu = (4, 3, 2)$ then we can nest their Young diagrams as follows:



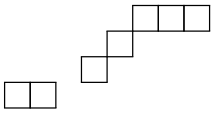
Then the skew diagram λ/μ is drawn as follows:

(3.44)



A *path* in a skew diagram θ is a sequence $x_0, x_1, \ldots, x_m$ of boxes in θ such that x_{i-1} and x_i have a common side, for $1 \leq i \leq m$. A subset φ of θ is *connected* if any two squares in φ can be joined by a path in φ. The *connected components* of θ are the maximal connected subsets of θ, which are themselves skew diagrams. For example, the skew diagram (3.44) has three connected components.

The *conjugate* of a skew diagram λ/μ is the skew diagram λ'/μ'. We say that a skew diagram is a *horizontal strip* if it has at most one square in each column. Thus λ/μ is a horizontal strip if and only if $\lambda'_i - \mu'_i \leq 1$ for all $i \geq 1$. For example, the skew diagram (3.44) is not a horizontal strip, but the following skew diagram is:



We say that λ/μ is a *vertical strip* if it has at most one square in each row. Equivalently, λ/μ is a vertical strip if and only if λ'/μ' is a horizontal strip.

Proposition 3.4.6. *Suppose* $\lambda, \mu \in \mathrm{Par}$.

(a) *We have*

$$s_{\lambda/\mu} = 0 \text{ unless } \lambda \supseteq \mu. \tag{3.45}$$

(b) *If* $\theta^1, \theta^2, \ldots, \theta^r$ *are the components of the skew diagram* λ/μ, *then* $s_{\lambda/\mu} = \prod_{i=1}^{r} s_{\theta^i}$.

Proof. Fix $n \geq \ell(\lambda), \ell(\mu)$.

(a) Suppose $\lambda \not\supseteq \mu$. Then $\lambda_r < \mu_r$ for some r, and so

$$\lambda_i \leq \lambda_r < \mu_r \leq \mu_j \qquad \text{for} \quad 1 \leq j \leq r \leq i \leq n.$$

Therefore, for $1 \leq j \leq r \leq i \leq n$, we have $\lambda_i - \mu_j - i + j < 0$ and so $h_{\lambda_i - \mu_j - i + j} = 0$. Therefore, in the matrix appearing in (3.39), the first r entries of the last $n - r + 1$ rows are zero. This implies that these rows are linearly dependent, and thus the determinant is zero.

(b) Note that as we move down the rows of the skew diagram λ/μ, we start a new connected component whenever $\mu_r \geq \lambda_{r+1}$ for some $1 \leq r < n$. In this case, we have

$$\lambda_i \leq \lambda_{r+1} \leq \mu_r \leq \mu_j \qquad \text{for} \quad 1 \leq j \leq r < i \leq n,$$

and so

$$\lambda_i - \mu_j - i + j < 0 \qquad \text{for} \quad 1 \leq j \leq r < i \leq n.$$

Therefore, the matrix $(h_{\lambda_i-\mu_j-i+j})_{1\le i,j\le n}$ can be written in block form

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix},$$

where A is of size $r\times r$ and B is of size $(n-r)\times(n-r)$. Thus

$$\begin{aligned} s_{\lambda/\mu} = \det(h_{\lambda_i-\mu_j-i+j})_{1\le i,j\le n} &= \det(A)\det(B) \\ &= \det(h_{\lambda_i-\mu_j-i+j})_{1\le i,j\le r}\det(h_{\lambda_i-\mu_j-i+j})_{r+1\le i,j\le n} = s_{\theta^1}s_{\theta^2}, \end{aligned}$$

where

$$\theta^1 = (\lambda_1,\dots,\lambda_r)/(\mu_1,\dots,\mu_r) \quad \text{and} \quad \theta^2 = (\lambda_{r+1},\dots,\lambda_n)/(\mu_{r+1},\dots,\mu_n).$$

The result then follows by induction. □

Example 3.4.7. Suppose λ/μ is a horizontal strip and let $\nu_1,\dots,\nu_r$ be the lengths of the components of the strip. Then, by Proposition 3.4.6, we have

$$s_{\lambda/\mu} = s_{(\nu_1)}s_{(\nu_2)}\cdots s_{(\nu_r)} \overset{(3.16)}{=} h_\nu = h_{\nu_1}h_{\nu_2}\cdots h_{\nu_r}.$$

Similarly, if λ/μ is a vertical strip, then

$$s_{\lambda/\mu} = e_\nu = e_{\nu_1}e_{\nu_2}\cdots e_{\nu_r},$$

where ν_i are the lengths of the components of the strip. △

In finitely many variables, we have an even stricter condition for the skew Schur functions to be nonzero.

Lemma 3.4.8. *We have*

$$s_{\lambda/\mu}(x_1,\dots,x_n) = 0 \text{ unless } 0\le \lambda'_i-\mu'_i\le n \text{ for all } i\ge 1.$$

Proof. We already know that $s_{\lambda/\mu}=0$ if $\lambda'_r<\mu'_r$ for some $r\ge 1$ by Proposition 3.4.6(a). Now suppose that $\lambda'_r-\mu'_r>n$ for some $r\ge 1$. In particular, $\lambda'_r>0$, and so $r\le\lambda_1=\ell(\lambda')$. Choose $m\ge\ell(\lambda'),\ell(\mu')$, so that $r\le m$. Then we have

$$\lambda'_i-\mu'_j\ge\lambda'_r-\mu'_r>n \qquad \text{for} \quad 1\le i\le r\le j\le m.$$

Therefore

$$\lambda'_i-\mu'_j-i+j>n \qquad \text{for} \quad 1\le i\le r\le j\le m.$$

Since $e_k(x_1,\dots,x_n)=0$ for $k>n$, this implies that the matrix $\left(e_{\lambda'_i-\mu'_j-i+j}(x_1,\dots,x_n)\right)_{1\le i,j\le m}$ has an $r\times(m+1-r)$ block of zeros in the top-right corner. So it has block form

$$\begin{pmatrix} A & 0_{(r-1)\times 1} & 0_{(r-1)\times(m-r)} \\ B & 0_{1\times 1} & 0_{1\times(m-r)} \\ C & D & E \end{pmatrix},$$

where $0_{p\times q}$ denotes a $p\times q$ zero matrix, A has size $(r-1)\times(r-1)$, and E has size $(m-r)\times(m-r)$. It follows from (3.40) that

$$s_{\lambda/\mu}(x_1,\dots,x_n) = \det\left(e_{\lambda'_i-\mu'_j-i+j}(x_1,\dots,x_n)\right)_{1\le i,j\le m} = \det(A)\det(0_{1\times 1})\det(E) = 0. \quad \square$$

Let $x = (x_1, x_2, \dots)$, $y = (y_1, y_2, \dots)$, $z = (z_1, z_2, \dots)$ be three sets of independent variables and let $s_\lambda(x, y)$ denote the Schur function corresponding to λ in the set of variables $(x_1, y_1, x_2, y_2, \dots)$.

Proposition 3.4.9. *For $\lambda, \mu \in \mathrm{Par}$ with $\lambda \supseteq \mu$, we have*

$$s_{\lambda/\mu}(x, y) = \sum_{\nu : \lambda \supseteq \nu \supseteq \mu} s_{\lambda/\nu}(x) s_{\nu/\mu}(y). \tag{3.46}$$

Proof. We have

$$\begin{aligned}
\sum_{\lambda,\nu \in \mathrm{Par}} s_{\lambda/\nu}(x) s_\lambda(z) s_\nu(y) &\overset{(3.38)}{=} \sum_{\lambda,\nu,\mu \in \mathrm{Par}} c^\lambda_{\nu\mu} s_\mu(x) s_\lambda(z) s_\nu(y) \\
&\overset{(3.36)}{=} \sum_{\nu,\mu \in \mathrm{Par}} s_\mu(x) s_\nu(z) s_\mu(z) s_\nu(y) \\
&\overset{(3.28)}{=} \left(\prod_{i,k=1}^{\infty} (1 - x_i z_k)^{-1} \right) \left(\prod_{j,k=1}^{\infty} (1 - y_j z_k)^{-1} \right) \\
&= \prod_{k=1}^{\infty} \left(\prod_{i=1}^{\infty} (1 - x_i z_k)^{-1} \prod_{j=1}^{\infty} (1 - y_j z_k)^{-1} \right) \\
&\overset{(3.28)}{=} \sum_{\lambda \in \mathrm{Par}} s_\lambda(x, y) s_\lambda(z).
\end{aligned}$$

Comparing coefficients of $s_\lambda(z)$ then gives

$$s_\lambda(x, y) = \sum_{\nu \in \mathrm{Par}} s_{\lambda/\nu}(x) s_\nu(y) \overset{(3.45)}{=} \sum_{\nu \subseteq \lambda} s_{\lambda/\nu}(x) s_\nu(y). \tag{3.47}$$

This is the special case of (3.46) where $\mu = \varnothing$. Then

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

In the first equality above, we have used (3.47) after replacing the variables x with $(x_1, y_1, x_2, y_2, \dots)$ and the variables y with z. The second and third equalities are is similar. Finally, comparing coefficients of $s_\mu(z)$ yields (3.46). □

Corollary 3.4.10. *Suppose $x^{(1)}, \dots, x^{(n)}$ are n sets of variables and $\lambda, \mu \in \mathrm{Par}$. Then*

$$s_{\lambda/\mu}\left(x^{(1)}, \dots, x^{(n)}\right) = \sum_{(\nu)} \prod_{i=1}^{n} s_{\nu^i/\nu^{i-1}}\left(x^{(i)}\right), \tag{3.48}$$

where the sum is over all sequences $(\nu) = (\nu^0, \nu^1, \dots, \nu^n)$ of partitions such that

$$\mu = \nu^0 \subseteq \nu^1 \subseteq \cdots \subseteq \nu^n = \lambda.$$

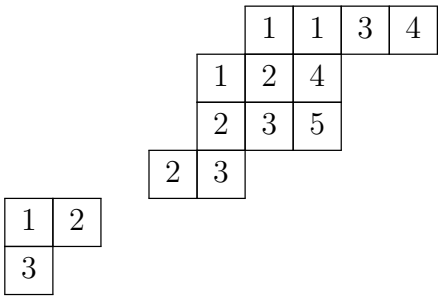
Proof. This follows from Proposition 3.4.9 by induction on n. □

Recall the definition of semistandard tableaux from Section 2.3. We define semistandard skew tableaux in an analogous way. Precisely, for $\lambda, \mu \in \mathrm{Par}$ with $\lambda \supseteq \mu$, a *semistandard skew tableau* of shape λ/μ is a filling of the boxes of the skew diagram λ/μ with positive integers such that the entries are weakly increasing along rows and strictly increasing down columns. We let $\mathrm{SST}(\lambda/\mu)$ denote the set of semistandard skew tableaux of shape λ/μ and, for $n \geq 1$, we let $\mathrm{SST}(\lambda/\mu, n)$ denote the set of such skew tableaux with entries chosen from the set $\{1, 2, \ldots, n\}$. We set $\mathrm{SST}(\lambda/\mu) = \mathrm{SST}(\lambda/\mu, n) = \varnothing$ for all $n \geq 1$ whenever $\lambda \not\supseteq \mu$. The weight $\mathrm{wt}(T)$ of a skew tableau is defined as for usual tableaux and we again use the notation $x^T := x^{\mathrm{wt}(T)}$.

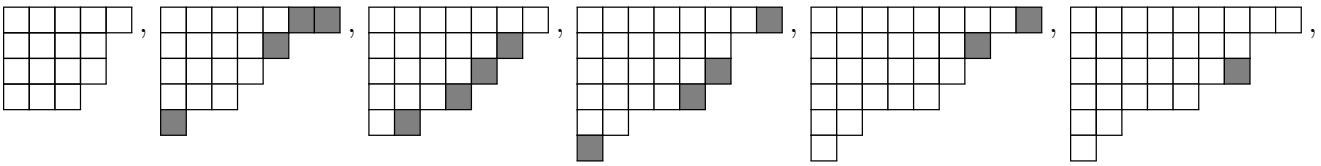
Now note that a sequence $(\nu) = (\nu^0, \nu^1, \ldots, \nu^n)$ of partitions with

$$\mu = \nu^0 \subseteq \nu^1 \subseteq \cdots \subseteq \nu^n = \lambda$$

such that each ν^i/ν^{i-1} is a horizontal strip is equivalent to a semistandard skew tableau of shape λ/μ with entries chosen from $\{1, 2, \ldots, n\}$. Namely, the skew tableau ν^i corresponds to the set of boxes with entries $\leq i$. For example, the semistandard skew tableau



corresponds to the sequence



where we have indicated in gray the boxes that have been added at each step.

Theorem 3.4.11 (Tableau description of skew Schur functions)**.** *For $\lambda, \mu \in \mathrm{Par}$ and $n \geq 1$, we have*

$$s_{\lambda/\mu} = \sum_{T \in \mathrm{SST}(\lambda/\mu)} x^T \qquad \text{and} \qquad s_{\lambda/\mu}(x_1, \ldots, x_n) = \sum_{T \in \mathrm{SST}(\lambda/\mu, n)} x^T. \tag{3.49}$$

Proof. It suffices to prove the second equality, since first then follows by taking the limit $n \to \infty$. For $1 \leq i \leq n$, let $x^{(i)}$ consist of the single variable x_i. Then Corollary 3.4.10 gives

$$s_{\lambda/\mu}(x_1, \ldots, x_n) = \sum_{(\nu)} \prod_{i=1}^{n} s_{\nu^i/\nu^{i-1}}(x_i).$$

Now, for a single variable x_i, it follows from Lemma 3.4.8 that $s_{\nu^i/\nu^{i-1}}(x_i)$ is zero unless ν^i/ν^{i-1} is a horizontal strip, in which case $s_{\nu^i/\nu^{i-1}}(x_i) = x_i^{|\nu^i|-|\nu^{i-1}|}$. Thus, each sequence (ν) appearing in the sum above corresponds to a tableau, and we have

$$s_{\lambda/\mu}(x_1,\dots,x_n) = \sum_{(\nu)} \prod_{i=1}^{n} x_i^{|\nu^i|-|\nu^{i-1}|} = \sum_{T\in \mathrm{SST}(\lambda/\mu,n)} x^T. \qquad \square$$

Corollary 3.4.12. *For $\lambda \in \mathrm{Par}$ and $n \geq 1$, we have*

$$s_\lambda = \sum_{T\in \mathrm{SST}(\lambda)} x^T \qquad \text{and} \qquad s_\lambda(x_1,\dots,x_n) = \sum_{T\in \mathrm{SST}(\lambda,n)} x^T. \tag{3.50}$$

Proof. This follows from taking $\mu = \varnothing$ in (3.49). $\square$

Example 3.4.13. The set $\mathrm{SST}((2,1,1),4)$ consists of the tableaux

$$\begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1} 3\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 1&2\\\hline 2\\\cline{1-1} 3\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1} 3\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 1&4\\\hline 2\\\cline{1-1} 3\\\cline{1-1}\end{array},$$

$$\begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1} 4\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 1&2\\\hline 2\\\cline{1-1} 4\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1} 4\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 1&4\\\hline 2\\\cline{1-1} 4\\\cline{1-1}\end{array},$$

$$\begin{array}{|c|c|}\hline 1&1\\\hline 3\\\cline{1-1} 4\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1} 4\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 1&3\\\hline 3\\\cline{1-1} 4\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 1&4\\\hline 3\\\cline{1-1} 4\\\cline{1-1}\end{array},$$

$$\begin{array}{|c|c|}\hline 2&2\\\hline 3\\\cline{1-1} 4\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 2&3\\\hline 3\\\cline{1-1} 4\\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline 2&4\\\hline 3\\\cline{1-1} 4\\\cline{1-1}\end{array}.$$

Thus

$$\begin{aligned} s_{(2,1,1)}(x_1,x_2,x_3,x_4) = {} & 3x_1x_2x_3x_4 + x_1^2x_2x_3 + x_1x_2^2x_3 + x_1x_2x_3^2 + x_1^2x_2x_4 + x_1x_2^2x_4 \\ & + x_1x_2x_4^2 + x_1^2x_3x_4 + x_1x_3^2x_4 + x_1x_3x_4^2 + x_2^2x_3x_4 + x_2x_3^2x_4 + x_2x_3x_4^2. \end{aligned} \qquad \triangle$$

Definition 3.4.14 (Kostka numbers)**.** For $\lambda,\mu,\nu \in \mathrm{Par}$ with $\lambda \supseteq \mu$, the *Kostka number* $K_{\lambda/\mu,\nu}$ is the number of semistandard skew tableaux of shape λ/μ and weight ν. We write $K_{\lambda,\mu}$ for $K_{\lambda/\varnothing,\mu}$.

The Kostka numbers are the coefficients giving the skew Schur functions in terms of the monomial symmetric functions.

Proposition 3.4.15. *For $\lambda,\mu \in \mathrm{Par}$ with $\lambda \supseteq \mu$, we have*

$$s_{\lambda/\mu} = \sum_{\nu\in\mathrm{Par}} K_{\lambda/\mu,\nu} m_\nu. \tag{3.51}$$

Proof. This follows from Theorem 3.4.11. □

Example 3.4.16. From Example 3.4.13, we see that

$$K_{(2,1,1),(1^4)} = 3, \quad K_{(2,1,1),(2,1,1)} = 1, \quad K_{(2,2,1),(2,2)} = 0, \quad K_{(2,2,1),(3,1)} = 0, \quad K_{(2,2,1),(4)} = 0.$$

Thus

$$s_{(2,1,1)}(x_1, x_2, x_3) = 3m_{(1^4)} + m_{2,1,1}.$$ △

Exercises.

3.4.1. Write the Schur function $s_{(2,2)}$ as a linear combination of the monomial symmetric functions.

3.4.2. Write the Schur function $s_{(3,2)}$ as a linear combination of the monomial symmetric functions.

3.4.3. Write the skew Schur function $s_{(3,2)/(1)}$ as a linear combination of monomial symmetric functions.

3.4.4. Write the skew Schur function $s_{(3,3,1)/(2,1)}$ as a linear combination of monomial symmetric functions.

3.4.5. Suppose $\lambda, \mu, \nu \in \text{Par}$ with $\lambda \supseteq \mu$. Show that

$$\langle s_{\lambda/\mu}, h_\nu \rangle = K_{\lambda/\mu,\nu} \qquad \text{and} \qquad \langle s_{\lambda/\mu}, e_\nu \rangle = K_{\lambda'/\mu',\nu}.$$

3.5 Transition matrices

In Propositions 2.4.4, 2.5.5, 2.6.7, and 3.4.15 we have given explicit formulas expressing bases for symmetric functions in terms of the monomial symmetric functions. Our goal in this section to complete this picture and give formulas for converting from any basis to any other. We follow the presentation of [Mac15, §6].

We will work with matrices whose rows and columns are indexed by partitions of a positive integer n. In doing so, we will order the partitions of n in reverse lexicographic order, so that (n) comes first and (1^n) comes last. Note from Exercise 1.2.4 that

$$\lambda > \mu \implies \lambda >_{\text{lex}} \mu, \tag{3.52}$$

but the converse implication does not hold.

We say that a matrix $M = (M_{\lambda,\mu})_{\lambda,\mu \in \text{Par}(n)}$ is *dominant upper triangular* if

$$M_{\lambda,\mu} \neq 0 \implies \lambda \geq \mu.$$

By (3.52), a dominant upper triangular matrix is upper triangular in the usual sense. We say the matrix is *dominant upper unitriangular* if it is dominant upper triangular and $M_{\lambda,\lambda} = 1$ for all $\lambda \in \mathrm{Par}(n)$. Similarly, it is *dominant lower triangular* if

$$M_{\lambda,\mu} \neq 0 \implies \lambda \leq \mu$$

and *dominant lower unitriangular* if we also have $M_{\lambda,\lambda} = 1$ for all $\lambda \in \mathrm{Par}(n)$.

Let UT_n (respectively, $\mathrm{UT}_n^\intercal$) denote the set of dominant upper (respectively, lower) unitriangular matrices with integer entries, indexed by the partitions of n.

Lemma 3.5.1. *The sets* UT_n *and* $\mathrm{UT}_n^\intercal$ *are groups with respect to matrix multiplication.*

Proof. Suppose $M, N \in \mathrm{UT}_n$. Then

$$(MN)_{\lambda,\mu} = \sum_{\nu \in \mathrm{Par}(n)} M_{\lambda,\nu} N_{\nu,\mu} = \sum_{\substack{\nu \in \mathrm{Par}(n) \\ \lambda \geq \nu \geq \mu}} M_{\lambda,\nu} N_{\nu,\mu}$$

is zero unless these exists $\nu \in \mathrm{Par}(n)$ such that $\lambda \geq \nu \geq \mu$. Such a ν exists if and only if $\lambda \geq \mu$. Furthermore, we have $(MN)_{\lambda,\lambda} = M_{\lambda,\lambda} N_{\lambda,\lambda} = 1$. Hence $MN \in \mathrm{UT}_n$.

Since the determinant of a triangular matrix is the product of its diagonal entries, we have $\det(M) = 1$, and so M is invertible. Since M is dominant upper unitriangular, it can be transformed to the identity matrix using only row operations that add a multiple of rows indexed by a partition λ to rows indexed by partitions $\mu < \lambda$. Since this same series of row operations transforms the identity matrix to M^{-1}, we see that $M^{-1} \in \mathrm{UT}_n$. This completes the proof that UT_n is a group.

It remains to prove that $\mathrm{UT}_n^\intercal$ is a group. Note first that

$$A \in \mathrm{UT}_n \iff A^\intercal \in \mathrm{UT}_n^\intercal.$$

Suppose $A, B \in \mathrm{UT}_n^\intercal$. Then

$$B^\intercal A^\intercal \in \mathrm{UT}_n \implies AB = \left(B^\intercal A^\intercal\right)^\intercal \in \mathrm{UT}_n^\intercal.$$

Furthermore A is invertible since $A^\intercal \in \mathrm{UT}_n$ is. Finally,

$$(A^{-1})^\intercal = (A^\intercal)^{-1} \in \mathrm{UT}_n \implies A^{-1} \in \mathrm{UT}_n^\intercal.$$

Hence $\mathrm{UT}_n^\intercal$ is also a group. □

Let $J = (J_{\lambda,\mu})_{\lambda,\mu \in \mathrm{Par}(n)}$ be the *transposition matrix* with entries

$$J_{\lambda,\mu} = \begin{cases} 1 & \text{if } \lambda' = \mu, \\ 0 & \text{otherwise.} \end{cases}$$

Strictly speaking, we have a transposition matrix J for each n. But we will use the same notation J for all of these; the n will be clear from the context. Note that

$$J^2 = I \qquad \text{and} \qquad J^\intercal = J, \tag{3.53}$$

where I denotes the identity matrix.

Lemma 3.5.2. *Suppose* $\lambda, \mu \in \mathrm{Par}(n)$. *Then*

$$\lambda \geq \mu \iff \mu' \geq \lambda'.$$

Proof. Since $\lambda'' = \lambda$, it suffices to prove one implication. Suppose that $\mu' \not\geq \lambda'$. We wish to show that $\lambda \not\geq \mu$.

Let $i \geq 1$ be the smallest index such that

$$\lambda'_1 + \cdots + \lambda'_i > \mu'_1 + \cdots + \mu'_i. \tag{3.54}$$

Thus

$$\lambda'_1 + \cdots + \lambda'_j \leq \mu'_1 + \cdots + \mu'_j, \qquad 1 \leq j \leq i-1,$$

and so $\lambda'_i > \mu'_i$.

Let $l = \lambda'_i$ and $m = \mu'_i$. Since $|\lambda| = |\mu|$, it follows from (3.54) that

$$\lambda'_{i+1} + \lambda'_{i+2} + \cdots < \mu'_{i+1} + \mu'_{i+2} + \cdots . \tag{3.55}$$

Since $\lambda'_{i+1} + \lambda'_{i+2} + \cdots$ is the number of boxes in the diagram of λ lying to the right of the i-th column, we have

$$\lambda'_{i+1} + \lambda'_{i+2} + \cdots = \sum_{j=1}^{l} (\lambda_j - i).$$

Similarly,

$$\mu'_{i+1} + \mu'_{i+2} + \cdots = \sum_{j=1}^{m} (\mu_j - i).$$

Therefore it follows from (3.55) that

$$\sum_{j=1}^{m} (\mu_j - i) > \sum_{j=1}^{l} (\lambda_j - i) \geq \sum_{j=1}^{m} (\lambda_j - i), \tag{3.56}$$

where the second inequality holds since $l > m$ and $\lambda_j \geq i$ for $1 \leq j \leq l$. It follows from (3.56) that

$$\mu_1 + \cdots + \mu_m > \lambda_1 + \cdots + \lambda_m.$$

Hence $\lambda \not\geq \mu$, as desired. □

Corollary 3.5.3. *A matrix* M *is dominant upper triangular (respectively, dominant upper unitriangular) if and only if* JMJ *is dominant lower triangular (respectively, dominant lower unitriangular).*

Proof. Let $N = JMJ$. Then $N_{\lambda,\mu} = M_{\lambda',\mu'}$. Then the result follows from Lemma 3.5.2. □

If $(u_\lambda)_{\lambda \in \mathrm{Par}(n)}$ and $(v_\lambda)_{\lambda \in \mathrm{Par}(n)}$ are two $\mathbb{Q}$-bases of $\mathrm{Sym}^n_{\mathbb{Q}}$, we let $M(u,v) = (M_{\lambda,\mu}(u,v))_{\lambda,\mu \in \mathrm{Par}(n)}$ denote the matrix of coefficients in the equations

$$u_\lambda = \sum_{\mu \in \mathrm{Par}(n)} M_{\lambda,\mu}(u,v) v_\mu, \qquad \lambda \in \mathrm{Par}(n). \tag{3.57}$$

We call $M(u,v)$ the *transition matrix* from the basis $(u_\lambda)_{\lambda \in \mathrm{Par}(n)}$ to the basis $(v_\lambda)_{\lambda \in \mathrm{Par}(n)}$. It is an invertible matrix of rational numbers. For an invertible matrix M, we define

$$M^{-\mathsf{T}} := \left(M^{\mathsf{T}}\right)^{-1} = \left(M^{-1}\right)^{\mathsf{T}}.$$

Proposition 3.5.4. *Suppose* $(u_\lambda)_{\lambda\in\mathrm{Par}(n)}$, $(v_\lambda)_{\lambda\in\mathrm{Par}(n)}$, *and* $(w_\lambda)_{\lambda\in\mathrm{Par}(n)}$ *are all* $\mathbb{Q}$*-bases of* $\mathrm{Sym}^n_{\mathbb{Q}}$. *Let* $(u'_\lambda)_{\lambda\in\mathrm{Par}(n)}$ *and* $(v'_\lambda)_{\lambda\in\mathrm{Par}(n)}$ *be the bases dual to* $(u_\lambda)_{\lambda\in\mathrm{Par}(n)}$ *and* $(v_\lambda)_{\lambda\in\mathrm{Par}(n)}$, *respectively, with respect to the Hall inner product. Then the following equalities holds.*

$$M(u,v)M(v,w) = M(u,w), \tag{3.58}$$

$$M(v,u) = M(u,v)^{-1}, \tag{3.59}$$

$$M(u',v') = M(v,u)^{\mathsf{T}} = M(u,v)^{-\mathsf{T}}, \tag{3.60}$$

$$M(\omega(u),\omega(v)) = M(u,v). \tag{3.61}$$

Proof. The proof of this proposition is left as Exercise 3.5.1. □

Let

$$K = M(s,m). \tag{3.62}$$

Note that this notation matches (3.51) and recall from Definition 3.4.14 that $K_{\lambda,\mu}$, called a *Kostka number*, is the number of semistandard tableaux of shape λ and weight μ.

Theorem 3.5.5. *The transition matrices between the bases*

$$(e_\lambda)_{\lambda\in\mathrm{Par}(n)}, \quad (h_\lambda)_{\lambda\in\mathrm{Par}(n)}, \quad (m_\lambda)_{\lambda\in\mathrm{Par}(n)}, \quad (f_\lambda)_{\lambda\in\mathrm{Par}(n)}, \quad \text{and} \quad (s_\lambda)_{\lambda\in\mathrm{Par}(n)}$$

are given in the following table, where the entry in row u *and column* v *is* $M(u,v)$*:*

	e	h	m	f	s
e	I	$K^{\mathsf{T}}JK^{-\mathsf{T}}$	$K^{\mathsf{T}}JK$	$K^{\mathsf{T}}K$	$K^{\mathsf{T}}J$
h	$K^{\mathsf{T}}JK^{-\mathsf{T}}$	I	$K^{\mathsf{T}}K$	$K^{\mathsf{T}}JK$	K^{T}
m	$K^{-1}JK^{-\mathsf{T}}$	$K^{-1}K^{-\mathsf{T}}$	I	$K^{-1}JK$	K^{-1}
f	$K^{-1}K^{-\mathsf{T}}$	$K^{-1}JK^{-\mathsf{T}}$	$K^{-1}JK$	I	$K^{-1}J$
s	$JK^{-\mathsf{T}}$	$K^{-\mathsf{T}}$	K	JK	I

All of the matrices in this table have integer entries.

Proof. By (3.18) and Theorem 3.3.7, we have

$$s' = s, \qquad e' = f, \qquad f' = e, \qquad h' = m, \qquad m' = h.$$

Thus

$$M(s,h) \overset{(3.60)}{=} M(s',h')^{-\mathsf{T}} = M(s,m)^{-\mathsf{T}} \overset{(3.62)}{=} K^{-\mathsf{T}}.$$

By (3.17), we have

$$M(\omega(s),s) = J.$$

Thus

$$M(s,e) \overset{(3.61)}{=} M(\omega(s),\omega(e)) \overset{(2.31)}{=} M(\omega(s),h) \overset{(3.58)}{=} M(\omega(s),s)M(s,h) = JK^{-\mathsf{T}}.$$

Finally,

$$M(s,f) \overset{(3.60)}{=} M(s',f')^{-\mathsf{T}} = M(s,e)^{-\mathsf{T}} = \left(JK^{-\mathsf{T}}\right)^{-\mathsf{T}} = J^{-\mathsf{T}}K \overset{(3.53)}{=} JK.$$

We can now use (3.58) and (3.59) to compute all the other entries of the given table, going via the Schur functions. For example

$$M(m,e) \overset{(3.58)}{=} M(m,s)M(s,e) \overset{(3.59)}{=} M(s,m)^{-1}M(s,e) = K^{-1}JK^{-\mathsf{T}}.$$

The fact that all the matrices have integer entries follows from the fact that all of the bases are bases of Sym over $\mathbb{Z}$. □

We can now deduce various triangularity properties of our various bases.

Proposition 3.5.6. (a) *The matrix K is dominant upper unitriangular.*

(b) *The matrices $M(s,h)$ and $M(h,s)$ are dominant lower unitriangular.*

(c) *The matrices $M(s,m)$ and $M(m,s)$ are dominant upper unitriangular.*

Proof. (a) Suppose T is a semistandard tableau of shape λ and weight μ. Then, for each $r \geq 1$, there are

$$\mu_1 + \cdots + \mu_r$$

entries $\leq r$ in T. These must all be located in the top r rows of T since the entries of T strictly increase down columns. Thus

$$\mu_1 + \cdots + \mu_r \leq \lambda_1 + \cdots + \lambda_r \qquad \text{for all } r \geq 1.$$

Hence $\mu \leq \lambda$. Thus K is dominant upper triangular. The same argument also shows that $K_{\lambda,\lambda} = 1$ for all λ, and so K is dominant upper unitriangular.

(b) Since K is dominant upper unitriangular, K^{T} is dominant lower unitriangular. Thus $K^{-\mathsf{T}}$ is also dominant lower unitriangular by Lemma 3.5.1. Hence $M(s,h) = K^{-\mathsf{T}}$ and $M(h,s) = K^{\mathsf{T}}$ are both dominant lower unitriangular.

(c) This follows immediately from the fact that $M(s,m) = K$, $M(m,s) = K^{-1}$, and K is dominant upper unitriangular. □

We can give additional combinatorial interpretations of the transition matrices $M(e,m)$ and $M(h,m)$. (Compare to Propositions 2.4.4 and 2.5.5.) If M is a matrix, its *i-th row sum* is the sum of the entries in its i-th row. Similarly, its *i-th column sum* is the sum of the entries in its i-th column.

Proposition 3.5.7. *Suppose $\lambda, \mu \in \mathrm{Par}(n)$.*

(a) *$M(e,m)_{\lambda,\mu} = \sum_{\nu \in \mathrm{Par}(n)} K_{\nu,\lambda}K_{\nu',\mu}$ is the number of matrices of zeros and ones whose i-th row sum is λ_i and whose i-th column sum is μ_i.*

(b) *$M(h,m)_{\lambda,\mu} = \sum_{\nu \in \mathrm{Par}(n)} K_{\nu,\lambda}K_{\nu,\mu}$ is the number of matrices of nonnegative integers whose i-th row sum is λ_i and show i-th column sum is μ_i.*

Proof. (a) First note that

$$M(e,m)_{\lambda,\mu} = \sum_{\nu,\sigma \in \mathrm{Par}(n)} (K^{\mathsf{T}})_{\lambda,\nu}J_{\nu,\sigma}K_{\sigma,\mu} = \sum_{\nu,\sigma \in \mathrm{Par}(n)} K_{\nu,\lambda}\delta_{\nu'\sigma}K_{\sigma,\mu} = \sum_{\nu \in \mathrm{Par}(n)} K_{\nu,\lambda}K_{\nu',\mu}.$$

Now, for $\mu \in \mathrm{Par}(n)$, consider the coefficient of the monomial x^μ in $e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots$. Each monomial in e_{λ_i} is of the form $\prod_j x_j^{a_{ij}}$, where $a_{ij} \in \{0,1\}$, and $\sum_j a_{ij} = \lambda_i$. Thus we have

$$\prod_{i,j} x_j^{a_{ij}} = \prod_j x_j^{\mu_j},$$

and hence $\sum_i a_{ij} = \mu_j$. Thus the matrix (a_{ij}) has i-th row sum equal to λ_i and j-th column sum equal to μ_j.

(b) The proof is analogous. The only difference is that we replace e_λ by h_λ, and thus the exponents a_{ij} can now be arbitrary nonnegative integers. □

Theorem 3.5.5 is missing the basis of power sum symmetric functions. We now consider the transition matrices involving the $\mathbb{Q}$-basis $(p_\lambda)_{\lambda \in \mathrm{Par}}$. Let

$$L = M(p,m), \qquad \text{so that} \qquad p_\lambda = \sum_{\mu \in \mathrm{Par}(n)} L_{\lambda,\mu} m_\mu, \quad \lambda \in \mathrm{Par}(n). \tag{3.63}$$

By Proposition 2.6.7, $L_{\lambda,\mu}$ is the number of Young tableaux of shape λ and weight μ in which the entries in each row are constant. It follows that, if $L_{\lambda,\mu} \neq 0$, then each μ_i is a sum of parts of λ. If this is the case, we say that λ is a *refinement* of μ. For example

$$(5,5,4,3,3,2,1,1,1,1) \quad \text{is a refinement of} \quad (9,7,6,2,1,1),$$

where the blue parts of the first partition have been added to obtain the blue part of the second partition, and similarly for the other colours.

The notion of refinement can be explained in an equivalent way as follows. For partitions λ and μ, we let $\lambda \cup \mu$ denote the partition obtained by taking all the parts of λ and μ and arranging them in weakly decreasing order. Thus, for example,

$$(6,4,3,3,1) \cup (7,5,4,3,2,1,1) = (7,6,5,4,4,3,3,3,2,1,1,1).$$

Then λ is a refinement of μ if λ is of the form $\bigcup_{i \geq 1} \lambda^{(i)}$, where each $\lambda^{(i)}$ is a partition of μ_i.

We write $\lambda \leq_R \mu$ if λ is a refinement of μ. It is straightforward to verify that $\leq_R$ is a partial order on $\mathrm{Par}(n)$ for all $n \geq 0$.

Lemma 3.5.8. *For $\lambda, \mu \vdash n$, we have*

$$\lambda \leq_R \mu \implies \lambda \leq \mu.$$

Proof. Suppose $\lambda, \mu \vdash n$ and $\lambda \leq_R \mu$. Thus we have a function $f\colon \{1,2,\ldots,\ell(\lambda)\} \to \{1,2,\ldots,\ell(\mu)\}$ such that

$$\mu_j = \sum_{i : f(i) = j} \lambda_i, \qquad 1 \leq j \leq \ell(\mu).$$

Then, for $1 \leq k \leq \ell(\lambda)$, we have

$$\lambda_1 + \cdots + \lambda_k \leq \sum_{i \in f(\{1,2,\ldots,k\})} \mu_i \leq \mu_1 + \cdots + \mu_k,$$

where the last inequality holds since $f(\{1,2,\ldots,k\})$ contains at most k elements and $\mu_1 + \ldots + \mu_k$ is the sum of the k *largest* parts of μ. □

Corollary 3.5.9. *The matrix L is dominant lower triangular.*

Proof. As discussed above, Proposition 2.6.7 implies that $L_{\lambda,\mu} = 0$ unless $\lambda \leq_R \mu$. Thus, by Lemma 3.5.8, $L_{\lambda,\mu} = 0$ unless $\lambda \leq \mu$. Hence L is dominant lower triangular. □

Theorem 3.5.10. *The transition matrices involving the power sum symmetric functions are given by*

$$M(p, e) = \varepsilon z L^{-\mathsf{T}}, \tag{3.64}$$
$$M(p, f) = \varepsilon L, \tag{3.65}$$
$$M(p, h) = z L^{-\mathsf{T}}, \tag{3.66}$$
$$M(p, s) = L K^{-1}, \tag{3.67}$$

where ε and z are diagonal matrices given by

$$\varepsilon_{\lambda,\mu} = \delta_{\lambda\mu}(-1)^{|\lambda|-\ell(\lambda)}, \qquad z_{\lambda,\mu} = \delta_{\lambda\mu} z_\lambda.$$

Proof. By (3.18) and Theorem 3.3.7, the bases dual to $(h_\lambda)_{\lambda\in\mathrm{Par}(n)}$ and $(p_\lambda)_{\lambda\in\mathrm{Par}(n)}$ are $(m_\lambda)_{\lambda\in\mathrm{Par}(n)}$ and $(z_\lambda^{-1} p_\lambda)_{\lambda\in\mathrm{Par}(n)}$, respectively. Note also that

$$M(zu, v) = z M(u, v) \qquad \text{and} \qquad M(\varepsilon u, v) = \varepsilon M(u, v)$$

for any bases $(u_\lambda)_{\lambda\in\mathrm{Par}(n)}$ and $(u_\lambda)_{\lambda\in\mathrm{Par}(n)}$ of $\mathrm{Sym}^n_{\mathbb{Q}}$. (See Exercise 3.5.2.) Thus

$$\begin{aligned}
M(p, h) &\overset{(3.60)}{=} M(z^{-1}p, m)^{-\mathsf{T}} = (z^{-1}L)^{-\mathsf{T}} = z^{\mathsf{T}} L^{-\mathsf{T}} = z L^{-\mathsf{T}}, \\
M(p, e) &\overset{(3.61)}{=} M(\omega(p), \omega(e)) \underset{(2.31)}{\overset{(2.45)}{=}} M(\varepsilon p, h) = \varepsilon M(p, h) = \varepsilon z L^{-\mathsf{T}}, \\
M(p, f) &\overset{(3.61)}{=} M(\omega(p), \omega(f)) \underset{(2.34)}{\overset{(2.45)}{=}} M(\varepsilon p, m) = \varepsilon M(p, m) = \varepsilon L, \\
M(p, s) &\overset{(3.58)}{=} M(p, m) M(m, s) \overset{(3.59)}{=} M(p, m) M(s, m)^{-1} = L K^{-1}.
\end{aligned}$$

□

We will see later (in Proposition 6.3.8) that $M(p, s)$ gives the character table of the symmetric group.

Exercises.

3.5.1. Prove Proposition 3.5.4.

3.5.2. Suppose that $n \geq 0$ and that $(u_\lambda)_{\lambda\in\mathrm{Par}(n)}$ and $(v_\lambda)_{\lambda\in\mathrm{Par}(n)}$ are two $\mathbb{Q}$-bases of $\mathrm{Sym}^n_{\mathbb{Q}}$. Furthermore, suppose that $D = (D_{\lambda,\mu})_{\lambda,\mu\in\mathrm{Par}(n)}$ is a diagonal matrix. Show that

$$M(Du, v) = D M(u, v).$$

3.5.3. By Proposition 3.5.6(a) and Exercise 1.2.4, we have

$$K_{\lambda,\mu} \neq 0 \implies \lambda \geq_{\text{lex}} \mu.$$

Show that the converse of this implication is false in general. That is, find $\lambda, \mu \in \text{Par}$, $|\lambda| = |\mu|$, such that $\lambda >_{\text{lex}} \mu$ but $K_{\lambda,\mu} = 0$.

3.5.4. In this exercise, we consider partitions of $n = 3$.

(a) Write the partitions of n in reverse lexicographic order.

(b) Write down the matrices J, K, L, ε, and z.

(c) Compute $M(h, f)$, $M(s, e)$, $M(p, f)$, and $M(p, h)$.

(d) Write $p_{(1^n)}$ as a linear combination of the forgotten symmetric functions.

(e) Write p_n as a linear combination of the complete homogeneous symmetric functions.

3.5.5. Repeat Exercise 3.5.4 for $n = 4$.

3.6 The Littlewood-Richardson rule

Our goal in this section is to prove the Littlewood–Richardson rule, which tells us how to write the product of two Schur functions as a linear combination of Schur functions. There are many different proofs of this important result. We will follow the relatively short proof found in [Ste02]. Other proofs can be found in [Egg19, Ch. 10] and [Mac15, §I.9].

Before proving the full Littlewood–Richardson rule, we begin with two important special cases, known as Pieri's formulas, which are much easier to prove. These formulas are corollaries of the following result.

Proposition 3.6.1. *For $\lambda, \mu \in \text{Par}$, we have*

$$h_\mu s_\lambda = \sum_{\nu \supseteq \lambda} K_{\nu/\lambda,\mu} s_\nu \qquad \text{and} \qquad e_\mu s_\lambda = \sum_{\nu \supseteq \lambda} K_{\nu'/\lambda',\mu} s_\nu.$$

Proof. We have

$$h_\mu s_\lambda = \sum_{\nu \in \text{Par}} \langle s_\nu, h_\mu s_\lambda \rangle s_\nu \overset{(3.35)}{=} \sum_{\nu \supseteq \lambda} \langle s_{\nu/\lambda}, h_\mu \rangle s_\nu \overset{(3.51)}{=} \sum_{\substack{\nu \supseteq \lambda \\ \rho \in \text{Par}}} K_{\nu/\lambda,\rho} \langle m_\rho, h_\mu \rangle s_\nu \overset{(3.18)}{=} \sum_{\nu \supseteq \lambda} K_{\nu/\lambda,\mu} s_\nu.$$

This proves the first equality. Applying ω then gives the second equality. □

Corollary 3.6.2 (Pieri's formula)**.** *For $\lambda \in \text{Par}$ and $r \in \mathbb{N}$, we have*

$$h_r s_\lambda = \sum_\nu s_\nu, \tag{3.68}$$

where the sum is over all partitions $\nu \supseteq \lambda$ such that ν/λ is a horizontal strip with r boxes.

Proof. Consider Proposition 3.6.1 with $\mu = (r)$. Since $K_{\nu/\lambda,(r)}$ is the number of semistandard skew tableaux of shape ν/λ whose entries are all 1s, we have

$$K_{\nu/\lambda,(r)} = \begin{cases} 1 & \text{if } \nu/\lambda \text{ is a horizontal strip,} \\ 0 & \text{otherwise.} \end{cases}$$

The result follows. □

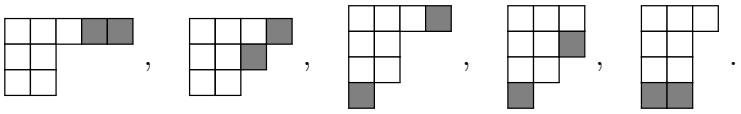
Corollary 3.6.3 (Dual Pieri's formula)**.** *For $\lambda \in \mathrm{Par}$ and $r \in \mathbb{N}$, we have*

$$e_r s_\lambda = \sum_\nu s_\nu, \tag{3.69}$$

where the sum is over all partitions $\nu \supseteq \lambda$ such that ν/λ is a vertical strip with r boxes.

Proof. We apply the involution ω to (3.68). □

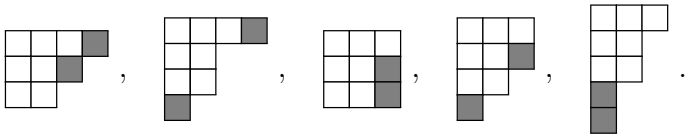
Example 3.6.4. The ways of adding a horizontal strip with 2 boxes to $(3,2,2)$ are



Thus we have

$$h_2 s_{(3,2,2)} = s_{(5,2,2)} + s_{(4,3,2)} + s_{(4,2,2,1)} + s_{(3,3,2,1)} + s_{(3,2,2,2)}.$$

On the other hand, the ways of adding a vertical strip with 2 boxes to $(3,2,2)$ are



Thus we have

$$e_2 s_{(3,2,2)} = s_{(4,3,2)} + s_{(4,2,2,1)} + s_{(3,3,3)} + s_{(3,3,2,1)} + s_{(3,2,2,1,1)}.$$ △

We now turn our attention to the full Littlewood–Richardson rule. A key element in the proof will be a certain set of involutions on tableaux, which we now describe.

Suppose T is a semistandard tableau and fix a positive integer j. We say that an entry j in T is *free* if there is no entry $j+1$ in the same column. Similarly, we say that an entry $j+1$ in T is *free* if there is no entry j in the same column. In any given row, the free entries must occur in consecutive columns. (See Exercise 3.6.1.) We define $\beta_j(T)$ to be the tableau obtained from T by interchanging the number of free j's and $(j+1)$'s in each row (always keeping j's to the left of $(j+1)$'s, so that the tableau remains semistandard). The operation β_j is called a *Bender–Knuth involution*. Since the non-free j's and $(j+1)$'s occur in pairs, we see that

$$\mathrm{wt}(\beta_j(T)) = \mathrm{wt}(T)^{\sigma_j}, \tag{3.70}$$

where σ_j is the transposition of j and $j+1$.

Example 3.6.5. Fix $j = 3$. In the following tableau, of all the entries equal to 3 or 4, the free entries are coloured green and the non-free entries are coloured red:

$$T = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1&1&1&1&1&2&2&2&2&2&2&3&4&6\\\hline 2&2&2&2&2&3&3&3&3&3&4&4&5&7\\\hline 3&3&3&4&4&4&5&5&6&6\\\cline{1-10} 4&4&5&5&6&7\\\cline{1-6} 6\\\cline{1-1}\end{array}$$

Applying the Bender–Knuth involution, we obtain

$$\beta_3(T) = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1&1&1&1&1&2&2&2&2&2&2&3&3&6\\\hline 2&2&2&2&2&3&3&4&4&4&4&4&5&7\\\hline 3&3&3&3&4&4&5&5&6&6\\\cline{1-10} 4&4&5&5&6&7\\\cline{1-6} 6\\\cline{1-1}\end{array}\,.$$

We have

$$\mathrm{wt}(T) = (5, 11, 9, 8, 5, 5, 2) \qquad \text{and} \qquad \mathrm{wt}(\beta_j(T)) = (5, 11, 8, 9, 5, 5, 2) = \mathrm{wt}(T)^{\sigma_3}. \qquad \triangle$$

For a skew tableau T and positive integer j, let $T_{\geq j}$ denote the subtableau of T consisting of columns $j, j+1, \ldots$. We define $T_{<j}$ and $T_{>j}$ analogously. For example, if

$$T = \begin{array}{|c|c|c|c|c|c|}\hline &&&1&2\\\cline{4-5} &&1&2&4\\\cline{3-5} &2&2&3\\\cline{2-4} 1&2\\\cline{1-2} 3\\\cline{1-1}\end{array}$$

then

$$T_{\geq 3} = \begin{array}{|c|c|c|c|}\hline &&1&2\\\cline{3-4} &1&2&4\\\cline{2-4} 2&2&3\\\cline{1-3}\end{array}\,, \quad T_{>3} = \begin{array}{|c|c|c|}\hline &1&2\\\cline{2-3} 1&2&4\\\cline{1-3} 2&3\\\cline{1-2}\end{array}\,, \quad \text{and} \quad T_{<3} = \begin{array}{|c|c|}\hline 1&2\\\cline{1-2} 3\\\cline{1-1}\end{array}\,.$$

Proposition 3.6.6. *For all $\lambda \in \mathrm{Par}_n$ and all $\mu, \nu \in \mathrm{Par}$ such that $\nu \subseteq \mu$, we have*

$$a_{\lambda+\delta^n}(x_1, \ldots, x_n) s_{\mu/\nu}(x_1, \ldots, x_n) = \sum_T a_{\lambda+\mathrm{wt}(T)+\delta^n}(x_1, \ldots, x_n), \tag{3.71}$$

where the sum is over all $T \in \mathrm{SST}(\mu/\nu, n)$ such that $\lambda + \mathrm{wt}(T_{\geq j}) \in \mathrm{Par}_n$ for all $j \geq 1$.

Proof. Fix $\lambda \in \mathrm{Par}_n$ and $\mu, \nu \in \mathrm{Par}$ such that $\nu \subseteq \mu$. Since $s_{\mu/\nu}$ is a symmetric function, we have

$$s_{\mu/\nu}(x_1, \ldots, x_n) = \pi^{-1} s_{\mu/\nu}(x_1, \ldots, x_n) \overset{(2.5)}{=} \sum_{T \in \mathrm{SST}(\mu/\nu, n)} x^{\mathrm{wt}(T)^\pi} \quad \text{for all } \pi \in \mathfrak{S}_n.$$

Thus

(3.72)
$$a_{\lambda+\delta^n} s_{\mu/\nu}(x_1, \ldots, x_n) = \sum_{\pi \in \mathfrak{S}_n} \sum_{T \in \mathrm{SST}(\mu/\nu, n)} \mathrm{sgn}(\pi) x^{(\lambda+\delta^n+\mathrm{wt}(T))^\pi} = \sum_{T \in \mathrm{SST}(\mu/\nu, n)} a_{\lambda+\mathrm{wt}(T)+\delta^n}.$$

We say that $T \in \mathrm{SST}(\mu/\nu, n)$ is a *gardyloo* if $\lambda + \mathrm{wt}(T_{\geq j}) \notin \mathrm{Par}_n$ for some $j \geq 1$. Therefore, T is a gardyloo if and only if

$$\lambda_k + \mathrm{wt}(T_{\geq j})_k < \lambda_{k+1} + \mathrm{wt}(T_{\geq j})_{k+1} \tag{3.73}$$

for some k, j. (Since $\mathrm{wt}(T)$ is the weight of T, $\mathrm{wt}(T)_k$ is the number of entries of T equal to k.) Our goal is to show that all gardyloo terms cancel in (3.72).

Suppose T is a gardyloo. Let j be the largest integer such $\lambda + \mathrm{wt}(T_{\geq j}) \notin \mathrm{Par}_n$. Then choose the smallest k such that (3.73) holds. Thus we have

$$\begin{aligned} &\lambda_k + \mathrm{wt}(T_{>j})_k \geq \lambda_{k+1} + \mathrm{wt}(T_{>j})_{k+1} \qquad \text{and} \\ &\lambda_k + \mathrm{wt}(T_{\geq j})_k < \lambda_{k+1} + \mathrm{wt}(T_{\geq j})_{k+1}. \end{aligned} \tag{3.74}$$

It follows that

$$\mathrm{wt}(T_{\geq j})_{k+1} - \mathrm{wt}(T_{>j})_{k+1} > \mathrm{wt}(T_{\geq j})_k - \mathrm{wt}(T_{>j})_k.$$

Now, $\mathrm{wt}(T_{\geq j})_i - \mathrm{wt}(T_{>j})_i$ is the number of i's in the j-th column of T. Since the entries in T strictly decrease down columns, this number is either zero or one. So we must have

$$\mathrm{wt}(T_{\geq j})_{k+1} - \mathrm{wt}(T_{>j})_{k+1} = 1 \qquad \text{and} \qquad \mathrm{wt}(T_{\geq j})_k - \mathrm{wt}(T_{>j})_k = 0.$$

It then follows from (3.74) that the j-th column of T contains a $k+1$ but no k, and that

$$\lambda_k + \mathrm{wt}(T_{\geq j})_k + 1 = \lambda_{k+1} + \mathrm{wt}(T_{\geq j})_{k+1}. \tag{3.75}$$

Let T^* denote the tableau obtained from T by applying the Bender–Knuth involution β_k to the subtableau $T_{<j}$ of T, leaving the remainder of T unchanged. This involves changing some subset of the entries of $T_{<j}$ from k to $k+1$ and vice-versa. Since the j-th column of T contains a $k+1$ but no k, the tableau T^* is semistandard. Furthermore, since $T_{\geq j} = (T^*)_{\geq j}$, we have $(T^*)^* = T$. Hence $T \mapsto T^*$ is an involution on the set of gardyloos.

Now we compare the contributions of T and T^* to the sum (3.72). It follows from (3.70) that

$$\mathrm{wt}(T_{<j})^{\sigma_k} = \mathrm{wt}\left((T^*)_{<j}\right).$$

Furthermore, it follows from (3.75) that

$$(\lambda + \mathrm{wt}(T_{\geq j}) + \delta^n)^{\sigma_k} = \lambda + \mathrm{wt}(T_{\geq j}) + \delta^n.$$

Thus

$$(\lambda + \mathrm{wt}(T) + \delta^n)^{\sigma_k} = \lambda + \mathrm{wt}(T^*) + \delta^n$$

and so

$$a_{\lambda+\mathrm{wt}(T)+\delta^n} = -\sigma_k a_{\lambda+\mathrm{wt}(T)+\delta^n} = -a_{\lambda+\mathrm{wt}(T^*)+\delta^n}.$$

(See Exercise 3.6.3.) Therefore, in the sum (3.72), the terms where T is a gardyloo all cancel. □

Corollary 3.6.7. *For all $\lambda \in \mathrm{Par}$ and all $\mu, \nu \in \mathrm{Par}$ such that $\nu \subseteq \mu$, we have*

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

where the sum ranges over all $T \in \mathrm{SST}(\mu/\nu)$ such that $\lambda + \mathrm{wt}(T_{\geq j}) \in \mathrm{Par}$ for all $j \geq 1$.

Proof. Fix $\lambda, \mu, \nu \in \mathrm{Par}$ such that $\nu \subseteq \mu$. Choose $n \geq |\lambda| + |\mu|$. Dividing both sides of (3.71) by a_{δ^n} gives that

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

where the sum ranges over all $T \in \mathrm{SST}(\mu/\nu, n)$ such that $\lambda + \mathrm{wt}(T_{\geq j}) \in \mathrm{Par}_n$ for all $j \geq 1$. Now note that, for $j \geq 1$,

$$\begin{aligned} \lambda + \mathrm{wt}(T_{\geq j}) \in \mathrm{Par} &\implies \ell(\lambda + \mathrm{wt}(T_{\geq j})) \leq |\lambda| + |\,\mathrm{wt}(T_{\geq j})| \leq |\lambda| + |\mu| \leq n \\ &\implies \lambda + \mathrm{wt}(T_{\geq j}) \in \mathrm{Par}_n. \end{aligned}$$

This also implies that any $T \in \mathrm{SST}(\mu/\nu)$ satisfying $\lambda + \mathrm{wt}(T_{\geq j}) \in \mathrm{Par}$ for all $j \geq 1$ lies in $\mathrm{SST}(\mu/\nu, n)$. □

Recall from (3.36) and (3.38) that the Littlewood–Richardson coefficients $c^\lambda_{\mu\nu}$ satisfy

$$s_\mu s_\nu = \sum_{\lambda \in \mathrm{Par}} c^\lambda_{\mu\nu} s_\lambda \qquad \text{and} \qquad s_{\lambda/\mu} = \sum_{\nu \in \mathrm{Par}} c^\lambda_{\mu\nu} s_\nu. \tag{3.76}$$

Definition 3.6.8 (Littlewood–Richardson tableau)**.** We say that a semistandard skew tableau T is a *Littlewood–Richardson tableau* if $\mathrm{wt}(T_{\geq j}) \in \mathrm{Par}$ for all $j \geq 1$.

Theorem 3.6.9 (Littlewood-Richardson rule)**.** *For $\lambda, \mu, \nu \in \mathrm{Par}$, the Littlewood–Richardson coefficient $c^\lambda_{\mu\nu}$ is equal to the number of Littlewood–Richardson tableaux of shape λ/μ and weight ν.*

Proof. If $\mu \not\subseteq \lambda$, then $s_{\lambda/\mu} = 0$ by (3.45), and so $c^\lambda_{\mu\nu} = 0$ for all $\nu \in \mathrm{Par}$ by (3.76), while there are no Littlewood–Richardson tableaux of shape λ/μ, since $\mathrm{SST}(\lambda/\mu) = \varnothing$. So we may assume that $\mu \subseteq \lambda$. Then, by Corollary 3.6.7 with (λ, μ, ν) replaced by $(\varnothing, \lambda, \mu)$, we have

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

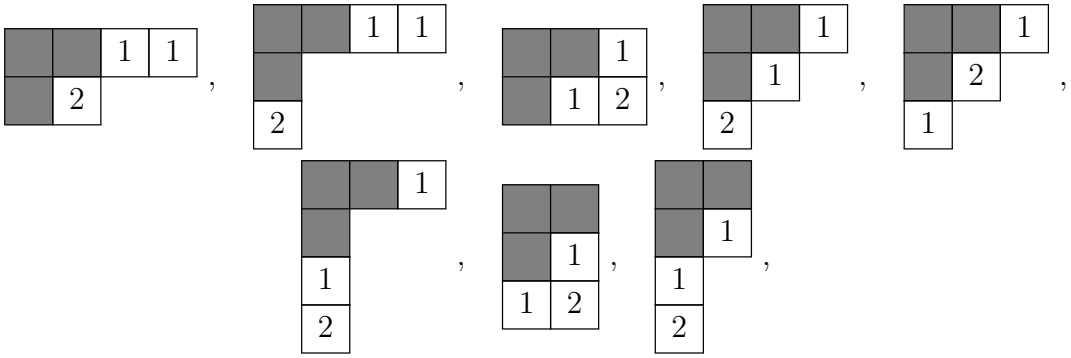
where the sum ranges over all $T \in \mathrm{SST}(\lambda/\mu)$ such that $\mathrm{wt}(T_{\geq j}) \in \mathrm{Par}$ for all $j \geq 1$. The result then follows from the second equality in (3.76). □

Example 3.6.10. We can recover Pieri's formula from the Littlewood–Richardson rule. Theorem 3.6.9 tells us that $c^{\nu}_{\lambda,(r)}$ is equal to the number of Littlewood–Richardson tableaux of shape ν/λ and weight (r). Since any semistandard skew tableau of weight (r) is automatically a Littlewood-Richardson tableau, $c^{\nu}_{\lambda,(r)}$ is equal to the number of semistandard skew tableaux of shape ν/λ and weight (r). There is one such tableau if ν/λ is a horizontal strip with r boxes, and no such tableaux otherwise. Thus

$$s_\lambda h_r = s_\lambda s_{(r)} = \sum_{\nu \in \mathrm{Par}} c^{\nu}_{\lambda,(r)} s_\nu = \sum_{\nu} s_\nu,$$

where the final sum is over all partitions $\nu \supseteq \lambda$ such that ν/λ is a horizontal strip with r boxes. △

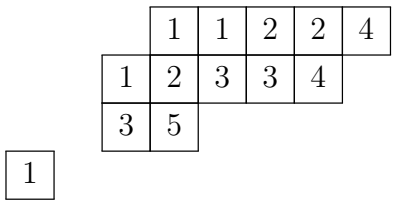
Example 3.6.11. Let's compute $s_{(2,1)}s_{(2,1)}$. So we are looking for Littlewood–Richardson tableaux of weight $(2,1)$ and shape $\lambda/(2,1)$ for some $\lambda \supseteq (2,1)$. The only such tableaux are



where we have indicated the partition $(2,2)$ in gray. Thus

$$s_{(2,1)}s_{(2,1)} = s_{(4,2)} + s_{(4,1,1)} + s_{(3,3)} + 2s_{(3,2,1)} + s_{(3,1,1,1)} + s_{(2,2,2)} + s_{(2,2,1,1)}.$$ △

Definition 3.6.8 is different from the traditional definition of Littlewood–Richardson tableaux, which involves lattice words. Nevertheless, the definitions are equivalent, as we now explain. Given a skew tableau T, we define its *word*, denoted $\mathrm{word}(T)$, to be the sequence of positive integers obtained by reading the entries of T from right to left in successive rows, starting with the top row. For example,

if $T =$  then $\mathrm{word}(T) = (4,2,2,1,1,4,3,3,2,1,5,3,1)$.

We say that a sequence of positive integers is a *lattice word* if, in every initial part of the sequence, any number i occurs at least as many times as the number $i+1$. For example

$$(1,1,2,3,2,1,3,4)$$

is a lattice word, but

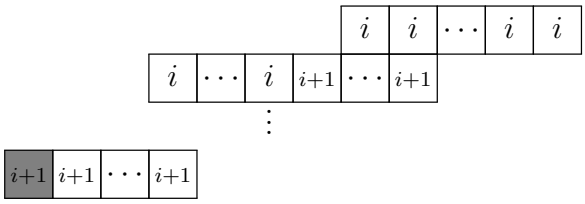
$$(1,2,2,3,1,1,3,4)$$

is not, since the initial part $(1,2,2)$ contains the number 2 more times than the number 1.

Proposition 3.6.12. *A semistandard skew tableau T is a Littlewood–Richardson tableau if and only if* word(T) *is a lattice word.*

Proof. Let T be a semistandard skew tableau. First suppose that word(T) is not a lattice word. Then there exists some i such that some initial part of word(T) contains the number $i+1$ more times than in contains the number i. Let k be the length of the smallest such initial part. It follows that the k-th entry in word(T) is $i+1$. Let j be the column number of this entry $i+1$. Then the entries i and $i+1$ in $T_{\geq j}$ are arranged as follows:

(3.77)



In particular, since T is semistandard, all of the entries $i+1$ in the length k initial part of word(T) occur in $T_{\geq j}$. It follows that $\mathrm{wt}(T_{\geq j})_{i+1} > \mathrm{wt}(T_{\geq j})_i$. Hence $\mathrm{wt}(T_{\geq j}) \notin \mathrm{Par}$, and so T is not a Littlewood–Richardson tableau.

Conversely, suppose T is not a Littlewood–Richardson tableau. Let j be the smallest integer such that $\mathrm{wt}(T_{\geq j}) \notin \mathrm{Par}$. Then there exists some i such that $\mathrm{wt}(T_{\geq j})_{i+1} > \mathrm{wt}(T_{\geq j})_i$. Then the entries i and $i+1$ in $T_{\geq j}$ are arranged as in (3.77), with all i entries preceding the gray box in word(T) lying in $T_{\geq j}$. (Any entry i in a column $< j$ and in a row above the grey box would have to have an entry $i+1$ immediately below it, contradicting our choice of j.) Hence word(T) is not a lattice word. □

It follows from the Littlewood–Richardson rule that the product of two Schur functions is a *nonnegative* integer linear combination of Schur functions. Later we will see how this fact also follows from a representation theoretic interpretation of the Littlewood–Richardson rule; see Remark 6.3.7. These applications to representation theory will also help explain the importance of the Littlewood–Richardson rule.

Exercises.

3.6.1. Prove that for a semistandard tableau T and a fixed positive integer j, the corresponding free entries in any row must lie in consecutive columns.

3.6.2. Consider the tableau

$$T = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|}\hline 1&1&1&1&2&2&2&3&3&3&4\\\hline 2&2&3&3&3&4&4&5\\\cline{1-8} 3&4&4&5&5&6\\\cline{1-6} 5&8\\\cline{1-2}\end{array}\ .$$

Compute $\beta_1(T)$, $\beta_2(T)$, $\beta_4(T)$ and $\beta_9(T)$.

3.6.3. Show that, for $\alpha \in \mathrm{WComp}_n$ and $\sigma \in \mathfrak{S}_n$, we have

$$\sigma a_\alpha = a_{\alpha^{\sigma^{-1}}}.$$

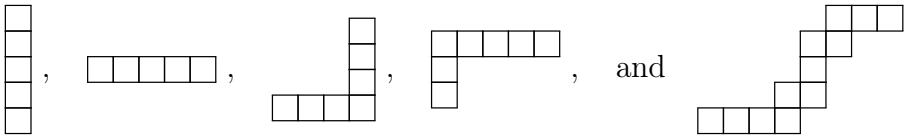
3.6.4. Deduce the dual Pieri formula (3.69) directly from the Littlewood–Richardson rule, as in Example 3.6.10.

3.6.5. Prove that $c^\lambda_{\mu\nu} = c^\lambda_{\nu\mu}$ for all $\lambda, \mu, \nu \in \mathrm{Par}$.

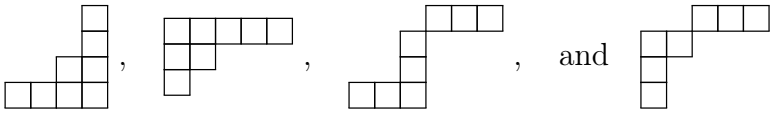
3.7 The Murnaghan–Nakayama rule

Our goal in this section is to prove the Murnaghan–Nakayama rule, which tells us how to write the product of a power sum and a Schur function as a linear combination of Schur functions. This result will be important in our proof of the boson-fermion correspondence. When we discuss the connection between symmetric functions and characters of the symmetric group, we will see that this rule also gives us a combinatorial method to compute the irreducible characters of the symmetric group.

A skew Young diagram is a *border strip* if it is connected and contains no 2×2 block of squares. Thus the successive rows (or columns) of a border strip overlap by exactly one square. For example,

(3.78) 

are border strips, but



are not, since the first two contain 2×2 boxes and the last two are not connected. The *height* $\mathrm{ht}(\theta)$ of a border strip θ is one less than the number of rows it contains. For example, the heights of the border strips in (3.78) are 4, 0, 3, 2, and 4, respectively.

Proposition 3.7.1 (Murnaghan–Nakayama rule)**.** *For all $\lambda \in \mathrm{Par}$ and $r \geq 1$, we have*

$$p_r s_\lambda = \sum_\mu (-1)^{\mathrm{ht}(\mu/\lambda)} s_\mu, \tag{3.79}$$

where the sum is over all partitions $\mu \supseteq \lambda$ such that μ/λ is a border strip with r boxes.

Proof. Fix $\lambda \in \mathrm{Par}$ and $r \geq 1$. It suffices to consider finitely many indeterminates $x_1, x_2, \ldots, x_n$, where $n \geq |\lambda| + r$. Let $\chi_i \in \mathrm{WComp}_n$ be the composition with 1 in the i-th position and 0's elsewhere. Then we have

$$\begin{aligned} p_r a_{\lambda+\delta^n} &= \sum_{\pi \in \mathfrak{S}_n} \sum_{i=1}^n \mathrm{sgn}(\pi) x^{(\lambda+\delta^n)^\pi + r\chi_i} = \sum_{\pi \in \mathfrak{S}_n} \mathrm{sgn}(\pi) \sum_{i=1}^n x^{(\lambda+\delta^n+r\chi_{\pi^{-1}(i)})^\pi} \\ &= \sum_{\pi \in \mathfrak{S}_n} \mathrm{sgn}(\pi) \sum_{i=1}^n x^{(\lambda+\delta^n+r\chi_i)^\pi} = \sum_{i=1}^n a_{\lambda+\delta^n+r\chi_i}. \end{aligned} \tag{3.80}$$

Now, we would like to rearrange the terms of $\lambda + \delta^n + r\chi_i$ in descending order. Since $\lambda \in \mathrm{Par}_n$, the terms of $\lambda + \delta^n$ are strictly decreasing. (See Lemma 3.1.5.) Thus we have

$$\lambda_1 + n - 1 > \lambda_2 + n - 2 > \cdots > \lambda_n.$$

The composition $\lambda + \delta^n + r\chi_i$ only differs from $\lambda + \delta^n$ in the i-th position. Thus, in order to arrange the terms of $\lambda + \delta^n + r\chi_i$ in descending order, we only need to move its i-th term $\lambda_i + n - i + r$ into the correct position. If $\lambda + \delta^n + r\chi_i$ has two equal terms, then the alternating polynomial $a_{\lambda+\delta^n+r\chi_i}$ in (3.80) is zero. Otherwise, there exists some $j \leq i$ such that

$$\lambda_{j-1} + n - j + 1 > \lambda_i + n - i + r > \lambda_j + n - j.$$

That is, we must move the term $\lambda_i + n - i + r$ to the j-th position in order to arrange the terms in descending order. (We adopt the convention here that $\lambda_0 = \infty$ when $j = 1$, so that the first inequality above is automatically satisfied.) Then we have

$$a_{\lambda+\delta^n+r\chi_i} = (-1)^{i-j} a_{\mu+\delta^n},$$

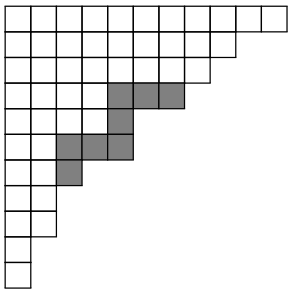
where

$$\mu = (\lambda_1, \ldots, \lambda_{j-1}, \lambda_i + j - i + r, \lambda_j + 1, \ldots, \lambda_{i-1} + 1, \lambda_{i+1}, \ldots, \lambda_n) \in \mathrm{Par}_n.$$

Note that

$$\mu_k = \lambda_{k-1} + 1 \text{ for } j < k \leq i.$$

So the Young diagrams look something like



where λ consists of the white boxes and μ consists of the white and gray boxes. In particular, we see that μ/λ is a border strip with r boxes. In the above example, $j = 4$ (the first row containing a gray box) and $i = 7$ (the last row containing a gray box). The values of i for which $\lambda + \delta^n + r\chi_i$ has two equal terms correspond precisely to those i where one cannot add to λ a border strip whose bottom row is the i-th row.

It follows from the above discussion that

$$p_r a_{\lambda+\delta^n} = \sum_{\mu} (-1)^{\mathrm{ht}(\mu/\lambda)} a_{\mu+\delta^n},$$

where the sum is over all partitions $\mu \supseteq \lambda$ such that μ/λ is a border strip with r boxes. Then the result follows from (3.7) after dividing both sides by a_{δ^n}. □

Exercises.

3.7.1. Write $p_1 s_{(5,3,2,2)}$ as a linear combination of Schur functions.

3.7.2. Write $p_3 s_{(4,4,2,1)}$ as a linear combination of Schur functions.

3.7.3. Use the Murnaghan—Nakayama rule to write the power sum p_r, $r \geq 1$, as a linear combination of Schur functions.

Chapter 4

Adjoint operators and Hopf algebras

So far we have considered Sym only as a ring. In this chapter we turn our attention to additional structure on Sym. We begin with a brief discussion of tensor products over commutative rings, since this concept is necessary for the additional structure. We then explore adjoint operators with respect the Hall inner product, which are a crucial ingredient of the Heisenberg algebra that we will discuss in the next chapter. Finally, we define Hopf algebras and define a Hopf algebra structure on Sym. In addition to being important in its own right, this structure will play an important role in our discussion of the Heisenberg algebra and applications to representation theory.

4.1 Tensor products

The material in this and following chapters will refer to tensor products of modules over a commutative ring $\Bbbk$. (In our case, $\Bbbk$ will usually be either $\mathbb{Z}$ or $\mathbb{Q}$.) In this section we give a brief treatment of such tensor products. We give their definition and some of their most important properties. We omit proofs, which can be found, for example, in [Sav, Ch. 3].

Fix a commutative ring $\Bbbk$. Let U and V be $\Bbbk$-modules. Let F be the free $\Bbbk$-module with basis $U \times V$. Let G be the submodule of F generated by the elements

$$\begin{aligned} (u+u',v)-(u,v)-(u',v), \qquad & (u,v+v')-(u,v)-(u,v'), \\ (cu,v)-c(u,v), \qquad & (u,cv)-c(u,v), \end{aligned}$$

for $u, u' \in U$, $v, v' \in V$, $c \in \Bbbk$. The *tensor product* of U and V over $\Bbbk$ is

$$U \otimes_{\Bbbk} V := F/G.$$

If the ground ring $\Bbbk$ is clear from the context, we sometimes write $U \otimes V$ for $U \otimes_{\Bbbk} V$. We set

$$u \otimes v := (u,v) + G, \quad u \in U,\ v \in V.$$

Therefore, every element of $U \otimes_{\Bbbk} V$ can be written in the form

$$t = \sum_{i=1}^{n} u_i \otimes v_i, \quad n \in \mathbb{N},\ u_1, \ldots, u_n \in U,\ v_1, \ldots, v_n \in V.$$

Keep in mind that such a representation is not unique in general.

Warning 4.1.1. It is *not* true that every element of $U \otimes V$ is of the form $u \otimes v$ for some $u \in U$, $v \in V$. An element of the form $u \otimes v$ is called a *simple tensor*. Every element can be written as a *finite sum* of simple tensors. Therefore, when defining a $\Bbbk$-linear map whose domain is $U \otimes V$, we will often define it on simple tensors. There is then at most one way to extend it to all of $U \otimes V$.

It follows from the definition that $u \otimes v$ is linear in each entry:

$$u \otimes (v + v') = u \otimes v + u \otimes v', \quad (u + u') \otimes v = u \otimes v + u' \otimes v, \tag{4.1}$$

$$c(u \otimes v) = cu \otimes v = u \otimes cv. \tag{4.2}$$

Thus,

$$h \colon U \times V \to U \otimes_R V, \quad h(m, n) = m \otimes n, \tag{4.3}$$

is an $\Bbbk$-bilinear map. The tensor product is uniquely characterized by the following universal property.

Proposition 4.1.2 (Universal property of the tensor product)**.** *For every $\Bbbk$-module W and bilinear map $\varphi \colon U \times V \to W$, there exists a unique $\Bbbk$-module homomorphism $\tilde{\varphi} \colon T \to W$ such that $\tilde{\varphi} \circ h = \varphi$; that is, the diagram*

$$\begin{array}{ccc} U \times V & \xrightarrow{h} & U \otimes_{\Bbbk} V \\ & \searrow^{\varphi} & \downarrow{\scriptstyle \tilde{\varphi}} \\ & & W \end{array}$$

commutes.

Proof. For a proof, see [Sav, Prop. 3.2.2]. □

Defining a map on a tensor product can be subtle, since elements of a tensor product are equivalence classes. See Exercise 4.1.2. Typically the best way to define a map whose domain is a tensor product $U \otimes_{\Bbbk} V$ is to use Proposition 4.1.2. More precisely, you define the map on $U \times V$, show that the map is bilinear, and then use Proposition 4.1.2 to conclude that the map factors through $U \otimes_{\Bbbk} V$.

Example 4.1.3. Fix $m, n \in \mathbb{Z}_{>0}$. Let's prove that we have an isomorphism of $\mathbb{Z}$-modules

$$(\mathbb{Z}/m\mathbb{Z}) \otimes_{\mathbb{Z}} (\mathbb{Z}/n\mathbb{Z}) \cong \mathbb{Z}/\gcd(m, n)\mathbb{Z}.$$

For $x \in \mathbb{Z}$, let $[x]_k$ denote its image in $\mathbb{Z}/k\mathbb{Z}$. Let $d = \gcd(m, n)$ and consider the map

$$\varphi \colon (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z}) \to \mathbb{Z}/d\mathbb{Z}, \qquad \varphi([a]_m, [b]_n) = [ab]_d.$$

We leave it as Exercise 4.1.1 to prove that this map is well defined (that is, independent of the choice of representatives a and b). Since it is clearly $\mathbb{Z}$-bilinear, it induces a map

$$\tilde{\varphi} \colon (\mathbb{Z}/m\mathbb{Z}) \otimes_{\mathbb{Z}} (\mathbb{Z}/n\mathbb{Z}) \to \mathbb{Z}/\gcd(m, n)\mathbb{Z}, \qquad \tilde{\varphi}([a]_m \otimes [b]_n) = [ab]_d.$$

Since

$$\tilde{\varphi}([a]_m \otimes [1]_n) = [a]_d,$$

for all $a \in \mathbb{Z}$, the map $\tilde{\varphi}$ is surjective. It remains to show that it is injective. Suppose

$$\tilde{\varphi}\left(\sum_{i=1}^{k} [a_i]_m \otimes [b_i]_n\right) = \sum_{i=1}^{k} [a_i b_i]_d = 0.$$

Then $\sum_{i=1}^{k} a_i b_i \in d\mathbb{Z}$ and so $\sum_{i=1}^{k} a_i b_i = dz$ for some $z \in \mathbb{Z}$. Choose $x, y \in \mathbb{Z}$ such that $d = xm + yn$. (Here we use the fact $m\mathbb{Z} + n\mathbb{Z} = \gcd(m,n)\mathbb{Z}$.) Then we have

$$\begin{aligned}
\sum_{i=1}^{k} ([a_i]_m \otimes [b_i]_n) &= \sum_{i=1}^{k} ([a_i]_m \otimes b_i[1]_n) \\
&= \sum_{i=1}^{k} (b_i[a_i]_m \otimes [1]_n) \\
&= \sum_{i=1}^{k} ([a_i b_i]_m \otimes [1]_n) \\
&= [dz]_m \otimes [1]_n \\
&= d[z]_m \otimes [1]_n \\
&= xm[z]_m \otimes [1]_n + yn[z]_m \otimes [1]_n \\
&= [0]_m \otimes [1]_n + [yz]_m \otimes n[1]_n \\
&= [0]_d + [yz]_m \otimes [0]_n \\
&= [0]_d.
\end{aligned}$$

Thus $\tilde{\varphi}$ is injective. △

The following two propositions summarize the most important properties of the tensor product.

Proposition 4.1.4 (Properties of the tensor product I)**.** *Suppose U, V, and W are $\Bbbk$-modules. In the following statements, all isomorphisms are isomorphisms of $\Bbbk$-modules and all tensor products are over $\Bbbk$.*

(a) *We have an isomorphism*

$$U \otimes V \cong V \otimes U, \qquad u \otimes v \mapsto v \otimes u.$$

(b) *We have an isomorphism*

$$(U \otimes V) \otimes W \cong U \otimes (V \otimes W), \qquad (u \otimes v) \otimes w \mapsto u \otimes (v \otimes w).$$

(c) *We have an isomorphism*

$$\begin{aligned}
&U \otimes (V \oplus W) \cong (U \otimes V) \oplus (U \otimes W), \\
&\quad u \otimes (v + w) \mapsto u \otimes v + u \otimes w, \quad u \in U,\ v \in V,\ w \in W.
\end{aligned}$$

(d) *Viewing* $\Bbbk$ *as a module over itself via multiplication, we have an isomorphism*

$$\Bbbk \otimes U \cong U, \qquad c \otimes u \mapsto cu. \tag{4.4}$$

Proof. Proofs of these statements can be found in [Sav, §3.3]. □

In light of Proposition 4.1.4(b), we will often omit the parentheses in multiple tensor products. For example, we will write

$$U \otimes V \otimes W,$$

and elements of this triple tensor product will be written as finite linear combinations of simple tensors $u \otimes v \otimes w$, for $u \in U$, $v \in V$, $w \in W$. Similarly, we will often *identify* $\Bbbk \otimes U$, $U \otimes \Bbbk$, and U.

Proposition 4.1.5 (Properties of the tensor product II)**.** *Suppose U and V are $\Bbbk$-modules. In the following statements, all isomorphisms are isomorphisms of $\Bbbk$-modules and all tensor products are over $\Bbbk$.*

(a) *If U and V are generated, as $\Bbbk$-modules, by the sets X and Y respectively, then $U \otimes V$ is generated by $\{u \otimes v : u \in X,\ v \in Y\}$.*

(b) *If U and V are free modules with bases X and Y respectively, then $U \otimes V$ is a free module with basis $\{u \otimes v : u \in X,\ v \in Y\}$.*

(c) *If $\Bbbk$ is a field and U and V are finite-dimensional vector spaces over $\Bbbk$, then $U \otimes V$ is a finite-dimensional vector space over $\Bbbk$, and*

$$\dim(U \otimes V) = (\dim U)(\dim V).$$

(d) *We have an isomorphism $\Bbbk^m \otimes \Bbbk^n \cong \Bbbk^{mn}$.*

Proof. Proofs of these statements can be found in [Sav, §3.3]. □

Example 4.1.6. Our main examples will be the $\mathbb{Z}$-algebra Sym and the $\mathbb{Q}$-algebra $\mathrm{Sym}_{\mathbb{Q}}$. It follows from Proposition 4.1.5(b) that

$$u_\lambda \otimes v_\mu, \qquad \lambda, \mu \in \mathrm{Par},$$

is a basis of $\mathrm{Sym} \otimes_{\mathbb{Z}} \mathrm{Sym}$, where $(u_\lambda)_{\lambda \in \mathrm{Par}}$ and $(v_\lambda)_{\lambda \in \mathrm{Par}}$ are any bases of Sym. For example $e_\lambda \otimes h_\mu$, $\lambda, \mu \in \mathrm{Par}$, is a basis of $\mathrm{Sym} \otimes \mathrm{Sym}$. If we instead work over $\mathbb{Q}$, then we may also use the $\mathbb{Q}$-basis $(p_\lambda)_{\lambda \in \mathrm{Par}}$ consisting of power sum symmetric functions. △

We can also take tensor products of $\Bbbk$-module homomorphisms. Suppose

$$f_1 \colon U_1 \to V_1 \qquad \text{and} \qquad f_2 \colon U_2 \to V_2$$

are $\Bbbk$-module homomorphisms. Then we have a $\Bbbk$-module homomorphism

$$f_1 \otimes f_2 \colon U_1 \otimes U_2 \to V_1 \otimes V_2, \qquad (f_1 \otimes f_2)(u_1 \otimes u_2) := f_1(u_1) \otimes f_2(u_2). \tag{4.5}$$

(We then extend to finite sums of simple tensors by linearity.) One can check that the map $f_1 \otimes f_2$ is a well-defined $\Bbbk$-module homomorphism; see [Sav, §3.4].

If A and B are $\Bbbk$-algebras, then their tensor product $A \otimes B$ is naturally a $\Bbbk$-algebra with multiplication given by

$$(a_1 \otimes b_1)(a_2 \otimes b_2) = a_1 a_2 \otimes b_1 b_2, \qquad a_1, a_2 \in A, b_1, b_2 \in B.$$

For example, $\mathrm{Sym} \otimes \mathrm{Sym}$ is a $\mathbb{Z}$-algebra. We can define an inner product on this algebra by

$$\langle u_1 \otimes v_1, u_2 \otimes v_2 \rangle := \langle u_1, u_2 \rangle \langle v_1, v_2 \rangle, \tag{4.6}$$

where the angled brackets on the right-hand side denote the Hall inner product.

Exercises.

4.1.1. Prove that the map φ of Example 4.1.3 is well defined.

4.1.2. Suppose $\Bbbk$ is a field and U and V are nonzero finite-dimensional vector spaces. Is the map

$$\varphi \colon U \otimes V \to U, \qquad \varphi(u \otimes v) = u,$$

well defined?

4.1.3. Let $n \in \mathbb{Z}$, $n \geq 1$. Prove that $\mathbb{Q} \otimes_{\mathbb{Z}} (\mathbb{Z}/n\mathbb{Z})$ is the zero $\mathbb{Z}$-module.

4.1.4. Suppose V and W are finite-dimensional vector spaces over a field $\Bbbk$. Let $V^* = \mathrm{Hom}_{\Bbbk}(V, \Bbbk)$ be the dual space of V. Show that the map

$$\psi \colon W \otimes_{\Bbbk} V^* \to \mathrm{Hom}_{\Bbbk}(V, W),$$
$$\psi(w \otimes f)(v) = f(v)w, \quad w \in W,\ f \in V^*,\ v \in V,$$

is an isomorphism of vector spaces.

4.2 Adjoint operators

In this section we discuss the adjoint operators, with respect to the Hall inner product, of multiplication by symmetric functions. These adjoint operators will form "half" of the Heisenberg algebra that we will introduce in Section 5.1.

Every ring acts on itself by multiplication. In particular we have an injective ring homomorphism

$$\mathrm{Sym} \to \mathrm{End}(\mathrm{Sym}), \quad f \mapsto (v \mapsto fv), \tag{4.7}$$

sending $f \in \mathrm{Sym}$ to the operation of multiplication by f. (See Exercise 4.2.1.)

Recall the Hall inner product $\langle\ ,\ \rangle$ from Definition 3.3.1. We deduced several important properties of this inner product in Theorem 3.3.7. In particular, $\langle\ ,\ \rangle$ is symmetric and positive definite. This allows us to define the adjoint of any linear operator on Sym. In particular, we have the adjoint $f^\perp$ of multiplication by any symmetric function f, as described in the following proposition.

Proposition 4.2.1. *For each $f \in \mathrm{Sym}$ there exists a unique linear operator $f^\perp \in \mathrm{End}(\mathrm{Sym})$ satisfying*

$$\langle f^\perp u, v\rangle = \langle u, fv\rangle \qquad \text{for all } u, v \in \mathrm{Sym}. \tag{4.8}$$

The map

$$\mathrm{Sym} \to \mathrm{End}(\mathrm{Sym}), \quad f \mapsto f^\perp,$$

is an injective ring homomorphism.

Proof. Let $f \in \mathrm{Sym}$. We first prove existence of a linear operator $f^\perp$ satisfying (4.8). Define $f^\perp \in \mathrm{End}(\mathrm{Sym})$ by

$$f^\perp u := \sum_{\lambda \in \mathrm{Par}} \langle u, fs_\lambda\rangle s_\lambda, \qquad u \in \mathrm{Sym}. \tag{4.9}$$

The fact that $f^\perp$ is a linear map follows easily from the fact that $\langle\ ,\ \rangle$ is bilinear. (Exercise 4.2.2.) For all $v \in \mathrm{Sym}$ we have

$$\langle f^\perp u, v\rangle = \sum_{\lambda \in \mathrm{Par}} \langle u, fs_\lambda\rangle\langle s_\lambda, v\rangle = \left\langle u, f \sum_{\lambda \in \mathrm{Par}} \langle s_\lambda, v\rangle s_\lambda \right\rangle = \langle u, fv\rangle.$$

Thus $f^\perp$ satisfies (4.8).

Next we show uniqueness. Suppose that $T \in \mathrm{End}(\mathrm{Sym})$ satisfies

$$\langle Tu, v\rangle = \langle u, fv\rangle \qquad \text{for all } u, v \in \mathrm{Sym}.$$

Then, for all $u, v \in \mathrm{Sym}$, we have

$$\left\langle \left(T - f^\perp\right) u, v\right\rangle = \langle Tu, v\rangle - \langle f^\perp u, v\rangle = \langle u, fv\rangle - \langle u, fv\rangle = 0.$$

Thus $(T - f^\perp)(u) = 0$ for all u, and hence $T = f^\perp$.

Next we show that the map $f \mapsto f^\perp$ is injective. Suppose that $f \in \mathrm{Sym}$ and $f^\perp = 0$. Then

$$0 = \langle f^\perp f, 1\rangle = \langle f, f\rangle.$$

Since $\langle\ ,\ \rangle$ is positive definite, this implies that $f = 0$.

It remains to show that the map $f \mapsto f^\perp$ is a ring homomorphism. For all $u, v, f, g \in \mathrm{Sym}$, we have

$$\langle (f+g)^\perp u, v\rangle = \langle u, (f+g)v\rangle = \langle u, fv\rangle + \langle u, gv\rangle = \langle f^\perp u, v\rangle + \langle g^\perp u, v\rangle = \langle (f^\perp + g^\perp)u, v\rangle.$$

It then follows from the uniqueness proved above that $(f+g)^\perp = f^\perp + g^\perp$. Similarly,

$$\langle (fg)^\perp u, v\rangle = \langle u, fgv\rangle = \langle u, gfv\rangle = \langle g^\perp u, fv\rangle = \langle f^\perp g^\perp u, v\rangle,$$

and so $(fg)^\perp = f^\perp g^\perp$. Finally,

$$\langle 1^\perp u, v\rangle = \langle u, 1v\rangle = \langle 1u, v\rangle,$$

and so $1^\perp = 1$. □

Lemma 4.2.2. *For $f \in \mathrm{Sym}^n$, we have*

$$f^\perp(\mathrm{Sym}^m) \subseteq \mathrm{Sym}^{m-n}.$$

(We say that $f^\perp$ is an operator of degree $-n$.)

Proof. We leave the proof of this result as Exercise 4.2.3. □

We can reinterpret the skew Schur functions in terms of adjoint operators.

Lemma 4.2.3. *For $\lambda, \mu \in \mathrm{Par}$, we have*

$$s_{\lambda/\mu} = s_\mu^\perp s_\lambda. \tag{4.10}$$

Proof. We have

$$s_{\lambda/\mu} \overset{(3.34)}{=} \sum_{\nu\in\mathrm{Par}} \langle s_\lambda, s_\mu s_\nu\rangle s_\nu = \sum_{\nu\in\mathrm{Par}} \langle s_\mu^\perp s_\lambda, s_\nu\rangle s_\nu = s_\mu^\perp s_\lambda.$$ □

Let $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ be two sets of indeterminates.

Lemma 4.2.4. *For all $f \in \mathrm{Sym}$, we have*

$$f(x,y) = \sum_{\mu\in\mathrm{Par}} (s_\mu^\perp f)(x) s_\mu(y). \tag{4.11}$$

Proof. First note that

$$s_\lambda(x,y) \overset{(3.46)}{=} \sum_{\mu\in\mathrm{Par}} s_{\lambda/\mu}(x) s_\mu(y) = \sum_{\mu\in\mathrm{Par}} (s_\mu^\perp s_\lambda)(x) s_\mu(y).$$

Thus, (4.11) holds for $f = s_\lambda$, $\lambda \in \mathrm{Par}$. Since both sides of (4.11) are linear in f, and the Schur functions form a basis of Sym, it follows that (4.11) holds for all $f \in \mathrm{Sym}$. □

Lemma 4.2.5. *For $\lambda \in \mathrm{Par}$ and $n \geq 1$, we have*

$$h_\lambda^\perp m_\mu = \begin{cases} m_\nu & \text{if } \mu = \lambda \cup \nu, \\ 0 & \text{if } \mu \text{ is not of the form } \lambda \cup \nu \text{ for some } \nu \in \mathrm{Par}. \end{cases}$$

Proof. Since the basis $(h_\lambda)_{\lambda\in\mathrm{Par}}$ is dual to the basis $(m_\lambda)_{\lambda\in\mathrm{Par}}$ by (3.18), the coefficient of m_ν when we write $h_\lambda^\perp m_\mu$ in the basis of monomial symmetric functions is $\langle h_\lambda^\perp m_\mu, h_\nu\rangle$. Then the result follows from the computation

$$\langle h_\lambda^\perp m_\mu, h_\nu\rangle = \langle m_\mu, h_\lambda h_\nu\rangle = \langle m_\mu, h_{\lambda\cup\nu}\rangle \overset{(3.18)}{=} \delta_{\mu,\lambda\cup\nu}. \qquad \square$$

We now turn our attention to the operators $p_n^\perp$. Since the power sum symmetric functions are only a basis over $\mathbb{Q}$, we work over $\mathbb{Q}$ for the remainder of the section. Recall from Theorems 2.4.6, 2.5.3, and 2.6.4 that

$$\mathrm{Sym}_{\mathbb{Q}} = \mathbb{Q}[e_1, e_2, \dots] = \mathbb{Q}[h_1, h_2, \dots] = \mathbb{Q}[p_1, p_2, \dots].$$

Thus, for $n \geq 1$, we have linear operators

$$\frac{\partial}{\partial e_n}, \quad \frac{\partial}{\partial h_n}, \quad \frac{\partial}{\partial p_n} \quad \in \mathrm{End}(\mathrm{Sym}_{\mathbb{Q}}). \tag{4.12}$$

The operator $\frac{\partial}{\partial e_n}$ acts on an element $f \in \mathrm{Sym}_{\mathbb{Q}}$ by acting on its expression as a polynomial in the elementary symmetric functions, and similarly for $\frac{\partial}{\partial h_n}$ and $\frac{\partial}{\partial p_n}$.

Example 4.2.6. Using the transition matrix $M(e,h) = K^\intercal J K^{-\intercal}$ from Theorem 3.5.5, we can compute

$$e_3 = h_{(3)} - 2h_{(2,1)} + h_{(1,1,1)} = h_3 - 2h_2h_1 + h_1^3.$$

Thus

$$\frac{\partial e_3}{\partial h_1} = \frac{\partial}{\partial h_1}\left(h_3 - 2h_2h_1 + h_1^3\right) = -2h_2 + 3h_1^2. \qquad \triangle$$

Proposition 4.2.7. *For $n \geq 1$, we have*

$$p_n^\perp = n\frac{\partial}{\partial p_n}. \tag{4.13}$$

Proof. Since the p_λ, $\lambda \in \mathrm{Par}$, form a basis of $\mathrm{Sym}_{\mathbb{Q}}$, it suffices to prove that

$$\langle p_n^\perp p_\lambda, p_\mu\rangle = n\left\langle \frac{\partial p_\lambda}{\partial p_n}, p_\mu \right\rangle \qquad \text{for all } \lambda, \mu \in \mathrm{Par}. \tag{4.14}$$

We have

$$\langle p_n^\perp p_\lambda, p_\mu\rangle = \langle p_\lambda, p_n p_\mu\rangle = \langle p_\lambda, p_{\mu\cup(n)}\rangle \overset{(3.29)}{=} \begin{cases} z_\lambda & \text{if } \lambda = \mu\cup(n), \\ 0 & \text{otherwise.}\end{cases}$$

On the other hand

$$\frac{\partial p_\lambda}{\partial p_n} = \begin{cases} m_n(\lambda)p_\mu & \text{if } \lambda = \mu\cup(n) \text{ for some } \mu, \\ 0 & \text{if } m_n(\lambda) = 0.\end{cases}$$

Thus, the equality (4.14) holds when λ has no part equal to n. On the other hand, if $\lambda = \mu\cup(n)$, we have

$$\left\langle \frac{\partial p_\lambda}{\partial p_n}, p_\mu \right\rangle = m_n(\lambda)\langle p_\mu, p_\mu\rangle \overset{(3.29)}{=} m_n(\lambda)z_\mu.$$

Since

$$z_\lambda z_\mu^{-1} = \left(\prod_{i\geq 1} i^{m_i(\lambda)} m_i(\lambda)!\right)\left(\prod_{i\geq 1} i^{m_i(\mu)} m_i(\mu)!\right)^{-1} = n m_n(\lambda),$$

we have

$$\langle p_n^\perp p_\lambda, p_\mu\rangle = z_\lambda = n m_n(\lambda) z_\mu = n\left\langle \frac{\partial p_\lambda}{\partial p_n}, p_\mu\right\rangle,$$

as desired. □

If $\Bbbk$ is a commutative ring and A is a $\Bbbk$-algebra, a $\Bbbk$-*derivation* (or simply a *derivation*, if the ground ring $\Bbbk$ is clear from the context) of A is a $\Bbbk$-linear map $D\colon A \to A$ satisfying the Leibniz rule:

$$D(ab) = D(a)b + aD(b), \qquad a, b \in A.$$

A derivation is uniquely determined by its action on a set of generators, since one can inductively use the Leibniz rule to compute its value on any product of the generators.

Corollary 4.2.8. *The operator $p_n^\perp$ is a $\mathbb{Q}$-derivation of* $\mathrm{Sym}_\mathbb{Q}$.

Proof. This follows from (4.13) and the fact that the partial derivative $\frac{\partial}{\partial p_n}$ satisfies the Leibniz rule. □

Proposition 4.2.9. *For $n \geq 1$, we have*

$$p_n^\perp = \sum_{r=0}^{\infty} h_r \frac{\partial}{\partial h_{n+r}} = (-1)^{n-1}\sum_{r=0}^{\infty} e_r \frac{\partial}{\partial e_{n+r}}. \tag{4.15}$$

Proof. Since $p_n^\perp$ is a derivation by Corollary 4.2.8, it suffices to check the given equalities on a set of generators. For the first equality, we check on the generators h_m, $m \geq 1$. First note that, for all $\lambda \in \mathrm{Par}(n)$, we have

$$\langle h_n, p_\lambda\rangle \overset{(2.49)}{=} \sum_{\mu\in\mathrm{Par}(n)} z_\mu^{-1}\langle p_\mu, p_\lambda\rangle \overset{(3.29)}{=} 1.$$

Thus, for all $m \geq n \geq 1$, $\lambda \in \mathrm{Par}(m-n)$, we have

$$\langle p_n^\perp h_m, p_\lambda\rangle = \langle h_m, p_n p_\lambda\rangle = 1 = \langle h_{m-n}, p_\lambda\rangle.$$

Since $p_n^\perp h_m \in \mathrm{Sym}^{m-n}$ by Lemma 4.2.2, and the p_λ, $\lambda \in \mathrm{Par}(m-n)$, form a basis of Sym^{m-n}, it follows that

$$p_n^\perp h_m = h_{m-n} = \sum_{r=0}^{\infty} h_r \frac{\partial h_m}{\partial h_{n+r}}.$$

This proves the first equality in (4.15).

Next, for $\lambda \in \mathrm{Par}(m-n)$, we have

$$\begin{aligned}\langle p_n^\perp e_m, p_\lambda\rangle &= \langle e_m, p_n p_\lambda\rangle \\ &\overset{(3.31)}{=} \langle \omega(e_m), \omega(p_n p_\lambda)\rangle\end{aligned}$$

$$\begin{aligned}
&\overset{(2.31)}{\underset{(2.45)}{=}} (-1)^{n-1}\langle h_m, p_n\omega(p_\lambda)\rangle \\
&= (-1)^{n-1}\langle p_n^\perp h_m, \omega(p_\lambda)\rangle \\
&= (-1)^{n-1}\langle h_{m-n}, \omega(p_\lambda)\rangle \\
&\overset{(3.31)}{=} (-1)^{n-1}\langle \omega(h_{m-n}), p_\lambda\rangle \\
&= (-1)^{n-1}\langle e_{m-n}, p_\lambda\rangle.
\end{aligned}$$

Since $p_n^\perp e_m \in \mathrm{Sym}^{m-n}$ by Lemma 4.2.2, and the p_λ, $\lambda \in \mathrm{Par}(m-n)$, form a basis of Sym^{m-n}, it follows that

$$p_n^\perp e_m = (-1)^{n-1} e_{m-n} = (-1)^{n-1} \sum_{r=0}^{\infty} e_r \frac{\partial e_m}{\partial e_{n+r}}.$$

This proves the second equality in (4.15). □

Remark 4.2.10. Note that even though the sums in (4.15) are infinite, all but finitely many terms act as zero on any given element of Sym. Thus, these infinite sums are well-defined operators on Sym. △

The involution ω of (2.31) is compatible with adjoints in the following sense.

Proposition 4.2.11. *For all* $f, g \in \mathrm{Sym}$,

$$\omega(f^\perp g) = \omega(f)^\perp \omega(g). \tag{4.16}$$

Proof. For all $f, g, h \in \mathrm{Sym}$, we have

$$\left\langle \omega(f^\perp g), h\right\rangle = \left\langle f^\perp g, \omega(h)\right\rangle = \left\langle g, f\omega(h)\right\rangle \overset{(3.31)}{=} \left\langle \omega(g), \omega(f)h\right\rangle = \left\langle \omega(f)^\perp \omega(g), h\right\rangle.$$

Since this holds for all $h \in \mathrm{Sym}$, the result follows. □

Exercises.

4.2.1. Show that (4.7) is an injective ring homomorphism.

4.2.2. Show that $f^\perp$, as defined by (4.9), is a linear map $\mathrm{Sym} \to \mathrm{Sym}$.

4.2.3. Prove Lemma 4.2.2.

4.2.4. Suppose that $\Bbbk$ is a commutative ring, A is a $\Bbbk$-algebra, and $D\colon A \to A$ is a $\Bbbk$-derivation. Prove that $D(1_A) = 0$.

4.2.5. Show that, for $f, g \in \mathrm{Sym}$, we have $(f \otimes g)^\perp = f^\perp \otimes g^\perp$. Here $(f \otimes g)^\perp$ denotes the operator $\mathrm{Sym} \otimes \mathrm{Sym} \to \mathrm{Sym} \otimes \mathrm{Sym}$ adjoint, with respect to the inner product (4.6), to multiplication by $f \otimes g$.

4.2.6. Prove that

$$\{f^\perp : f \in \mathrm{Sym}_{\mathbb{Q}}\} = \mathrm{Span}_{\mathbb{Q}}\left\{\frac{\partial^k}{\partial p_{i_1}\partial p_{i_2}\cdots\partial p_{i_k}} : k \in \mathbb{N},\ i_1,\ldots,i_k \geq 1\right\}.$$

Here we view both sides as subspaces of $\mathrm{End}_{\mathbb{Q}}(\mathrm{Sym})$.

4.3 Hopf algebras

In Section 4.4 will define the structure of a *Hopf algebra* on the ring of symmetric functions. In preparation for this, we give in this section the definition of a Hopf algebra. For a more detailed treatment, see, for example, [DNR01].

Throughout this section we fix a commutative ring $\Bbbk$. (In most of our applications, we will take $\Bbbk = \mathbb{Z}$ or $\Bbbk = \mathbb{Q}$.) All tensor products are over $\Bbbk$. If U and V are two $\Bbbk$-modules, we let

$$\mathrm{flip}\colon U \otimes V \xrightarrow{\cong} V \otimes U, \qquad \mathrm{flip}(u \otimes v) = v \otimes u,$$

be the isomorphism of Proposition 4.1.4(a).

Definition 4.3.1 (Algebra)**.** A *(unital associative) algebra* over $\Bbbk$ is an $\Bbbk$-module A together with $\Bbbk$-linear maps

$$\nabla\colon A \otimes A \to A \qquad \text{and} \qquad \eta\colon \Bbbk \to A,$$

such that the following two diagrams commute:

(4.17)
$$\begin{array}{ccc} A\otimes A\otimes A & \xrightarrow{\nabla\otimes\mathrm{id}} & A\otimes A \\ {\scriptstyle \mathrm{id}\otimes\nabla}\downarrow & & \downarrow{\scriptstyle \nabla} \\ A\otimes A & \xrightarrow[\nabla]{} & A \end{array} \qquad \begin{array}{ccc} A & \xrightarrow{\eta\otimes\mathrm{id}} & A\otimes A \\ {\scriptstyle \mathrm{id}\otimes\eta}\downarrow & \searrow^{\mathrm{id}} & \downarrow{\scriptstyle \nabla} \\ A\otimes A & \xrightarrow[\nabla]{} & A \end{array}$$

The map ∇ is called the *multiplication* (or *product*) of A and η is the *unit* of A. We often refer to A itself as being an algebra, leaving the maps ∇ and ε implied. We say that A is *commutative* if $\nabla \circ \mathrm{flip} = \nabla$.

Note that, in (4.17), when we write $\eta \otimes \mathrm{id}\colon A \to A \otimes A$, we really mean the composite

$$A \xrightarrow[(4.4)]{\cong} \Bbbk \otimes A \xrightarrow{\eta\otimes A} A \otimes A,$$

and similarly for $\mathrm{id} \otimes \eta\colon A \to A \otimes A$. When comparing Definition 4.3.1 to the definition of $\Bbbk$-algebra given in Section 1.3, the unit η corresponds to the map

$$\eta\colon \Bbbk \to A, \quad \eta(c) = c1_A,$$

where 1_A is the multiplicative identity of A. It is then straightforward to see that, if A_1 and A_2 are algebras, a $\Bbbk$-linear map $\varphi\colon A_1 \to A_2$ is a morphism of algebras (also called a

homomorphism of algebras) in the usual sense if and only if the following diagrams commute:

$$\begin{array}{ccc} A_1 \otimes A_1 & \xrightarrow{\nabla_1} & A_1 \\ {\scriptstyle \varphi\otimes\varphi}\downarrow & & \downarrow{\scriptstyle \varphi} \\ A_2 \otimes A_2 & \xrightarrow{\nabla_2} & A_2 \end{array} \qquad \begin{array}{ccc} & \Bbbk & \\ {\scriptstyle \eta_1}\swarrow & & \searrow{\scriptstyle \eta_2} \\ A_1 & \xrightarrow{\varphi} & A_2 \end{array}$$

Here ∇_i and η_i denote the multiplication and unit of A_i, respectively, for $i = 1, 2$.

Reversing all the arrows in the diagrams (4.17) gives the following definition of a coalgebra.

Definition 4.3.2 (Coalgebra)**.** A *(counital coassociative) coalgebra* over $\Bbbk$ is a $\Bbbk$-module C together with $\Bbbk$-linear maps

$$\Delta\colon C \to C \otimes C \qquad \text{and} \qquad \varepsilon\colon C \to \Bbbk,$$

such that the following two diagrams commute:

$$\begin{array}{ccc} C & \xrightarrow{\Delta} & C \otimes C \\ {\scriptstyle \Delta}\downarrow & & \downarrow{\scriptstyle \mathrm{id}\otimes\Delta} \\ C \otimes C & \xrightarrow[\Delta\otimes\mathrm{id}]{} & C \otimes C \otimes C \end{array} \qquad \begin{array}{ccc} C & \xrightarrow{\Delta} & C \otimes C \\ {\scriptstyle \Delta}\downarrow & \searrow^{\mathrm{id}} & \downarrow{\scriptstyle \mathrm{id}\otimes\varepsilon} \\ C \otimes C & \xrightarrow[\varepsilon\otimes\mathrm{id}]{} & C \end{array} \tag{4.18}$$

The map Δ is called the *comultiplication* (or *coproduct*) of C and ε is the *counit* of C. We often refer to C itself as a coalgebra, leaving the maps Δ and ε implied. We say that C is *cocommutative* if $\mathrm{flip} \circ \Delta = \Delta$.

Note that, in (4.18), when we write $\varepsilon \otimes \mathrm{id}\colon C \otimes C \to C$, we really mean the composite

$$C \otimes C \xrightarrow{\varepsilon\otimes\mathrm{id}} \Bbbk \otimes C \xrightarrow[(4.4)]{\cong} C,$$

and similarly for $\mathrm{id} \otimes \varepsilon\colon C \otimes C \to C$.

It is useful to define a consistent notation for writing out the comultiplication explicitly. If C is a coalgebra and $c \in C$, we write

$$\Delta(c) = \sum_{(c)} c_{(1)} \otimes c_{(2)}. \tag{4.19}$$

This is known as *Sweedler notation*.

Definition 4.3.3 (Coalgebra morphism)**.** Suppose $(C_1, \Delta_1, \varepsilon_1)$ and $(C_2, \Delta_2, \varepsilon_2)$ are coalgebras. A $\Bbbk$-linear map $\varphi\colon C_1 \to C_2$ is a *morphism of coalgebras* if the following diagrams commute:

$$\begin{array}{ccc} C_1 \otimes C_1 & \xleftarrow{\Delta_1} & C_1 \\ {\scriptstyle \varphi\otimes\varphi}\downarrow & & \downarrow{\scriptstyle \varphi} \\ C_2 \otimes C_2 & \xleftarrow{\Delta_2} & C_2 \end{array} \qquad \begin{array}{ccc} & \Bbbk & \\ {\scriptstyle \varepsilon_1}\nearrow & & \nwarrow{\scriptstyle \varepsilon_2} \\ C_1 & \xrightarrow{\varphi} & C_2 \end{array}$$

Definition 4.3.4 (Bialgebra)**.** A *bialgebra* over $\Bbbk$ is a tuple $(B, \nabla, \eta, \Delta, \varepsilon)$ such that (B, ∇, η) is an algebra, (B, Δ, ε) is a coalgebra, and the following diagrams commute

(4.20)
$$\begin{array}{ccccc} B \otimes B & & \xrightarrow{\nabla} & & B \xrightarrow{\Delta} B \otimes B \\ {\scriptstyle \Delta\otimes\Delta}\downarrow & & & & \uparrow{\scriptstyle \nabla\otimes\nabla} \\ B \otimes B \otimes B \otimes B & & \xrightarrow{\mathrm{id}\otimes\mathrm{flip}\otimes\mathrm{id}} & & B \otimes B \otimes B \otimes B \end{array}$$

(4.21)
$$\begin{array}{ccc} B \otimes B & \xrightarrow{\nabla} & B \\ & {\scriptstyle \varepsilon\otimes\varepsilon}\searrow \quad \swarrow{\scriptstyle \varepsilon} & \\ & \Bbbk & \end{array} \qquad \begin{array}{ccc} & \Bbbk & \\ {\scriptstyle \eta}\swarrow & & \searrow{\scriptstyle \eta\otimes\eta} \\ B & \xrightarrow{\Delta} & B \otimes B \end{array} \qquad \begin{array}{ccc} & B & \\ {\scriptstyle \eta}\nearrow & & \searrow{\scriptstyle \varepsilon} \\ \Bbbk & \xrightarrow{\mathrm{id}} & \Bbbk \end{array}$$

We often refer to B itself as a bialgebra, leaving the maps ∇, η, Δ, and ε implied. A *morphism of bialgebras* is a map that is both a morphism of algebras and a morphism of coalgebras.

Remark 4.3.5. Commutativity of the diagrams (4.20) and (4.21) corresponds to the statement that ∇ and η are morphisms of coalgebras or, equivalently, that Δ and ε are morphisms of algebras. △

Definition 4.3.6 (Hopf algebra)**.** A *Hopf algebra* H over $\Bbbk$ is a bialgebra over $\Bbbk$, together with a $\Bbbk$-linear map $S\colon H \to H$, called the *antipode* such that the following diagram commutes:

(4.22)
$$\begin{array}{ccccc} & & H \otimes H & \xrightarrow{S\otimes\mathrm{id}} & H \otimes H & & \\ & {\scriptstyle \Delta}\nearrow & & & & \searrow{\scriptstyle \nabla} & \\ H & & \xrightarrow{\varepsilon} & K & \xrightarrow{\eta} & & H \\ & {\scriptstyle \Delta}\searrow & & & & \nearrow{\scriptstyle \nabla} & \\ & & H \otimes H & \xrightarrow{\mathrm{id}\otimes S} & H \otimes H & & \end{array}$$

If H and H' are Hopf algebras, then a map $\varphi\colon H \to H'$ is a *morphism of Hopf algebras* if φ is a morphism of bialgebras. It then follows automatically (see [DNR01, Prop. 4.2.5]) that $f \circ S = S' \circ f$, where S and S' are the antipodes of H and H' respectively.

Remark 4.3.7. It can be shown (see [DNR01, Rem. 4.2.3]) that if a bialgebra admits an antipode, then the antipode is unique. Therefore, the antipode is not really additional structure. Being a Hopf algebra is a property that a bialgebra may possess. Furthermore, the antipode is an antihomomorphism of algebras and coalgebras (see [DNR01, Prop. 4.2.6]). In other words,

$$S \circ \eta = \eta, \qquad S \circ \nabla = \nabla \circ \mathrm{flip} \circ (S \otimes S), \tag{4.23}$$

$$\varepsilon \circ S = \varepsilon, \qquad \Delta \circ S = (S \otimes S) \circ \mathrm{flip} \circ \Delta. \tag{4.24}$$

The equations (4.23) are equivalent to the statements

$$S(1) = 1, \qquad S(ab) = S(b)S(a) \qquad \text{for all } a, b. \tag{4.25}$$

△

Examples 4.3.8. (a) If G is a group, then the group algebra $\Bbbk G$ is a cocommutative Hopf algebra with

$$\Delta g = g \otimes g, \qquad \varepsilon(g) = 1, \qquad S(g) = g^{-1} \qquad \text{for all } g \in G.$$

It is commutative if and only if G is abelian.

(b) If V is a vector space over a field $\Bbbk$, then the *tensor algebra* $T(V) := \bigoplus_{n=0}^{\infty} V^{\otimes n}$ is an algebra with multiplication given by tensor product and multiplicative identity given by the identity element of $V^{\otimes 0} \cong \Bbbk$. It is also a cocommutative Hopf algebra with

$$\Delta 1 = 1 \otimes 1, \quad \Delta v = v \otimes 1 + 1 \otimes v, \quad \varepsilon(v) = 0, \quad S(v) = -v, \qquad \text{for all } v \in V.$$

We extended these definitions to higher tensor powers using the fact that Δ and ε must be morphisms of algebras and S is an antihomomorphism of algebras. This Hopf algebra is commutative if and only if V has dimension less than or equal to 1.

(c) If $\mathfrak{g}$ is a Lie algebra, then its universal enveloping algebra $U(\mathfrak{g})$ is a cocommutative Hopf algebra with

$$\Delta a = a \otimes 1 + 1 \otimes a, \quad \varepsilon(a) = 0, \quad S(a) = -a, \qquad \text{for all } a \in \mathfrak{g}.$$

We then extend these definitions to all of $U(\mathfrak{g})$ as above. This Hopf algebra is commutative if and only if $\mathfrak{g}$ is an abelian Lie algebra.

(d) Quantized enveloping algebras provide examples of Hopf algebras that are neither commutative nor cocommutative. △

An element g in a coalgebra is said to be *grouplike* if

$$\varepsilon(g) = 1 \qquad \text{and} \qquad \Delta g = g \otimes g.$$

For example, in the Hopf algebra $\Bbbk G$ of Example 4.3.8(a), the elements of the group G are grouplike. (In fact, this is where the term *grouplike* comes from.) An element a in a bialgebra is said to be *primitive* if

$$\Delta a = a \otimes 1 + 1 \otimes a.$$

For example, in the Hopf algebra $U(\mathfrak{g})$ of Example 4.3.8(c), the elements of $\mathfrak{g}$ are primitive.

Remark 4.3.9. The main motivation behind the definition of bialgebras and Hopf algebras is the additional structure they provide to modules. If A and B are $\Bbbk$-algebras, U is an A-module, and V is a B-module, then $U \otimes V$ is naturally an $(A \otimes B)$-module with action given by

$$(a \otimes b)(u \otimes v) = au \otimes bv, \qquad a \in A,\ b \in B,\ u \in U,\ v \in V.$$

Thus, if U and V are both A-modules, $U \otimes V$ is naturally an $(A \otimes A)$-module. However, if A is a bialgebra, then $U \otimes V$ is also an A-module with action given by

$$a \cdot (u \otimes v) := \Delta(a)(u \otimes v), \tag{4.26}$$

extended by linearity. Furthermore, if A is a Hopf algebra, then the dual $V^* = \mathrm{Hom}_{\Bbbk}(V, \Bbbk)$ is also an A-module, with action

$$(af)(v) := f(S(a)v), \quad a \in A,\ f \in V^*,\ v \in V. \tag{4.27}$$

△

Examples 4.3.10. (a) Suppose G is a group and U and V are $\Bbbk G$-modules. Then $U \otimes V$ and V^* are $\Bbbk G$-modules with actions given by

$$g \cdot (u \otimes v) = gu \otimes gv, \quad (gf)(v) = f(g^{-1}v), \qquad g \in G,\ u \in U,\ v \in V,\ f \in V^*.$$

(b) Suppose $\mathfrak{g}$ is a Lie algebra and U and V are $\mathfrak{g}$-modules. Then $U \otimes V$ and V^* are $\mathfrak{g}$-modules with actions given by

$$x \cdot (u \otimes v) = xu \otimes v + u \otimes xv, \quad (xf)(v) = -f(xv), \qquad x \in \mathfrak{g},\ u \in U,\ v \in V,\ f \in V^*.$$

Here we use the fact that a $\mathfrak{g}$-module is equivalent to a $U(\mathfrak{g})$-module. △

Exercises.

4.3.1. An element g in a bialgebra is said to be *monoidlike* if $\Delta(g) = g \otimes g$. Show that the monoidlike elements form a monoid under multiplication. Show that, if $\Bbbk$ is a field, then the monoidlike elements are precisely the grouplike elements and 0.

4.3.2. Prove that the set of grouplike elements of a Hopf algebra is a group under the multiplication ∇ of the bialgebra.

4.3.3. Prove that the set of primitive elements of a bialgebra over $\Bbbk$ is a $\Bbbk$-submodule.

4.3.4. In the notation of Remark 4.3.9, prove that (4.26) defines the structure of an A-module on $U \otimes_{\Bbbk} V$ and that (4.27) defines the structure of an A-module on V^*.

4.3.5. Suppose (C, Δ, ε) is a counital coassociative coalgebra over a field $\Bbbk$. Show that the dual space $C^* = \mathrm{Hom}_{\Bbbk}(C, \Bbbk)$ is a unital associative algebra with multiplication

$$\nabla\colon C^* \otimes C^* \to C^*, \qquad \nabla(f \otimes g)(x) = (f \otimes g) \circ \Delta(x), \qquad f, g \in C^*,\ x \in C,$$

extended by linearity, and unit

$$\eta\colon \Bbbk \to C^*, \qquad \eta(a) = a\varepsilon, \qquad a \in \Bbbk.$$

4.4 Hopf algebra structure on the ring of symmetric functions

In this section we endow the ring of symmetric functions with the structure of a Hopf algebra. This structure is interesting in its own right, and will also be important in the applications to representation theory that we will discuss in the sequel. Throughout this section we work over the ground ring $\mathbb{Z}$.

We can identify

$$\mathrm{Sym} \otimes \mathrm{Sym}$$

with the ring of functions of bounded degree in two sets of indeterminates $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2 \dots)$ that are symmetric in each set separately via the map

$$f \otimes g \mapsto f(x)g(y), \qquad f, g \in \mathrm{Sym}.$$

This map sends the basis $m_\lambda \otimes m_\mu$, $\lambda, \mu \in \mathrm{Par}$, to the basis $m_\lambda(x)m_\mu(y)$, $\lambda, \mu \in \mathrm{Par}$. On the other hand, for any $f \in \mathrm{Sym}$, the function $f(x, y)$ is symmetric in $(x_1, y_1, x_2, y_2, \cdots)$. Therefore it also lies in $\mathrm{Sym} \otimes \mathrm{Sym}$. So we have a *comultiplication* map

$$\Delta \colon \mathrm{Sym} \to \mathrm{Sym} \otimes \mathrm{Sym}, \qquad (\Delta f)(x, y) = f(x, y). \tag{4.28}$$

We also define a *counit* map

$$\varepsilon := \rho_0 \colon \mathrm{Sym} \to \mathbb{Z}, \tag{4.29}$$

where ρ_0 is the map of (2.11) and we recall that $\mathrm{Sym}_0 = \mathbb{Z}$. Thus $\varepsilon(f)$ is the constant term of f. Finally, we define an antipode

$$S \colon \mathrm{Sym} \to \mathrm{Sym}, \qquad S(h_n) = (-1)^n e_n, \tag{4.30}$$

extended to all of Sym by (4.25).

Theorem 4.4.1. *The maps Δ, ε, and S endow* Sym *with the structure of a commutative and cocommutative Hopf algebra.*

Proof. We first show that $(\mathrm{Sym}, \Delta, \varepsilon)$ is a coalgebra. To do this, we identify $\mathrm{Sym}\otimes\mathrm{Sym}\otimes\mathrm{Sym}$ with the ring of functions in three sets of indeterminates $x = (x_1, x_2, \dots)$, $y = (y_1, y_2, \dots)$, and $z = (z_1, z_2, \dots)$ that are symmetric in each set of variables separately. Then, for $f \in \mathrm{Sym}$, we have

$$\big((\mathrm{id} \otimes \Delta) \circ \Delta(f)\big)(x, y, z) = f(x, y, z) = \big((\Delta \otimes \mathrm{id}) \circ \Delta(f)\big)(x, y, z),$$

and so the first diagram in (4.18) commutes. Next, note that $(\mathrm{id} \otimes \varepsilon) \circ \Delta(f)$ is the constant term, with respect to the y, of $f(x, y)$. This is precisely, $f(x)$. Thus

$$(\mathrm{id} \otimes \varepsilon) \circ \Delta(f) = f.$$

Similarly,

$$(\varepsilon \otimes \mathrm{id}) \circ \Delta(f) = f.$$

Thus the second diagram in (4.18) commutes. This completes the proof that $(\mathrm{Sym}, \Delta, \varepsilon)$ is a coalgebra.

Next we consider (4.20). For $f, g \in \mathrm{Sym}$, we have

$$\Delta \circ \nabla(f \otimes g) = \Delta(fg) = (fg)(x, y).$$

On the other hand

$$\begin{aligned}(\nabla \otimes \nabla)\circ(\mathrm{id} \otimes \mathrm{flip} \otimes \mathrm{id}) \circ (\Delta \otimes \Delta)(f \otimes g)& \\ = (\nabla \otimes \nabla) \circ (\mathrm{id} \otimes \mathrm{flip} \otimes \mathrm{id})(f(x, y) \otimes g(x, y))&\end{aligned}$$

$$= (fg)(x,y)$$
$$= \Delta \circ \nabla(f \otimes g).$$

Thus, diagram (4.20) commutes.

Next we compute that, for $f, g \in \mathrm{Sym}$,

$$\varepsilon \circ \nabla(f \otimes g) = \varepsilon(fg) = \varepsilon(f)\varepsilon(g),$$

since the constant term of fg is the product of the constant terms of f and g. So the first diagram in (4.21) commutes. It is straightforward to verify that the last two diagrams in (4.21) also commute. (See Exercise 4.4.1.)

It remains to show that S is an antipode, that is, that (4.22) commutes. Since Sym is commutative, S is a ring homomorphism by (4.25) (well defined by Theorem 2.5.3), and $\nabla\colon \mathrm{Sym} \otimes \mathrm{Sym} \to \mathrm{Sym}$ is a ring homomorphism. Since Δ is also a ring homomorphism, it follows that

$$\nabla \circ (S \otimes \mathrm{id}) \circ \Delta, \qquad \nabla \circ (\mathrm{id} \otimes S) \circ \Delta, \qquad \text{and} \qquad \eta \circ \varepsilon$$

are ring homomorphisms $\mathrm{Sym} \to \mathrm{Sym}$. By Theorem 2.5.3, it therefore suffices to show that they agree on h_n for $n \geq 1$. By (4.31) (whose proof does not use the present theorem), we have $\Delta h_n = \sum_{r=0}^{n} h_r \otimes h_{n-r}$, and so

$$\nabla \circ (S \otimes \mathrm{id}) \circ \Delta(h_n) = \sum_{r=0}^{n} (-1)^r e_r h_{n-r} \overset{(2.30)}{=} 0 = \eta \circ \varepsilon(h_n).$$

Similarly, replacing r by $n-r$,

$$\nabla \circ (\mathrm{id} \otimes S) \circ \Delta(h_n) = \sum_{r=0}^{n} (-1)^{n-r} h_r e_{n-r} = \sum_{r=0}^{n} (-1)^r e_r h_{n-r} = 0 = \eta \circ \varepsilon(h_n). \qquad \square$$

Proposition 4.4.2. *For all $n \geq 1$, we have*

$$\Delta h_n = \sum_{r=0}^{n} h_r \otimes h_{n-r}, \qquad \Delta e_n = \sum_{r=0}^{n} e_r \otimes e_{n-r}, \qquad \Delta p_n = p_n \otimes 1 + 1 \otimes p_n. \tag{4.31}$$

Proof. By definition, $h_n(x,y)$ is the sum of all monomials of total degree n in the indeterminates $x_1, y_1, x_2, y_2, \ldots$. Any such monomial is a product of a monomial in the x_i of degree r times a monomial of degree $n-r$ in the y_i, for some $0 \leq r \leq n$. Thus

$$(\Delta h_n)(x,y) = h_n(x,y) = \sum_{r=0}^{n} \sum_{\substack{1 \leq i_1 \leq i_2 \leq \cdots \leq i_r \\ 1 \leq j_1 \leq j_2 \leq \cdots \leq j_{n-r}}} x_{i_1} x_{i_2} \cdots x_{i_r} y_{j_1} y_{j_2} \cdots y_{j_{n-r}} = \sum_{r=0}^{n} h_r(x) h_{n-r}(y).$$

Similarly,

$$(\Delta e_n)(x,y) = e_n(x,y) = \sum_{r=0}^{n} \sum_{\substack{1 \leq i_1 < i_2 < \cdots < i_r \\ 1 \leq j_1 < j_2 < \cdots < j_{n-r}}} x_{i_1} x_{i_2} \cdots x_{i_r} y_{j_1} y_{j_2} \cdots y_{j_{n-r}} = \sum_{r=0}^{n} e_r(x) e_{n-r}(y).$$

Finally, we have

$$(\Delta p_n)(x,y) = p_n(x,y) = \sum_{i=1}^{\infty} (x_i^n + y_i^n) = p_n(x) + p_n(y). \qquad \square$$

Proposition 4.4.3. *For $\lambda \in \mathrm{Par}$, we have*

$$\Delta s_\lambda = \sum_{\mu \subseteq \lambda} s_{\lambda/\mu} \otimes s_\mu. \tag{4.32}$$

Proof. This follows immediately from Proposition 3.4.9. □

Proposition 4.4.4. *The involution ω of (2.31) is a morphism of Hopf algebras.*

Proof. Since we already know that ω is a morphism of algebras, it remains to prove that it is also a morphism of coalgebras. For $\mu \in \mathrm{Par}$, we have

$$\Delta(\omega(s_\lambda)) \overset{(3.17)}{=} \Delta(s_{\lambda'}) \overset{(4.32)}{=} \sum_{\mu' \subseteq \lambda'} s_{\lambda'/\mu'} \otimes s_{\mu'} \overset{(3.17)}{=} \sum_{\mu \subseteq \lambda} \omega(s_{\lambda/\mu}) \otimes \omega(s_\mu) = (\omega \otimes \omega)(\Delta(s_\lambda)),$$

where, in the third equality, we used the fact that $\mu \subseteq \lambda \iff \mu' \subseteq \lambda'$. Since the Schur functions span Sym, this implies that $\Delta \circ \omega = (\omega \otimes \omega) \circ \Delta$. Furthermore, since the constant term of $\omega(f)$ is the same as the constant term of f for all $f \in \mathrm{Sym}$, we also have $\varepsilon \circ \omega = \varepsilon$. □

Proposition 4.4.5. *Suppose $(u_\lambda)_{\lambda \in \mathrm{Par}}$ and $(v_\lambda)_{\lambda \in \mathrm{Par}}$ are dual bases of* Sym*. Then, for all $f \in \mathrm{Sym}$,*

$$\Delta f = \sum_{\lambda \in \mathrm{Par}} u_\lambda^\perp f \otimes v_\lambda.$$

Proof. It follows from Lemma 4.2.3 and Proposition 4.4.3 that

$$\Delta s_\mu = \sum_{\lambda \in \mathrm{Par}} s_{\mu/\lambda} \otimes s_\lambda = \sum_{\lambda \in \mathrm{Par}} s_\lambda^\perp s_\mu \otimes s_\lambda.$$

Since the Schur functions span Sym, it follows by linearity that

$$\Delta f = \sum_{\lambda \in \mathrm{Par}} s_\lambda^\perp f \otimes s_\lambda$$

for all $f \in \mathrm{Sym}$. This proves the proposition for the special case where $u_\lambda = v_\lambda = s_\lambda$ for all $\lambda \in \mathrm{Par}$.

Now let $(u_\lambda)_{\lambda \in \mathrm{Par}}$ and $(v_\lambda)_{\lambda \in \mathrm{Par}}$ be arbitrary dual bases of Sym. Let $M = M(s, v)$ be the transition matrix defined as in (3.57). Then, by (3.30) and (3.60), we have $M(s, u) = M^{-\mathsf{T}}$. It thus follows from Proposition 4.2.1 (in particular, the fact that the map $f \mapsto f^\perp$ is linear) that

$$s_\lambda^\perp = \sum_{\nu \in \mathrm{Par}} (M^{-1})_{\nu,\lambda} u_\nu^\perp.$$

Thus

$$\begin{aligned} \Delta f &= \sum_{\lambda \in \mathrm{Par}} s_\lambda^\perp f \otimes s_\lambda \\ &= \sum_{\lambda,\mu,\nu \in \mathrm{Par}} (M^{-1})_{\nu,\lambda} u_\nu^\perp f \otimes M_{\lambda,\mu} v_\mu \end{aligned}$$

$$= \sum_{\mu,\nu\in\mathrm{Par}} \left(\sum_{\lambda\in\mathrm{Par}} (M^{-1})_{\nu,\lambda} M_{\lambda,\mu} \right) u_\nu^\perp f \otimes v_\mu$$
$$= \sum_{\mu,\nu\in\mathrm{Par}} \delta_{\nu\mu} u_\nu^\perp f \otimes v_\mu$$
$$= \sum_{\mu\in\mathrm{Par}} u_\mu^\perp f \otimes v_\mu. \qquad \square$$

Theorem 4.4.6. *For all $f, g, h \in \mathrm{Sym}$, we have*

$$\langle \Delta f, g \otimes h \rangle = \langle f, gh \rangle, \tag{4.33}$$

where, on the left-hand side, we use the bilinear form on $\mathrm{Sym} \otimes \mathrm{Sym}$ *defined as in (4.6).*

Proof. Since the equation to be proved is linear in f, g, and h, it suffices to prove it when f, g, h are Schur functions, as these span Sym. For $\lambda, \mu, \nu \in \mathrm{Par}$, we have

$$\begin{aligned}
\langle \Delta s_\lambda, s_\mu \otimes s_\nu \rangle &\overset{(4.32)}{=} \sum_{\rho\subseteq\lambda} \langle s_{\lambda/\rho} \otimes s_\rho, s_\mu \otimes s_\nu \rangle \\
&= \sum_{\rho\in\mathrm{Par}} \langle s_{\lambda/\rho}, s_\mu \rangle \langle s_\rho, s_\nu \rangle \\
&\overset{(4.10)}{=} \sum_{\rho\in\mathrm{Par}} \langle s_\lambda, s_\rho s_\mu \rangle \langle s_\rho, s_\nu \rangle \\
&= \langle s_\lambda, s_\mu s_\nu \rangle,
\end{aligned}$$

where, in the second equality, we used the fact that $s_{\lambda/\rho} = 0$ unless $\rho \subseteq \lambda$ to change to a sum over all $\rho \in \mathrm{Sym}$. $\square$

Remark 4.4.7. Theorem 4.4.6 states that the comultiplication Δ is the adjoint of the multiplication map $\mathrm{Sym} \otimes \mathrm{Sym} \to \mathrm{Sym}$. Since the Hall inner product is positive definite (hence nondegenerate), this implies that the ring of symmetric functions is a *self-dual Hopf algebra*. (One can also show that the unit and counit are adjoint.) $\triangle$

Proposition 4.4.2 implies that the power sum p_n are primitive elements of Sym. The set of primitive elements is a $\mathbb{Z}$-submodule of Sym; see Exercise 4.3.3.

Proposition 4.4.8. *The p_n, $n \geq 1$, are a $\mathbb{Z}$-basis of the set of primitive elements of* Sym.

Proof. Since we already know that the p_n are linearly independent (Theorem 2.6.4), it remains to show that they span the set of primitive elements. Suppose $p \in \mathrm{Sym}$ is primitive, and write $p = \sum_{n\geq 0} p^{(n)}$ with $p^{(n)} \in \mathrm{Sym}^n$. Since $\Delta(\mathrm{Sym}^n) \subseteq \bigoplus_{k=0}^n \mathrm{Sym}^k \otimes \mathrm{Sym}^{n-k}$, comparing the components of both sides of $\Delta(p) = p\otimes 1 + 1 \otimes p$ lying in $\bigoplus_{k=0}^n \mathrm{Sym}^k \otimes \mathrm{Sym}^{n-k}$ shows that each $p^{(n)}$ is primitive. For $n = 0$, we have $p^{(0)} = c \in \mathbb{Z}$, and $c(1 \otimes 1) = \Delta(c) = 2c(1 \otimes 1)$ forces $c = 0$, since $1 \otimes 1$ is a basis element of the free $\mathbb{Z}$-module $\mathrm{Sym} \otimes \mathrm{Sym}$ by Proposition 4.1.5(b). So it suffices to consider a primitive element $p \in \mathrm{Sym}^n$ with $n \geq 1$. Then, for all homogeneous $f, g \in \mathrm{Sym}$ of positive degree, we have

$$\langle p, fg \rangle \overset{(4.33)}{=} \langle \Delta p, f \otimes g \rangle = \langle 1 \otimes p + p \otimes 1, f \otimes g \rangle \overset{(4.6)}{=} \langle 1, f \rangle \langle p, g \rangle + \langle p, f \rangle \langle 1, g \rangle = 0,$$

since the Hall inner product of 1 and any element of positive degree is zero. It follows that

$$\langle p, h_\lambda \rangle = 0 \qquad \text{for all } \lambda \in \mathrm{Par}(n),\ \lambda \neq (n).$$

Writing p as a linear combination of monomial symmetric functions, we have

$$p = \sum_{\mu \in \mathrm{Par}(n)} d_\mu m_\mu, \qquad d_\mu \in \mathbb{Z}.$$

Then, for all $\lambda \in \mathrm{Par}(n)$, $\lambda \neq (n)$, we have

$$0 = \langle p, h_\lambda \rangle = \sum_{\mu \in \mathrm{Par}(n)} d_\mu \langle m_\mu, h_\lambda \rangle \overset{(3.18)}{=} d_\lambda.$$

Thus

$$p = d_{(n)} m_{(n)} = d_{(n)} p_n$$

is a multiple of the power sum p_n. □

We conclude this section with some results that will be useful when we discuss the Heisenberg algebra in Section 5.1.

Proposition 4.4.9. *For all $f, g, h \in \mathrm{Sym}$,*

$$f^\perp(gh) = \sum_{(f)} \left(f^\perp_{(1)} g\right)\left(f^\perp_{(2)} h\right). \tag{4.34}$$

Proof. For any $u \in \mathrm{Sym}$, we have

$$\begin{aligned}
\langle u, f^\perp(gh) \rangle &= \langle fu, gh \rangle \\
&\overset{(4.33)}{=} \langle \Delta(fu), g \otimes h \rangle \\
&\overset{(4.20)}{=} \langle (\Delta f)(\Delta u), g \otimes h \rangle \\
&= \sum_{(f)} \langle (f_{(1)} \otimes f_{(2)})(\Delta u), g \otimes h \rangle \\
&= \sum_{(f)} \left\langle \Delta u, f^\perp_{(1)} g \otimes f^\perp_{(2)} h \right\rangle \\
&\overset{(4.33)}{=} \sum_{(f)} \left\langle u, \left(f^\perp_{(1)} g\right)\left(f^\perp_{(2)} h\right) \right\rangle,
\end{aligned}$$

where the fifth equality uses Exercise 4.2.5. Since $u \in \mathrm{Sym}$ was arbitrary, (4.34) follows. □

Proposition 4.4.10. *For all $f, g \in \mathrm{Sym}$,*

$$f^\perp g = \sum_{(g)} \langle f, g_{(1)} \rangle g_{(2)}. \tag{4.35}$$

Proof. For all $f, g, h \in \mathrm{Sym}$, we have

$$\langle f^\perp g, h \rangle = \langle g, fh \rangle \overset{(4.33)}{=} \langle \Delta g, f \otimes h \rangle = \sum_{(g)} \langle g_{(1)}, f \rangle \langle g_{(2)}, h \rangle = \left\langle \sum_{(g)} \langle f, g_{(1)} \rangle g_{(2)}, h \right\rangle.$$

Since h was arbitrary, the result follows. □

Exercises.

4.4.1. Complete the proof of Theorem 4.4.1 by showing that the last two diagrams in (4.21) commute.

4.4.2. Find the group of grouplike elements of $\mathrm{Sym}_{\mathbb{Q}}$. (See Exercise 4.3.2.)

4.4.3. Show that $S(f)(x_1, x_2, \dots) = \omega(f)(-x_1, -x_2, \dots)$ for all $f \in \mathrm{Sym}$.

4.4.4. Give explicit formulas for the action of the antipode S on e_λ, h_λ, p_λ, and s_λ, for $\lambda \in \mathrm{Par}$.

4.4.5. Prove that $\nabla \circ \Delta(p_\lambda) = 2^{\ell(\lambda)} p_\lambda$ for all $\lambda \in \mathrm{Par}$, where ∇ is the multiplication map $\mathrm{Sym} \otimes \mathrm{Sym} \to \mathrm{Sym}$.

4.4.6. Prove that, for all $m, n \geq 1$,

$$h_m^\perp h_n = h_{n-m} \qquad \text{and} \qquad h_m^\perp e_n = \begin{cases} e_n & \text{if } m = 0, \\ e_{n-1} & \text{if } m = 1, \\ 0 & \text{if } m > 1, \end{cases} \tag{4.36}$$

where we recall the convention that $h_r = 0$ for $r < 0$.

4.4.7. Prove that, for all $m, n \geq 1$,

$$e_m^\perp e_n = e_{n-m} \qquad \text{and} \qquad e_m^\perp h_n = \begin{cases} h_n & \text{if } m = 0, \\ h_{n-1} & \text{if } m = 1, \\ 0 & \text{if } m > 1, \end{cases} \tag{4.37}$$

where we recall the convention that $e_r = 0$ for $r < 0$.

Chapter 5

The boson-fermion correspondence

In this section we explore the boson-fermion correspondence, which is a precise relationship between a natural representation, known as bosonic Fock space, of the infinite rank Heisenberg algebra, and a natural representation, known as fermionic Fock space, of the infinite rank Clifford algebra. We begin by defining the Heisenberg algebra and its action on bosonic Fock space. We then define the Clifford algebra and its action on fermionic Fock space. Finally we describe the two halves of the boson-fermion correspondence: bosonization and fermionization.

5.1 The Heisenberg algebra

In this section we construct the Heisenberg algebra from the ring of symmetric functions. We start with a rather abstract definition, as the Heisenberg double of the Hopf algebra of symmetric functions, and then give a more concrete presentation that matches the usual definition of the Heisenberg algebra.

Recall from (4.7) that we have an injective ring homomorphism

$$\text{Sym} \to \text{End}(\text{Sym}), \quad f \mapsto f^+, \quad \text{where } f^+(v) = fv, \tag{5.1}$$

sending $f \in \text{Sym}$ to the operation f^+ of multiplication by f. We let Sym^+ denote the image of this map. Recall also, from Proposition 4.2.1, that we have an injective ring homomorphism

$$\text{Sym} \to \text{End}(\text{Sym}), \quad f \mapsto f^- := f^\perp, \tag{5.2}$$

where $f^\perp$ is the operator adjoint to f^+ with respect to the Hall inner product. We let Sym^- denote the image of this map. Thus Sym^+ and Sym^- are both subrings of End(Sym) isomorphic to Sym.

Definition 5.1.1 (Heisenberg double)**.** We define the *Heisenberg double* Heis of Sym to be the subring of End(Sym) generated by Sym^+ and Sym^-. In other words Heis is the subring of End(Sym) generated by the operations of multiplication by elements of Sym and the adjoints of these operations with respect to the Hall inner product.

We can also work over $\mathbb{Q}$ and define $\text{Heis}_\mathbb{Q}$ as above, replacing Sym by $\text{Sym}_\mathbb{Q}$ everywhere. Then we have

$$\text{Heis}_\mathbb{Q} := \mathbb{Q} \otimes_\mathbb{Z} \text{Heis}.$$

Remark 5.1.2. The Heisenberg double is a more general construction than what we have presented here. More generally, one can form a Heisenberg double from a pair of dual Hopf algebras. Since Sym is self-dual as a Hopf algebra (Remark 4.4.7), we are able to take the Heisenberg double of Sym with itself. △

We define an involution

$$\omega\colon \mathrm{Heis} \to \mathrm{Heis}, \qquad \omega(a) := \omega \circ a \circ \omega, \quad a \in \mathrm{Heis}, \tag{5.3}$$

where, on the right-hand side, ω is the involution of Sym from (2.31). We leave it as Exercise 5.1.1 to show that

$$\omega(f^{\pm}) = \omega(f)^{\pm} \qquad \text{for all } f \in \mathrm{Sym}. \tag{5.4}$$

Thus, under the inclusion (5.1) of Sym in Heis, the involution (5.3) extends the involution (2.31). This justifies our use of the same notation for both.

Our goal now is to give a *presentation* of the Heisenberg double. In other words, we want a description in terms of generators and relations. We begin by computing a key commutation relation.

Proposition 5.1.3. *For $f, g \in \mathrm{Sym}$, we have*

$$f^- g^+ = \sum_{(f)} \left(f^{\perp}_{(1)} g\right)^+ f^-_{(2)}, \tag{5.5}$$

where we have used Sweedler notation (4.19).

Proof. For $f, g, h \in \mathrm{Sym}$,

$$(f^- g^+)h = f^{\perp}(gh) \overset{(4.34)}{=} \sum_{(f)} \left(f^{\perp}_{(1)} g\right)\left(f^{\perp}_{(2)} h\right) = \left(\sum_{(f)} \left(f^{\perp}_{(1)} g\right)^+ f^-_{(2)}\right) h. \qquad \square$$

Proposition 5.1.4. *We have an isomorphism of $\mathbb{Z}$-modules*

$$\beta\colon \mathrm{Sym} \otimes_{\mathbb{Z}} \mathrm{Sym} \xrightarrow{\cong} \mathrm{Heis}, \qquad f \otimes g \mapsto f^+ g^-.$$

Proof. Since the map

$$\mathrm{Sym} \times \mathrm{Sym} \to \mathrm{Heis}, \qquad (f, g) \mapsto f^+ g^-,$$

is bilinear, it induces the map β, which is therefore well defined. By definition, elements of Heis are linear combinations of products of $f^{\pm}$, $f \in \mathrm{Sym}$. Proposition 5.1.3 allows us to move elements of the form f^+ to the left of elements of the form f^-. Thus the map β is surjective.

It remains to prove that β is injective. Suppose, towards a contradiction, that $\beta(a) = 0$ for some $a \in \mathrm{Sym} \otimes \mathrm{Sym}$, $a \neq 0$. By Proposition 4.1.5(b), we can write

$$a = \sum_{\lambda,\mu \in \mathrm{Par}} a_{\lambda\mu} s_\lambda \otimes s_\mu, \qquad a_{\lambda\mu} \in \mathbb{Z},$$

where only finitely many of the $a_{\lambda\mu}$ are nonzero. Choose ν minimal, with respect to the partial order $\subseteq$, amongst the partitions such that $a_{\lambda\nu} \neq 0$ for some λ. Then we have

$$0 = \beta(a)(s_\nu) = \sum_{\lambda,\mu\in\mathrm{Par}} a_{\lambda\mu} s_\lambda s_\mu^\perp s_\nu \overset{(4.10)}{=} \sum_{\lambda,\mu\in\mathrm{Par}} a_{\lambda\mu} s_\lambda s_{\nu/\mu} \overset{(3.45)}{=} \sum_{\lambda\in\mathrm{Par}} a_{\lambda\nu} s_\lambda.$$

Since the Schur functions are linearly independent, this implies that $a_{\lambda\nu} = 0$ for all $\lambda \in \mathrm{Par}$, which is a contradiction. Hence β is injective. □

Corollary 5.1.5. *If $(u_\lambda)_{\lambda\in\mathrm{Par}}$ and $(v_\lambda)_{\lambda\in\mathrm{Par}}$ are two bases of* Sym *(respectively,* $\mathrm{Sym}_\mathbb{Q}$*), then $u_\lambda^+ v_\mu^-$, $\lambda, \mu \in \mathrm{Par}$, is a $\mathbb{Z}$-basis of* Heis *(respectively, a $\mathbb{Q}$-basis of* $\mathrm{Heis}_\mathbb{Q}$*).*

Proof. This follows immediately from Proposition 5.1.4. □

Propositions 5.1.3 and 5.1.4 allow us to find many different presentations of Heis and $\mathrm{Heis}_\mathbb{Q}$. We can choose any set of generators for Sym^+, any set of generators for Sym^-, and then use Proposition 5.1.3 to compute the relations satisfied by these elements. The following theorem is one example of this.

Theorem 5.1.6. *As a $\mathbb{Q}$-algebra,* $\mathrm{Heis}_\mathbb{Q}$ *is generated by $p_n^\pm$, $n \geq 1$, subject to the relations*

$$p_m^+ p_n^+ = p_n^+ p_m^+, \quad p_m^- p_n^- = p_n^- p_m^-, \quad p_m^- p_n^+ = p_n^+ p_m^- + n\delta_{mn}, \qquad m, n \geq 1. \tag{5.6}$$

Proof. Let H denote the $\mathbb{Q}$-algebra generated by $p_n^\pm$, $n \geq 1$, subject to the relations (5.6). In $\mathrm{Heis}_\mathbb{Q}$, we have $p_m^+ p_n^+ = p_n^+ p_m^+$, $p_m^- p_n^- = p_n^- p_m^-$, and

$$p_m^- p_n^+ \overset{(5.5)}{\underset{(4.31)}{=}} (1^\perp p_n)^+ p_m^- + (p_m^\perp p_n)^+ 1^- \overset{(4.13)}{=} p_n^+ p_m^- + n\delta_{mn}.$$

Thus we have a map $H \to \mathrm{Heis}_\mathbb{Q}$, $p_n^\pm \mapsto p_n^\pm$. This map is surjective by Theorem 2.6.4. Using the relations (5.6), we can move every p_m^- to the right of every p_n^+ in any product of generators of H. Thus H is spanned by the products $p_{\lambda_1}^+ \cdots p_{\lambda_k}^+ p_{\mu_1}^- \cdots p_{\mu_l}^-$ with $(\lambda_1, \ldots, \lambda_k), (\mu_1, \ldots, \mu_l) \in \mathrm{Par}$. Their images $p_\lambda^+ p_\mu^-$ are linearly independent by Theorem 2.6.4 and Corollary 5.1.5. Hence the map is injective. (See Remark 5.1.7.) □

The *infinite rank Heisenberg algebra* is the $\mathbb{Q}$-algebra generated by $p_n^\pm$, $n \geq 1$, subject to the relations (5.6). Thus, Theorem 5.1.6 states that the Heisenberg double is isomorphic to the infinite rank Heisenberg algebra. From now on, we will identify the two, and simply refer to $\mathrm{Heis}_\mathbb{Q}$ as the *Heisenberg algebra*. The relations (5.6), which are sometimes called the *canonical commutations relations*, are the relations satisfied by position and momentum operators in quantum mechanics. These relations imply the Heisenberg uncertainty principle.

Remark 5.1.7. The *Heisenberg Lie algebra* $\mathfrak{h}$ is the Lie algebra spanned (over $\mathbb{Q}$) by $p_n^\pm$, $n \geq 1$, and c, with Lie bracket given by

$$[p_m^+, p_n^+] = 0 = [p_m^-, p_n^-], \qquad [p_m^-, p_n^+] = n\delta_{mn} c, \qquad [c, p_n^\pm] = 0, \qquad m, n \geq 1.$$

Then $\mathrm{Heis}_\mathbb{Q}$ is the quotient of the universal enveloping algebra of $\mathfrak{h}$ by the relation $c = 1$. It follows from the Poincaré–Birkhoff–Witt theorem that this quotient is isomorphic to $\mathbb{Q}[p_1^+, p_2^+, \ldots] \otimes_\mathbb{Z} \mathbb{Q}[p_1^-, p_2^-, \ldots]$ as a $\mathbb{Z}$-module. △

Remark 5.1.8. In order to work with the generators $p_n^{\pm}$, we need to work over $\mathbb{Q}$ instead of over $\mathbb{Z}$ (see Remark 2.6.5). However, if we instead choose generators $e_n^{\pm}$ or $h_n^{\pm}$, we can give a presentation of Heis, which is an integral form of the Heisenberg algebra. See Exercises 5.1.2 to 5.1.4. △

The definition of a $\mathbb{Z}$-graded ring is obtained from Definition 1.3.1 by replacing $\mathbb{N}$ by $\mathbb{Z}$ everywhere, and the definition of a $\mathbb{Z}$-graded algebra is analogous. The Heisenberg algebra is a $\mathbb{Z}$-graded algebra, with gradation given by

$$\text{(5.7)} \qquad \mathrm{Heis}_{\mathbb{Q}} = \bigoplus_{n\in\mathbb{Z}} \mathrm{Heis}_{\mathbb{Q}}^n, \qquad \mathrm{Heis}_{\mathbb{Q}}^n := \{a \in \mathrm{Heis}_{\mathbb{Q}} : a\mathrm{Sym}_{\mathbb{Q}}^m \subseteq \mathrm{Sym}_{\mathbb{Q}}^{m+n} \text{ for all } m \in \mathbb{N}\},$$

where we adopt the convention that $\mathrm{Sym}^m = 0$ for $m < 0$.

Proposition 5.1.9. *For all $n \in \mathbb{Z}$, we have*

$$\mathrm{Heis}_{\mathbb{Q}}^n = \mathrm{Span}_{\mathbb{Q}}\{s_\lambda^+ s_\mu^- : \lambda, \mu \in \mathrm{Par},\ |\lambda| - |\mu| = n\}.$$

Proof. By Corollary 5.1.5, $s_\lambda^+ s_\mu^-$, $\lambda, \mu \in \mathrm{Par}$, is a $\mathbb{Q}$-basis of $\mathrm{Heis}_{\mathbb{Q}}$. Thus, it suffices to show that

$$s_\lambda^+ s_\mu^- \in \mathrm{Heis}_{\mathbb{Q}}^{|\lambda|-|\mu|} \qquad \text{for all } \lambda, \mu \in \mathrm{Par}.$$

Suppose $f \in \mathrm{Sym}_{\mathbb{Q}}^m$. Then, by Lemma 4.2.2, we have $s_\mu^\perp f \in \mathrm{Sym}_{\mathbb{Q}}^{m-|\mu|}$. Thus

$$s_\lambda^+ s_\mu^- f = s_\lambda s_\mu^\perp f \in \mathrm{Sym}_{\mathbb{Q}}^{m+|\lambda|-|\mu|}.$$

□

Exercises.

5.1.1. Prove that the involution ω of (5.4) satisfies $\omega(f^{\pm}) = \omega(f)^{\pm}$ for all $f \in \mathrm{Sym}$.

5.1.2. Prove that

$$h_m^- e_n^+ = e_n^+ h_m^- + e_{n-1}^+ h_{m-1}^- \quad \text{and} \quad e_m^- h_n^+ = h_n^+ e_m^- + h_{n-1}^+ e_{m-1}^-$$

for all $m, n \in \mathbb{N}$.

5.1.3. Prove that

$$h_m^- h_n^+ = \sum_{r=0}^{\min(m,n)} h_{n-r}^+ h_{m-r}^- \quad \text{and} \quad e_m^- e_n^+ = \sum_{r=0}^{\min(m,n)} e_{n-r}^+ e_{m-r}^-$$

for all $m, n \in \mathbb{N}$.

5.1.4. Prove that

$$s_\lambda^- s_\mu^+ = \sum_{\substack{\nu \subseteq \lambda \\ \rho \subseteq \mu}} c_{\nu\rho}^\lambda s_{\mu/\rho}^+ s_\nu^-$$

for all $\lambda, \mu \in \mathrm{Par}$.

5.1.5. Let $m, n \in \mathbb{N}$. In $\mathrm{Heis}_\mathbb{Q}$, write $p_m^- e_n^+$ and $p_m^- h_n^+$ as linear combinations of elements of the form $f^+ g^-$, $f, g \in \mathrm{Sym}$.

5.2 Bosonic Fock space

In this section we define bosonic Fock space, which is a natural action of the Heisenberg algebra on the space of symmetric functions. Throughout this section we will work over the field $\mathbb{Q}$ of rational numbers. (However, see Remark 5.2.4.)

Recall that if A is a $\mathbb{Q}$-algebra, then a left A-module is equivalent to a $\mathbb{Q}$-algebra homomorphism

$$A \to \mathrm{End}_\mathbb{Q}(V) \tag{5.8}$$

for some $\mathbb{Q}$-module V. Here $\mathrm{End}_\mathbb{Q}(V)$ denotes the $\mathbb{Q}$-algebra of $\mathbb{Q}$-linear endomorphisms of V. The map (5.8) is called a *representation* of A.

Since $\mathrm{Heis}_\mathbb{Q}$ is, by definition, a subring of $\mathrm{End}_\mathbb{Q}(\mathrm{Sym}_\mathbb{Q})$, we have the inclusion map

$$\mathrm{Heis}_\mathbb{Q} \hookrightarrow \mathrm{End}_\mathbb{Q}(\mathrm{Sym}_\mathbb{Q}).$$

This is called the *Fock space representation* of the Heisenberg algebra $\mathrm{Heis}_\mathbb{Q}$, and the $\mathrm{Heis}_\mathbb{Q}$-module $\mathrm{Sym}_\mathbb{Q}$ is called *bosonic Fock space*.

Remark 5.2.1. The term *boson* refers to a type of subatomic particle with the property that many particles can occupy the same state. If we view p_n as a particle in state n, then, for example, $p_1^3 p_2^2 p_5$ corresponds to a system with three particles in state 1, two particles in state 2, and one particle in state 5. Soon we will encounter *fermionic* Fock space, where there can be at most one particle in each state. See Remark 5.3.4. △

Recall that a left ideal I of a $\mathbb{Q}$-algebra A is *maximal* if the only left ideal of A properly containing I is equal to A itself.

Lemma 5.2.2. *The set*

$$I^- := \mathrm{Span}_\mathbb{Q}\{f^+ g^- : f \in \mathrm{Sym}_\mathbb{Q},\ g \in \mathrm{Sym}_\mathbb{Q}^n,\ n \geq 1\}$$

is a maximal left ideal of $\mathrm{Heis}_\mathbb{Q}$.

Proof. The set I^- is clearly closed under addition. For $p, f, g \in \mathrm{Sym}_\mathbb{Q}$, $h \in \mathrm{Sym}_\mathbb{Q}^n$, $n \geq 1$, we have

$$\left(p^+ f^-\right)\left(g^+ h^-\right) \overset{(5.5)}{=} \sum_{(f)} p^+ \left(f_{(1)}^\perp g\right)^+ f_{(2)}^- h^- = \sum_{(f)} \left(p f_{(1)}^\perp g\right)^+ \left(f_{(2)} h\right)^- \in I^-.$$

Hence I^- is a left ideal of Heis.

It remains to show that I^- is maximal. It follows from Corollary 5.1.5 that we have a decomposition of $\mathbb{Q}$-modules

$$\mathrm{Heis}_{\mathbb{Q}} = I^- \oplus \mathrm{Sym}^+_{\mathbb{Q}}.$$

Thus $I^- \neq \mathrm{Heis}_{\mathbb{Q}}$. Now suppose J is a left ideal of $\mathrm{Heis}_{\mathbb{Q}}$ properly containing I^-. Then it follows from the above decomposition that J contains a nonzero element $f^+ \in \mathrm{Sym}^+_{\mathbb{Q}}$. Amongst all such elements, choose one for which f has the smallest degree. If we can show that f is a nonzero constant polynomial, then J contains the multiplicative identity 1, and so $J = \mathrm{Heis}_{\mathbb{Q}}$, and we are done.

Suppose, towards a contradiction, that f is nonconstant. By Theorem 2.6.4, f is nonconstant polynomial in the power sums p_n. So there exists $n \in \mathbb{N}$ such that $\frac{\partial f}{\partial p_n} \neq 0$. Now note that

$$p_n^- f^+ \underset{(4.31)}{\overset{(5.5)}{=}} (p_n^\perp f)^+ + f^+ p_n^-,$$

and so

$$n\left(\frac{\partial f}{\partial p_n}\right)^+ \overset{(4.13)}{=} (p_n^\perp f)^+ = p_n^- f^+ - f^+ p_n^- \in J,$$

(We use here the fact that $f^+ p_n^- \in I^- \subseteq J$ by the definition of I^-, and $p_n^- f^+ \in J$ since J is a left ideal.) Since the degree of $n\frac{\partial f}{\partial p_n}$ is less than the degree of f, this contradicts our choice of f. □

Recall that the *annihilator* of an element v in an A-module V is

$$\mathrm{Ann}_A(v) = \{a \in A : av = 0\}.$$

The annihilator $\mathrm{Ann}_A(v)$ is a left ideal of A. If V is a $\mathrm{Heis}_{\mathbb{Q}}$-module, then we call a nonzero element $v \in V$ a *highest weight vector* if $I^- v = 0$. In other words, v is a highest weight vector if $I^- \subseteq \mathrm{Ann}_{\mathrm{Heis}}(v)$. For example, any nonzero constant is a highest weight vector in $\mathrm{Sym}_{\mathbb{Q}}$.

The following result gives an important characterization of bosonic Fock space. It is closely related to the Stone–von Neumann theorem. Recall that a nonzero A-module V is *simple* if its only submodules are $\{0\}$ and V.

Theorem 5.2.3. *Bosonic Fock space is a simple* $\mathrm{Heis}_{\mathbb{Q}}$*-module. Furthermore, any* $\mathrm{Heis}_{\mathbb{Q}}$*-module generated by a highest weight vector is isomorphic to bosonic Fock space.*

Proof. Since I^- is maximal left ideal of $\mathrm{Heis}_{\mathbb{Q}}$ by Lemma 5.2.2, the quotient $\mathrm{Heis}_{\mathbb{Q}}/I^-$ is a simple $\mathrm{Heis}_{\mathbb{Q}}$-module. Since bosonic Fock space is generated by the highest weight element $1 \in \mathrm{Sym}_{\mathbb{Q}}$, it thus suffices to prove that any $\mathrm{Heis}_{\mathbb{Q}}$-module generated by a highest weight vector is isomorphic to $\mathrm{Heis}_{\mathbb{Q}}/I^-$.

Suppose V is a $\mathrm{Heis}_{\mathbb{Q}}$-module generated by a highest weight vector v. (Since highest weight vectors are nonzero by definition, this implies that V is nonzero.) Then we have a surjective homomorphism of left $\mathrm{Heis}_{\mathbb{Q}}$-modules

$$\varphi\colon \mathrm{Heis}_{\mathbb{Q}} \to V, \qquad a \mapsto av.$$

By the definition of a highest weight vector, we have $I^- \subseteq \ker(\varphi)$. Since I^- is maximal by Lemma 5.2.2, and $\ker(\varphi) \neq \mathrm{Heis}_{\mathbb{Q}}$ since V is nonzero, it follows that $I^- = \ker(\varphi)$. Thus, by the first isomorphism theorem for modules, we have

$$V \cong \mathrm{Heis}_{\mathbb{Q}}/I^-. \qquad \square$$

It follows immediately from the definition (5.7) of the $\mathbb{Z}$-grading on $\mathrm{Heis}_{\mathbb{Q}}$ that bosonic Fock space is a graded module, in the sense that

$$\mathrm{Heis}^m_{\mathbb{Q}}\mathrm{Sym}^n_{\mathbb{Q}} \subseteq \mathrm{Sym}^{m+n}_{\mathbb{Q}}, \qquad m, n \in \mathbb{Z}. \tag{5.9}$$

Remark 5.2.4. Throughout this section we worked over the ground ring $\mathbb{Q}$. It is also possible to work over $\mathbb{Z}$ with some minor modifications. In particular, bosonic Fock space is no longer a simple Heis module over $\mathbb{Z}$, since $n\mathrm{Sym}$ is a Heis-submodule of Sym for all $n \in \mathbb{Z}$. $\triangle$

Exercises.

5.2.1. Suppose that

$$f = \sum_{\lambda \in \mathrm{Par}} c_\lambda p_\lambda$$

for some $c_\lambda \in \mathbb{Q}$, $\lambda \in \mathrm{Par}$, only finitely many of which are nonzero. Choose λ of maximal size amongst the partitions such that $c_\lambda \neq 0$. Show that

$$p^-_\lambda f \overset{(4.13)}{=} \left(\prod_{i=1}^{\ell(\lambda)} \lambda_i \frac{\partial}{\partial p_{\lambda_i}} \right) f$$

is a nonzero constant polynomial.

5.2.2. Give an example of a $\mathrm{Heis}_{\mathbb{Q}}$-module that is *not* generated by a highest weight vector.

5.3 Fermionic Fock space

In this section we define the infinite rank Clifford algebra and its natural action on fermionic Fock space. Throughout this section we work over the field $\mathbb{Q}$ of rational numbers.

Definition 5.3.1 (Clifford algebra)**.** The *infinite rank Clifford algebra* Cl is the $\mathbb{Q}$-algebra generated by $\psi^\pm_i$, $i \in \mathbb{Z}$, subject to the relations

$$\psi^+_i\psi^+_j + \psi^+_j\psi^+_i = 0, \qquad \psi^-_i\psi^-_j + \psi^-_j\psi^-_i = 0, \qquad \psi^+_i\psi^-_j + \psi^-_j\psi^+_i = \delta_{ij}, \tag{5.10}$$

for all $i, j \in \mathbb{Z}$.

Let V be the $\mathbb{Q}$-module with basis v_i, $i \in \mathbb{Z}$. Then define $\mathcal{F}$ to be the $\mathbb{Q}$-module with basis

$$v_{i_0} \wedge v_{i_1} \wedge v_{i_2} \wedge \cdots, \qquad i_0 > i_1 > i_2 > \cdots, \quad i_{k+1} = i_k - 1 \text{ for } k \gg 0. \tag{5.11}$$

We view these basis elements as formal symbols, which we call *semi-infinite wedges*. (They are also sometimes called *semi-infinite monomials*.) We refer to $\mathcal{F}$ as the *semi-infinite wedge space* of V. The condition $i_{k+1} = i_k - 1$ for $k \gg 0$ means that there exists $N \in \mathbb{N}$ such that $i_{k+1} = i_k - 1$ for all $k \geq N$. We will also consider symbols as in (5.11), but without the condition $i_0 > i_1 > i_2 > \cdots$, using the convention

$$v_{i_0} \wedge \cdots \wedge v_{i_l} \wedge v_{i_{l+1}} \wedge v_{i_{l+2}} \wedge v_{i_{l+3}} \wedge \cdots = -v_{i_0} \wedge \cdots \wedge v_{i_l} \wedge v_{i_{l+2}} \wedge v_{i_{l+1}} \wedge v_{i_{l+3}} \wedge \cdots \tag{5.12}$$

(We allow $l = -1$ here, corresponding to interchanging the first two terms.) In other words, interchanging neighbouring terms introduces a minus sign. Note that this convention implies that

$$v_{i_0} \wedge v_{i_1} \wedge \cdots = 0 \quad \text{if} \quad i_k = i_l \text{ for some } k \neq l. \tag{5.13}$$

So a semi-infinite wedge with repeated terms is zero. Then any semi-infinite wedge

$$v_{i_0} \wedge v_{i_1} \wedge \cdots, \qquad i_{k+1} = i_k - 1 \text{ for } k \gg 0,$$

is equal to zero or a semi-infinite wedge of the form (5.11), up to sign, since we can reorder its terms.

In order to simplify notation, we define

$$v_{\mathbf{i}} := v_{i_0} \wedge v_{i_1} \wedge \cdots \quad \text{where} \quad \mathbf{i} = (i_0, i_1, \dots), \ i_{k+1} = i_k - 1 \text{ for } k \gg 0. \tag{5.14}$$

The condition $i_{k+1} = i_k - 1$ for $k \gg 0$ implies that there exists some $c \in \mathbb{Z}$ such that $i_k = c - k$ for $k \gg 0$. For $c \in \mathbb{Z}$, define

$$I^{(c)} := \{(i_0, i_1, \dots) : i_0 > i_1 > \cdots, \ i_k = c - k \text{ for } k \gg 0\}, \qquad I = \bigcup_{c \in \mathbb{Z}} I^{(c)}.$$

Then we have a decomposition of $\mathbb{Q}$-modules

$$\mathcal{F} = \bigoplus_{c \in \mathbb{Z}} \mathcal{F}^{(c)}, \qquad \mathcal{F}^{(c)} := \mathrm{Span}_{\mathbb{Q}} \left\{ v_{\mathbf{i}}, \ \mathbf{i} \in I^{(c)} \right\}. \tag{5.15}$$

Nonzero elements of $\mathcal{F}^{(c)}$ are said to be of *charge* c.

Lemma 5.3.2. *For $c \in \mathbb{Z}$, we have a bijection of sets*

$$\mathrm{Par} \to I^{(c)}, \qquad \lambda \mapsto (\lambda_1 + c, \lambda_2 + c - 1, \lambda_3 + c - 2, \dots).$$

Proof. The proof of this lemma is left as Exercise 5.3.1. □

For $c \in \mathbb{Z}$ and $\lambda \in \mathrm{Par}$, define

$$v_\lambda^{(c)} := v_{\lambda_1+c} \wedge v_{\lambda_2+c-1} \wedge v_{\lambda_3+c-2} \wedge \cdots .$$

It follows from Lemma 5.3.2 that, for $c \in \mathbb{Z}$,

$$\mathcal{F}^{(c)} \quad \text{has basis} \quad v_\lambda^{(c)},\ \lambda \in \mathrm{Par}. \tag{5.16}$$

Our goal in this section is to endow $\mathcal{F}$ with the structure of a simple Cl-module.

Proposition 5.3.3 (Fermionic Fock space)**.** *Semi-infinite wedge space $\mathcal{F}$ is a* Cl*-module under the action*

$$\begin{aligned} \psi_i^+ v_{\mathbf{i}} &= v_i \wedge v_{\mathbf{i}}, \\ \psi_i^- v_{\mathbf{i}} &= \begin{cases} (-1)^k v_{i_0} \wedge \cdots \wedge v_{i_{k-1}} \wedge v_{i_{k+1}} \wedge \cdots & \text{if } i_k = i, \\ 0 & \text{if } i_k \neq i \text{ for all } k \in \mathbb{N}, \end{cases} \end{aligned} \tag{5.17}$$

extended by linearity, for

$$\mathbf{i} = (i_0, i_1, \dots), \qquad i_0 > i_1 > i_2 > \cdots, \quad i_{k+1} = i_k - 1 \text{ for } k \gg 0.$$

Viewed as a Cl*-module with this action, $\mathcal{F}$ is called* fermionic Fock space*.*

Proof. To prove that $\mathcal{F}$ is a Cl-module with the given action, we must verify that the relations (5.10) are satisfied. We leave it as Exercise 5.3.2 to verify that the formulas (5.17) also hold when

$$\mathbf{i} = (i_0, i_1, \dots), \qquad i_{k+1} = i_k - 1 \text{ for } k \gg 0,$$

has no repeated terms. That is, it is not necessary to suppose that $i_0 > i_1 > i_2 > \cdots$.

Now, for all $i, j \in \mathbb{Z}$ and all semi-infinite wedges $v_{\mathbf{i}}$, $\mathbf{i} = (i_0, i_1, \dots,)$, we have

$$\psi_i^+ \psi_j^+ v_{\mathbf{i}} = v_i \wedge v_j \wedge v_{\mathbf{i}} \overset{(5.12)}{=} -v_j \wedge v_i \wedge v_{\mathbf{i}} = -\psi_j^+ \psi_i^+ v_{\mathbf{i}}.$$

Thus the first relation in (5.10) is satisfied. We leave it as Exercise 5.3.3 to verify that the second relation in (5.10) is satisfied.

It remains to consider the third relation in (5.10). First consider the case $i \neq j$. We have

$$\psi_i^+ \psi_j^- v_{\mathbf{i}} = \begin{cases} (-1)^k v_i \wedge v_{i_0} \wedge \cdots \wedge v_{i_{k-1}} \wedge v_{i_{k+1}} \wedge \cdots & \text{if } i_k = j, \\ 0 & \text{if } i_k \neq j \text{ for all } k \in \mathbb{N}, \end{cases}$$

and

$$\psi_j^- \psi_i^+ v_{\mathbf{i}} = \begin{cases} (-1)^{k+1} v_i \wedge v_{i_0} \wedge \cdots \wedge v_{i_{k-1}} \wedge v_{i_{k+1}} \wedge \cdots & \text{if } i_k = j, \\ 0 & \text{if } i_k \neq j \text{ for all } k \in \mathbb{N}. \end{cases}$$

Hence $\psi_i^+ \psi_j^- v_{\mathbf{i}} = -\psi_j^- \psi_i^+ v_{\mathbf{i}}$, as desired. Now consider the case $i = j$. We have

$$\psi_i^+ \psi_i^- v_{\mathbf{i}} = \begin{cases} v_{\mathbf{i}} & \text{if } i_k = i \text{ for some } k \in \mathbb{N}, \\ 0 & \text{if } i_k \neq i \text{ for all } k \in \mathbb{N}, \end{cases} \tag{5.18}$$

and

$$\psi_i^- \psi_i^+ v_{\mathbf{i}} = \begin{cases} 0 & \text{if } i_k = i \text{ for some } k \in \mathbb{N}, \\ v_{\mathbf{i}} & \text{if } i_k \neq i \text{ for all } k \in \mathbb{N}. \end{cases} \tag{5.19}$$

Hence $(\psi_i^+ \psi_i^- + \psi_i^- \psi_i^+) v_{\mathbf{i}} = v_{\mathbf{i}}$, as desired. □

Remark 5.3.4. The operators ψ_i^+ are called *wedging* or *creation* operators, and the operators ψ_i^- are called *contraction* or *annihilation* operators. A *fermion* is a type of subatomic particle with the property that at most one particle can occupy each state. This is sometimes referred to as the *Pauli exclusion principle*. If we view v_i as a particle in state i, then this property corresponds to (5.13). Compare to Remark 5.2.1. △

Lemma 5.3.5. *For all $i, c \in \mathbb{Z}$, we have*

$$\psi_i^+ \mathcal{F}^{(c)} \subseteq \mathcal{F}^{(c+1)}, \qquad \psi_i^- \mathcal{F}^{(c)} \subseteq \mathcal{F}^{(c-1)}.$$

Proof. We leave the proof of this lemma as Exercise 5.3.5. □

Proposition 5.3.6. *The elements*

$$\begin{gathered} \psi_{i_1}^- \cdots \psi_{i_r}^- \psi_{j_1}^+ \cdots \psi_{j_s}^+ \psi_{k_1}^- \cdots \psi_{k_t}^-, \\ r, s, t \geq 0, \quad 0 \geq i_1 > \cdots > i_r, \quad j_1 > \cdots > j_s, \quad k_1 > \cdots > k_t > 0, \end{gathered} \tag{5.20}$$

form a $\mathbb{Q}$-basis of Cl.

Proof. The defining relations (5.10) allow us to reorder the generators $\psi_i^\pm$ up to terms with fewer generators. It follows that the elements (5.20) span Cl. It remains to prove that they are linearly independent.

Since we may reorder the $\psi_i^\pm$ appearing in the elements (5.20), up to a sign and shorter expressions, it suffices to prove that the elements

$$\psi_{i_1}^+ \cdots \psi_{i_r}^+ \psi_{j_1}^- \cdots \psi_{j_s}^-, \qquad r, s \geq 0,\ i_1 > \cdots > i_r,\ j_1 > \cdots > j_s, \tag{5.21}$$

are linearly independent. (See Exercise 5.3.4.) Suppose, towards a contradiction, that we have a linear dependence relation amongst these elements. Let W denote the finite set of elements (5.21) appearing in this linear dependence relation with nonzero coefficient. So we can write this relation in the form

$$\sum_{w \in W} a_w w = 0, \qquad a_w \in \mathbb{Q}.$$

Choose $k \in \mathbb{Z}$ less than all the indices appearing in elements of W. Now choose an element

$$\psi_{i_1}^+ \cdots \psi_{i_r}^+ \psi_{j_1}^- \cdots \psi_{j_s}^- \in W$$

with s minimal, and define

$$v_{\mathbf{i}} = v_{j_1} \wedge \cdots \wedge v_{j_s} \wedge v_k \wedge v_{k-1} \wedge \cdots .$$

Then we leave it as Exercise 5.3.6 to show that

$$\sum_{w \in W} a_w w v_{\mathbf{i}} \neq 0, \tag{5.22}$$

which is a contradiction. □

Let J denote the $\mathbb{Q}$-span of the elements (5.20) for which

$$t > 0 \quad \text{or} \quad (j_a \leq 0 \text{ for some } 1 \leq a \leq s).$$

In other words, J is the $\mathbb{Q}$-span of the elements (5.20) that annihilate

$$v_\varnothing^{(0)} = v_0 \wedge v_{-1} \wedge v_{-2} \wedge \cdots \in \mathcal{F}.$$

Proposition 5.3.7. *The set J is a left ideal of* Cl*, and we have an isomorphism of* Cl*-modules*

$$\mathrm{Cl}/J \xrightarrow{\cong} \mathcal{F}, \quad a + J \mapsto a v_\varnothing^{(0)}. \tag{5.23}$$

Proof. The elements (5.20) not in J are those with $t = 0$ and $j_1 > \cdots > j_s > 0$. Such an element sends $v_\varnothing^{(0)}$ to $\pm$ the semi-infinite wedge obtained from $v_{j_1} \wedge \cdots \wedge v_{j_s} \wedge v_0 \wedge v_{-1} \wedge \cdots$ by removing the factors $v_{i_1}, \ldots, v_{i_r}$. Distinct such elements give distinct semi-infinite wedges, so their images are linearly independent. Combined with Proposition 5.3.6, this shows that $J = \mathrm{Ann}_{\mathrm{Cl}}\left(v_\varnothing^{(0)}\right)$. Thus J is the kernel of the Cl-module homomorphism

$$\mathrm{Cl} \to \mathcal{F}, \quad a \mapsto a v_\varnothing^{(0)}. \tag{5.24}$$

Hence $J = \mathrm{Ann}_{\mathrm{Cl}}(v_\varnothing^{(0)})$ is a left ideal of Cl. Since every element $v_{\mathbf{i}}$, $\mathbf{i} \in I$, can be obtained from $v_\varnothing^{(0)}$ by applying wedging and contracting operators, the map (5.24) is also surjective. Thus we have an isomorphism (5.23) by the first isomorphism theorem for modules. □

We would like an analogue of Theorem 5.2.3 for fermionic Fock space. We say that a nonzero element v in a Cl-module is a *vacuum vector* if

$$\psi_i^+ v = 0 \text{ for all } i \leq 0 \qquad \text{and} \qquad \psi_i^- v = 0 \text{ for all } i > 0. \tag{5.25}$$

For example, the vector $v_\varnothing^{(0)}$ is a vacuum vector.

Theorem 5.3.8. *Fermionic Fock space $\mathcal{F}$ is a* simple Cl*-module. Furthermore, any* Cl*-module generated by a vacuum vector v is isomorphic to fermionic Fock space via an isomorphism sending $v_\varnothing^{(0)}$ to v.*

Proof. Suppose U is a nonzero Cl-submodule of $\mathcal{F}$. We must show that $U = \mathcal{F}$. We first show that U contains a semi-infinite wedge. Let u be a nonzero element of U. Then, after multiplying by an element of $\mathbb{Q}$ if necessary, we have

$$u = v_{\mathbf{i}} + \sum_{\mathbf{j} \in X} c_{\mathbf{j}} v_{\mathbf{j}},$$

where $c_{\mathbf{j}} \in \mathbb{Q} \setminus \{0\}$, and the sum is over a finite set X not containing $\mathbf{i}$. We can choose u so that the cardinality of X is as small as possible. We aim to show that X is empty.

Suppose, towards a contradiction, that X is nonempty. Then choose $\mathbf{k} \in X$. Since $\mathbf{i} \neq \mathbf{k}$, we can choose $r \in \mathbb{Z}$ such that r appears in the sequence $\mathbf{i}$ but not in $\mathbf{k}$, or vice versa. If r appears in $\mathbf{i}$ but does not appear in $\mathbf{k}$, then, by (5.18), we have

$$\psi_r^+ \psi_r^- u = v_{\mathbf{i}} + \sum_{\mathbf{j} \in X'} c_{\mathbf{j}} v_{\mathbf{j}}$$

where X' is a subset of X not containing $\mathbf{k}$. In particular X' is *proper* subset of X, contradicting our choice of u. Similarly, if r appears in $\mathbf{k}$ but does not appear in $\mathbf{i}$, then, by (5.19), we have

$$\psi_r^- \psi_r^+ u = v_{\mathbf{i}} + \sum_{\mathbf{j} \in X'} c_{\mathbf{j}} v_{\mathbf{j}}$$

for a proper subset X' of X.

We now know that U contains some semi-infinite wedge $v_{\mathbf{i}}$. Then is easy to see that, using the operators $\psi_i^{\pm}$, $i \in \mathbb{Z}$, we can obtain all other semi-infinite wedges, up to sign. Thus U contains all semi-infinite wedges, and so $U = \mathcal{F}$, as desired.

Now suppose V is a Cl-module generated by a vacuum vector v, and consider the homomorphism of Cl-modules

$$\kappa \colon \mathrm{Cl} \to V, \quad a \mapsto av, \quad a \in \mathrm{Cl}.$$

The kernel of κ contains the left ideal J. Since κ is not the zero map and J is maximal, we must have $\ker(\kappa) = J$. On the other hand, since V is generated by v, the map κ is also surjective. Thus we have an isomorphism of Cl-modules

$$\mathrm{Cl}/J \xrightarrow{\cong} V, \quad a + J \mapsto av, \quad a \in \mathrm{Cl}.$$

Then it follows from (5.23) that we have an isomorphism of Cl-modules

$$\mathcal{F} \xrightarrow{\cong} V, \quad av_{\varnothing}^{(0)} \mapsto av, \quad a \in \mathrm{Cl}. \qquad \square$$

We conclude this section by defining a natural inner product on $\mathcal{F}$ that will correspond, under the boson-fermion correspondence, to the Hall inner product on the ring of symmetric functions. We define this inner product on $\mathcal{F}$ by

$$\langle\ ,\ \rangle \colon \mathcal{F} \times \mathcal{F} \to \mathbb{Q}, \qquad \langle v_{\mathbf{i}}, v_{\mathbf{j}} \rangle = \delta_{\mathbf{ij}}, \qquad \mathbf{i}, \mathbf{j} \in I, \tag{5.26}$$

extended by bilinearity.

Proposition 5.3.9. *With respect to the inner product (5.26), we have* $(\psi_i^{\pm})^{\perp} = \psi_i^{\mp}$ *for all* $i \in \mathbb{Z}$. *In other words,*

$$\langle \psi_i^{\pm} u, v \rangle = \langle u, \psi_i^{\mp} v \rangle \qquad \text{for all } i \in \mathbb{Z},\ u, v \in \mathcal{F}. \tag{5.27}$$

Proof. We leave the proof of this result as Exercise 5.3.7. $\square$

Exercises.

5.3.1. Prove Lemma 5.3.2.

5.3.2. Show that the formulas (5.17) also hold when

$$\mathbf{i} = (i_0, i_1, \dots), \qquad i_{k+1} = i_k - 1 \text{ for } k \gg 0,$$

has no repeated terms. That is, it is not necessary to suppose that $i_0 > i_1 > i_2 > \cdots$.

5.3.3. Show that the action defined in Proposition 5.3.3 satisfies the second relation in (5.10).

5.3.4. Let $\succeq_1$ and $\succeq_2$ be two total orders on the set

$$Y = \{\psi_i^{\pm} : i \in \mathbb{Z}\}.$$

Then, for $k \in \{1, 2\}$, let

$$Y_k = \{z_1 z_2 \dots z_l : l \geq 0,\ z_1 \succ_k z_2 \succ \cdots \succ_k z_l\} \subseteq \mathrm{Cl}.$$

Prove that Y_1 is linearly independent if and only if Y_2 is linearly independent. *Hint*: The relations (5.10) allow you to reorder products $z_1 z_2 \dots z_l$ up to shorter products.

5.3.5. Prove Lemma 5.3.5.

5.3.6. Prove (5.22).

5.3.7. Prove Proposition 5.3.9.

5.4 Bosonization

We are now ready to present the first half of the boson-fermion correspondence, known as *bosonization*. This consists of defining an action of the Heisenberg algebra on fermionic Fock space, and identifying the resulting Heis-module with a sum of countably many copies of bosonic Fock space.

For $r \in \mathbb{Z}$, $r \neq 0$, consider the infinite sum

$$P_r := \sum_{j \in \mathbb{Z}} \psi_{j+r}^{+} \psi_j^{-} \tag{5.28}$$

of products of elements of Cl. Since the sum is infinite, this is *not* an element of Cl. However, suppose

$$\mathbf{i} = (i_0, i_1, \dots), \quad i_0 > i_1 > \cdots, \quad i_{k+1} = i_k - 1 \text{ for } k \geq N.$$

Then (5.13) implies that

$$P_r v_{\mathbf{i}} = \begin{cases} \sum_{k=0}^{N} v_{i_0} \wedge v_{i_1} \wedge \cdots \wedge v_{i_{k-1}} \wedge v_{i_k + r} \wedge v_{i_{k+1}} \wedge \cdots & \text{if } r < 0, \\ \sum_{k=0}^{N+r+1} v_{i_0} \wedge v_{i_1} \wedge \cdots \wedge v_{i_{k-1}} \wedge v_{i_k + r} \wedge v_{i_{k+1}} \wedge \cdots & \text{if } r > 0. \end{cases}$$

In particular, P_r *is* well-defined as an element of $\mathrm{End}_{\mathbb{Q}}(\mathcal{F})$ since all but finitely many of the terms $\psi_{j+r}^{+}\psi_j^{-}$ act as zero on any given $v_{\mathbf{i}}$. It follows immediately from Lemma 5.3.5 that

$$P_r\mathcal{F}^{(c)} \subseteq \mathcal{F}^{(c)} \qquad \text{for all } c, r \in \mathbb{Z},\ r \neq 0. \tag{5.29}$$

Recall that if U and V are vector spaces equipped with inner products, then a linear map $f\colon U \to V$ is an *isometry* if

$$\langle f(u_1), f(u_2)\rangle = \langle u_1, u_2\rangle \qquad \text{for all } u_1, u_2 \in U.$$

If f is also a vector space isomorphism, then we say that f is an *isometric isomorphism.*

Theorem 5.4.1 (Bosonization)**.** *For all $c \in \mathbb{Z}$, the map*

$$\mathbf{A}^{(c)}\colon \mathcal{F}^{(c)} \to \mathrm{Sym}_{\mathbb{Q}}, \quad v_\lambda^{(c)} \mapsto s_\lambda, \tag{5.30}$$

is an isometric isomorphism. Furthermore, for all $r \geq 1$, the diagram

$$\begin{array}{ccc} \mathcal{F}^{(c)} & \xrightarrow{P_{\pm r}} & \mathcal{F}^{(c)} \\ {\scriptstyle \mathbf{A}^{(c)}}\downarrow & & \downarrow{\scriptstyle \mathbf{A}^{(c)}} \\ \mathrm{Sym}_{\mathbb{Q}} & \xrightarrow{p_r^{\pm}} & \mathrm{Sym}_{\mathbb{Q}} \end{array} \tag{5.31}$$

commutes. In particular,

$$\mathrm{Heis}_{\mathbb{Q}} \to \mathrm{End}_{\mathbb{Q}}\left(\mathcal{F}^{(c)}\right), \quad p_r^{\pm} \mapsto P_{\pm r}, \quad r \geq 1,$$

is a representation of $\mathrm{Heis}_{\mathbb{Q}}$ *and, under this action, $\mathcal{F}^{(c)}$ is isomorphic to bosonic Fock space.*

Proof. The fact that $\mathbf{A}^{(c)}$ is an isomorphism of vector spaces follows immediately from (5.16) and the fact that the Schur functions are a basis of $\mathrm{Sym}_{\mathbb{Q}}$. It is an isometric isomorphism by (3.30) and (5.26). Thus, it remains to prove that (5.31) commutes. To do this, we compare the action of P_r on $v_\lambda^{(n)}$ to the Murnaghan–Nakayama rule (Proposition 3.7.1).

Fix $j, r \in \mathbb{Z}$, $r \geq 1$, and $\lambda \in \mathrm{Par}$. Let

$$\mathbf{i} = (\lambda_1 + c, \lambda_2 + c - 1, \lambda_3 + c - 2, \dots)$$

be the element of $I^{(c)}$ corresponding to λ under the bijection of Lemma 5.3.2, so that $v_\lambda^{(c)} = v_{\mathbf{i}}$. If j does not appear in the sequence $\mathbf{i}$, then $\psi_{j+r}^{+}\psi_j^{-}v_{\mathbf{i}} = 0$. Otherwise, $\psi_{j+r}^{+}\psi_j^{-}v_{\mathbf{i}}$ is obtained from $v_{\mathbf{i}}$ by replacing the entry j by $j + r$. Then, repeating the argument in the proof of Proposition 3.7.1 shows that

$$\sum_{j\in\mathbb{Z}} \psi_{j+r}^{+}\psi_j^{-}v_{\mathbf{i}} = \sum_{\mu} (-1)^{\mathrm{ht}(\mu/\lambda)} v_\mu^{(c)},$$

where the sum is over all partitions $\mu \supseteq \lambda$ such that μ/λ is a border strip with r boxes. This proves that (5.31) commutes for P_r and p_r^{+}.

It remains to prove that (5.31) commutes for P_{-r} and p_r^-. For all $\lambda, \mu \in \mathrm{Par}$, we have

$$\begin{aligned}
\left\langle \mathbf{A}^{(c)}\left(P_{-r}v_\lambda^{(c)}\right), s_\mu\right\rangle &= \left\langle \mathbf{A}^{(c)}\left(P_{-r}v_\lambda^{(c)}\right), \mathbf{A}^{(c)}\left(v_\mu^{(c)}\right)\right\rangle \\
&= \left\langle P_{-r}v_\lambda^{(c)}, v_\mu^{(c)}\right\rangle && \text{(since } \mathbf{A}^{(c)} \text{ is an isometry)} \\
&\overset{(5.28)}{=} \sum_{j\in\mathbb{Z}} \left\langle \psi_{j-r}^+\psi_j^- v_\lambda^{(c)}, v_\mu^{(c)}\right\rangle \\
&\overset{(5.27)}{=} \sum_{j\in\mathbb{Z}} \left\langle v_\lambda^{(c)}, \psi_j^+\psi_{j-r}^- v_\mu^{(c)}\right\rangle \\
&\overset{(5.28)}{=} \left\langle v_\lambda^{(c)}, P_r v_\mu^{(c)}\right\rangle \\
&= \left\langle \mathbf{A}^{(c)}\left(v_\lambda^{(c)}\right), \mathbf{A}^{(c)}\left(P_r v_\mu^{(c)}\right)\right\rangle && \text{(since } \mathbf{A}^{(c)} \text{ is an isometry)} \\
&= \left\langle \mathbf{A}^{(c)}\left(v_\lambda^{(c)}\right), p_r^+\mathbf{A}^{(c)}\left(v_\mu^{(c)}\right)\right\rangle \\
&= \left\langle p_r^-\mathbf{A}^{(c)}\left(v_\lambda^{(c)}\right), s_\mu\right\rangle && \text{(since } p_r^- = p_r^\perp\text{)}.
\end{aligned}$$

Since the s_μ, $\mu \in \mathrm{Par}$, span Sym, it follows that

$$\mathbf{A}^{(c)}\left(P_{-r}v_\lambda^{(c)}\right) = p_r^-\mathbf{A}^{(c)}\left(v_\lambda^{(c)}\right) \qquad \text{for all } \lambda \in \mathrm{Par}.$$

Thus $\mathbf{A}^{(c)}P_{-r} = p_r^-\mathbf{A}^{(c)}$ by (5.16). □

It follows from Theorem 5.4.1 that $\mathcal{F} = \bigoplus_{c\in\mathbb{Z}} \mathcal{F}^{(c)}$ is isomorphic to a direct sum of countably many copies of bosonic Fock space. Let $\mathbb{Q}[q, q^{-1}]$ denote the ring of Laurent polynomials in q with rational coefficients. Then

$$\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]} := \mathbb{Q}[q, q^{-1}] \otimes_{\mathbb{Q}} \mathrm{Sym}_{\mathbb{Q}}$$

can be identified with the ring of symmetric functions with coefficients in $\mathbb{Q}[q, q^{-1}]$ and we have

$$\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]} = \bigoplus_{c\in\mathbb{Z}} q^c \mathrm{Sym}_{\mathbb{Q}}.$$

We can extend the action of $\mathrm{Heis}_{\mathbb{Q}}$ on $\mathrm{Sym}_{\mathbb{Q}}$ to an action on $\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$ by

$$x(av) = axv, \qquad x \in \mathrm{Heis}_{\mathbb{Q}},\ a \in \mathbb{Q}[q, q^{-1}],\ v \in \mathrm{Sym}_{\mathbb{Q}}.$$

Furthermore, we extend the Hall inner product to $\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$ by

$$\langle q^c f, q^d g\rangle = \delta_{c,d}\langle f, g\rangle, \qquad c, d \in \mathbb{Z},\ f, g \in \mathrm{Sym}_{\mathbb{Q}}.$$

Then we have the following corollary of Theorem 5.4.1.

Corollary 5.4.2. *We have an isometric isomorphism of* $\mathrm{Heis}_{\mathbb{Q}}$*-modules*

$$\mathbf{A}\colon \mathcal{F} \xrightarrow{\cong} \mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}, \quad v_\lambda^{(c)} \mapsto q^c s_\lambda, \qquad c \in \mathbb{Z},\ \lambda \in \mathrm{Par}. \tag{5.32}$$

Exercises.

5.4.1. Consider (5.28) with $r = 0$. Is P_0 a well-defined operator on $\mathcal{F}$?

5.4.2. It follows from Theorem 5.4.1 and (5.6) that $P_r P_{-r} = P_{-r} P_r - r$ for $r \geq 1$. So what is wrong with the following computation?

$$\begin{aligned}
P_r P_{-r} &= \sum_{j,k\in\mathbb{Z}} \psi^+_{j+r}\psi^-_j \psi^+_{k-r}\psi^-_k \\
&\overset{(5.10)}{=} -\sum_{j,k\in\mathbb{Z}} \psi^+_{j+r}\psi^+_{k-r}\psi^-_j \psi^-_k + \sum_{j,k\in\mathbb{Z}} \delta_{j,k-r}\psi^+_{j+r}\psi^-_k \\
&\overset{(5.10)}{=} -\sum_{j,k\in\mathbb{Z}} \psi^+_{k-r}\psi^+_{j+r}\psi^-_k \psi^-_j + \sum_{j\in\mathbb{Z}} \psi^+_{j+r}\psi^-_{j+r} \\
&\overset{(5.10)}{=} \sum_{j,k\in\mathbb{Z}} \psi^+_{k-r}\psi^-_k \psi^+_{j+r}\psi^-_j - \sum_{j,k\in\mathbb{Z}} \delta_{k,j+r}\psi^+_{k-r}\psi^-_j + \sum_{j\in\mathbb{Z}} \psi^+_{j+r}\psi^-_{j+r} \\
&= P_{-r}P_r - \sum_{j\in\mathbb{Z}} \psi^+_j\psi^-_j + \sum_{j\in\mathbb{Z}} \psi^+_{j+r}\psi^-_{j+r} \\
&= P_{-r}P_r.
\end{aligned}$$

5.4.3. It follows from Theorem 5.4.1 and (5.6) that $P_r P_s = P_s P_r$ for $r, s \in \mathbb{Z} \setminus \{0\}$, $r \neq s$. Prove this relation directly using (5.10) and (5.28).

5.4.4. Suppose U and V are $\mathbb{Q}$-vector spaces equipped with inner products. (By definition, the inner product is bilinear, symmetric, and positive definite.) Prove that every isometry $f\colon U \to V$ is injective.

5.5 Fermionization

In this section we present the second half of the boson-fermion correspondence, known as *fermionization*. This consists of defining an action of the Clifford algebra on $\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$ (which is a countable direct sum of copies of bosonic Fock space), and showing that the isometric isomorphism $\mathbf{A}$ of (5.32) is also an isomorphism of Cl-modules. Much of our presentation follows [Mac15, p. 95, Exercise I.5.29].

Recall the generating functions $E(t)$ and $H(t)$ from (2.21) and (2.28), and define the formal power series

$$E^+(t) = \sum_{n\geq 0} e^+_n t^n, \quad E^-(t) = \sum_{n\geq 0} e^-_n t^n, \quad H^+(t) = \sum_{n\geq 0} h^+_n t^n, \quad H^-(t) = \sum_{n\geq 0} h^-_n t^n. \tag{5.33}$$

These are all elements of $\mathrm{End}(\mathrm{Sym})[\![t]\!]$. They can also be viewed as maps $\mathrm{Sym} \to \mathrm{Sym}[\![t]\!]$.

Lemma 5.5.1. *Viewing $H^{\pm}(t)$ and $E^{\pm}(t)$ as maps* $\mathrm{Sym} \to \mathrm{Sym}[\![t]\!]$*, we have*

$$H^{\pm}(t) \circ \omega = \omega \circ E^{\pm}(t), \tag{5.34}$$

where, on the right-hand side, ω is the involution of $\mathrm{Sym}[\![t]\!]$ *that applies ω (from (2.31)) to each coefficient.*

Proof. For all $f \in \mathrm{Sym}$, we have

$$H^{+}(t) \circ \omega(f) = \sum_{n \geq 0} h_n \omega(f) t^n \overset{(2.31)}{=} \sum_{n \geq 0} \omega\big(e_n f\big) t^n = \omega \circ E^{+}(t)(f)$$

and

$$H^{-}(t) \circ \omega(f) = \sum_{n \geq 0} h_n^{\perp}(\omega(f)) t^n \overset{(4.16)}{=} \sum_{n \geq 0} \omega\left(\omega(h_n)^{\perp} f\right) t^n \overset{(2.32)}{=} \sum_{n \geq 0} \omega\left(e_n^{\perp} f\right) t^n = \omega \circ E^{-}(t) f. \qquad \square$$

Lemma 5.5.2. *The formal power series $E^{-}(t)$ and $H^{-}(t)$ are ring homomorphisms* $\mathrm{Sym} \to \mathrm{Sym}[t]$.

Proof. For any $f \in \mathrm{Sym}$, we have $e_n^{-} f = e_n^{\perp} f = 0$ for all but finitely many n. Thus, $E^{-}(t) f \in \mathrm{Sym}[t]$. Similarly, $H^{-}(t)$ is a map from Sym to $\mathrm{Sym}[t]$. It is clear that they are $\Bbbk$-linear maps and that they send 1 to 1. So it remains to check that they respect multiplication. Indeed, for $f, g \in \mathrm{Sym}$, we have

$$\begin{aligned} E^{-}(t)(fg) &= \sum_{n \geq 0} e_n^{\perp}(fg) t^n \\ &\overset{\substack{(4.34)\\(4.31)}}{=} \sum_{n \geq 0} \sum_{r=0}^{n} (e_r^{\perp} f)(e_{n-r}^{\perp} g) t^n \\ &= \left(\sum_{n \geq 0} \left(e_n^{\perp} f\right) t^n \right) \left(\sum_{m \geq 0} \left(e_m^{\perp} g\right) t^m \right) \\ &= \left(E^{-}(t) f\right) \left(E^{-}(t) g\right). \end{aligned}$$

Hence $E^{-}(t)$ is a ring homomorphism.

Now, by (5.34), we have

$$H^{-}(t) = \omega \circ E^{-}(t) \circ \omega.$$

Since ω is a ring homomorphism, it follows that $H^{-}(t)$ is also a ring homomorphism. $\square$

In what follows, we consider two indeterminates t and u. Recall from (1.11) that

$$(1 - tu)^{-1} = \sum_{n \geq 0} t^n u^n.$$

Proposition 5.5.3. *We have*

$$H^{-}(t) H^{+}(u) = (1 - tu)^{-1} H^{+}(u) H^{-}(t), \tag{5.35}$$

$$H^{-}(t) E^{+}(u) = (1 + tu) E^{+}(u) H^{-}(t), \tag{5.36}$$

$$E^{-}(t) E^{+}(u) = (1 - tu)^{-1} E^{+}(u) E^{-}(t), \tag{5.37}$$

$$E^{-}(t) H^{+}(u) = (1 + tu) H^{+}(u) E^{-}(t). \tag{5.38}$$

Proof. We first compute

$$H^-(t)H(u) = \sum_{m,n\geq 0} \left(h_m^\perp h_n\right) t^m u^n \overset{(4.36)}{=} \sum_{m,n\geq 0} (h_{n-m}u^{n-m})(tu)^m = (1-tu)^{-1}H(u).$$

Since $H^-(t)$ is a ring homomorphism by Lemma 5.5.2, we have, for $f \in \mathrm{Sym}$,

$$H^-(t)H^+(u)f = H^-(t)\left(H(u)f\right) = \left(H^-(t)H(u)\right)\left(H^-(t)f\right) = (1-tu)^{-1}H(u)H^-(t)f.$$

Thus (5.35) is satisfied.

Next we compute

$$H^-(t)E(u) = \sum_{m,n\geq 0} \left(h_m^\perp e_n\right) t^m u^n \overset{(4.36)}{=} \sum_{n\geq 0} (e_n + te_{n-1})u^n = (1+tu)E(u).$$

Since $H^-(t)$ is a ring homomorphism by Lemma 5.5.2, we have, for $f \in \mathrm{Sym}$,

$$H^-(t)E^+(u)f = H^-(t)\left(E(u)f\right) = \left(H^-(t)E(u)\right)\left(H^-(t)f\right) = (1+tu)E(u)H^-(t)f.$$

Thus (5.36) is satisfied.

Next, we compose (5.35) on the left and right with ω to obtain

$$\begin{aligned} &\omega \circ H^-(t)H^+(u) \circ \omega = (1-tu)^{-1}\omega \circ H^+(u)H^-(t) \circ \omega \\ \Longrightarrow\ &\omega \circ H^-(t) \circ \omega^2 \circ H^+(u) \circ \omega = (1-tu)^{-1}\omega \circ H^+(u) \circ \omega^2 \circ H^-(t) \circ \omega \\ \overset{(5.34)}{\Longrightarrow}\ &E^-(t)E^+(u) = (1-tu)^{-1}E^+(u)E^-(t). \end{aligned}$$

Relation (5.38) can be similarly deduced from (5.36). □

Remark 5.5.4. Equating coefficients of $t^m u^n$, relations (5.35) to (5.38) correspond to Exercises 5.1.2 and 5.1.3. △

Next we define

$$B^+(t) = \sum_{n\in\mathbb{Z}} B_n^+ t^n = H^+(t)E^-(-t^{-1}), \qquad B^-(t) = \sum_{n\in\mathbb{Z}} B_n^- t^n = E^+(-t^{-1})H^-(t). \tag{5.39}$$

Thus

$$B_n^+ = \sum_{r\geq 0} (-1)^r h_{n+r}^+ e_r^-, \qquad B_n^- = \sum_{r\geq 0} (-1)^r e_r^+ h_{n+r}^-. \tag{5.40}$$

The operators $B_n^\pm$ are called *Bernstein operators*. Note that they are *not* elements of Heis since the sums are infinite. However, they *are* elements of $\mathrm{End}_{\mathbb{Q}}(\mathrm{Sym}_{\mathbb{Q}})$ since, when they act on any element of $\mathrm{Sym}_{\mathbb{Q}}$, all but finitely many terms act as zero. In this sense, they are similar to the operators P_r from (5.28).

Note also that $B^\pm(t)$ are not formal power series, since we allow negative exponents. Thus, when multiplying them by other series, we must verify that the product is well-defined. When we write an equation involving such a product, we are implicitly asserting that these products are well defined.

If $t_1, t_2, \ldots, t_n$ are independent variables, define

$$B^+(t_1, \ldots, t_n) = H^+(t_1)\cdots H^+(t_n)E^-(-t_1^{-1})\cdots E^-(-t_n^{-1}).$$

When $n = 1$, this reduces to the $B^+(t)$ from (5.39).

Lemma 5.5.5. *For $n \geq 1$,*

$$B^+(t_1)\cdots B^+(t_n) = \prod_{1\leq i<j\leq n}\left(1-t_i^{-1}t_j\right) B^+(t_1,\ldots,t_n).$$

(When $n=1$, we have $\prod_{1\leq i<j\leq n}\left(1-t_i^{-1}t_j\right)=1$ by convention.)

Proof. This follows from repeated use of (5.38). We leave the details as Exercise 5.5.2. □

Corollary 5.5.6. *For $n \geq 1$,*

$$B^+(t_1)\cdots B^+(t_n)1 = \prod_{1\leq i<j\leq n}\left(1-t_i^{-1}t_j\right) H(t_1)\cdots H(t_n). \tag{5.41}$$

Proof. This follows immediately from Lemma 5.5.5 and the fact that $E^-(t)1=1$. □

The next result shows that the Bernstein operators are creation operators for Schur functions.[1]

Proposition 5.5.7. *For $\lambda \in \mathrm{Par}_n$, we have*

$$B^+_{\lambda_1}B^+_{\lambda_2}\cdots B^+_{\lambda_n}1 = s_\lambda. \tag{5.42}$$

Proof. We have

$$\begin{aligned}
B^+(t_1)\cdots B^+(t_n)1 &\overset{(5.41)}{=} \prod_{1\leq i<j\leq n}(1-t_i^{-1}t_j)H(t_1)\cdots H(t_n)\\
&= t^{-\delta^n}\prod_{1\leq i<j\leq n}(t_i-t_j)H(t_1)\cdots H(t_n)\\
&\overset{(3.5)}{\underset{(3.2)}{=}} t^{-\delta^n}\sum_{\pi\in\mathfrak{S}_n}\mathrm{sgn}(\pi)t^{(\delta^n)^\pi}H(t_1)\cdots H(t_n)\\
&= \sum_{\pi\in\mathfrak{S}_n}\mathrm{sgn}(\pi)t^{(\delta^n)^\pi-\delta^n}H(t_1)\cdots H(t_n).
\end{aligned}$$

Comparing coefficients of $t_1^{\lambda_1}\cdots t_n^{\lambda_n}$, we obtain

$$\begin{aligned}
B^+_{\lambda_1}\cdots B^+_{\lambda_n}1 &= \sum_{\pi\in\mathfrak{S}_n}\mathrm{sgn}(\pi)\left(\text{coefficient of } t_1^{\lambda_1}\cdots t_n^{\lambda_n} \text{ in } t^{(\delta^n)^\pi-\delta^n}H(t_1)\cdots H(t_n)\right)\\
&= \sum_{\pi\in\mathfrak{S}_n}\mathrm{sgn}(\pi)\left(\text{coefficient of } \prod_{i=1}^n t_i^{\lambda_i-(n-\pi(i))+(n-i)} \text{ in } H(t_1)\cdots H(t_n)\right)\\
&= \sum_{\pi\in\mathfrak{S}_n}\mathrm{sgn}(\pi)\left(\text{coefficient of } \prod_{i=1}^n t_i^{\lambda_i-i+\pi(i)} \text{ in } H(t_1)\cdots H(t_n)\right)\\
&= \sum_{\pi\in\mathfrak{S}_n}\mathrm{sgn}(\pi)\prod_{i=1}^n h_{\lambda_i-i+\pi(i)}\\
&\overset{(3.2)}{=} \det(h_{\lambda_i-i+j})_{1\leq i,j\leq n}\\
&\overset{(3.10)}{=} s_\lambda
\end{aligned}$$

□

[1] We thank Darij Grinberg for providing the proof given here.

Define

$$\varphi(tu) = \sum_{n\in\mathbb{Z}} t^n u^n. \tag{5.43}$$

Note that, like $B^{\pm}(t)$, $\varphi(tu)$ is not a formal power series. So we need to be careful when multiplying it by other series.

Lemma 5.5.8. *We have*

$$H(t)E(-u^{-1})\varphi(tu) = \varphi(tu). \tag{5.44}$$

Proof. We have

$$H(t)E(-u^{-1})\varphi(tu) = \sum_{m,n\geq 0}\sum_{r\in\mathbb{Z}} (-1)^n h_m e_n t^{m+r} u^{r-n}.$$

Now, for $k, l \in \mathbb{Z}$, the coefficient of $t^k u^l$ in the above sum is

$$\sum_{r\in\mathbb{Z}} (-1)^{r-l} h_{k-r} e_{r-l} = \sum_{r\in\mathbb{Z}} (-1)^r h_{k-l-r} e_r \overset{(2.30)}{=} \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{otherwise.} \end{cases}$$

This shows that the left-hand side of (5.44) is well-defined and equal to the right-hand side. □

Proposition 5.5.9. *We have*

$$B^+(t)B^+(u) + t^{-1}uB^+(u)B^+(t) = 0, \tag{5.45}$$

$$B^-(t)B^-(u) + tu^{-1}B^-(u)B^-(t) = 0, \tag{5.46}$$

$$B^+(t)B^-(u) + tuB^-(u)B^+(t) = \varphi(tu). \tag{5.47}$$

Proof. To prove (5.45), we compute

$$\begin{aligned} B^+(t)B^+(u) &= H^+(t)E^-(-t^{-1})H^+(u)E^-(-u^{-1}) \\ &\overset{(5.38)}{=} \left(1 - t^{-1}u\right) H^+(t)H^+(u)E^-(-t^{-1})E^-(-u^{-1}) \\ &= \left(1 - t^{-1}u\right) H^+(u)H^+(t)E^-(-u^{-1})E^-(-t^{-1}) \\ &= -t^{-1}u\left(1 - tu^{-1}\right) H^+(u)H^+(t)E^-(-u^{-1})E^-(-t^{-1}) \\ &\overset{(5.38)}{=} -t^{-1}uH^+(u)E^-(-u^{-1})H^+(t)E^-(-t^{-1}) \\ &= -t^{-1}uB^+(u)B^+(t). \end{aligned}$$

The proof of (5.46) is similar, using (5.36). It is left as Exercise 5.5.4.

It remains to prove (5.47). We have

$$\begin{aligned} B^+(t)B^-(u) &= H^+(t)E^-(-t^{-1})E^+(-u^{-1})H^-(u) \\ &\overset{(5.37)}{=} \left(1 - t^{-1}u^{-1}\right)^{-1} H^+(t)E^+(-u^{-1})E^-(-t^{-1})H^-(u). \end{aligned}$$

On the other hand

$$B^-(u)B^+(t) = E^+(-u^{-1})H^-(u)H^+(t)E^-(-t^{-1})$$
$$\overset{(5.35)}{=} (1-tu)^{-1}E^+(-u^{-1})H^+(t)H^-(u)E^-(-t^{-1}).$$

Since

$$(1-t^{-1}u^{-1})^{-1} + tu(1-tu)^{-1} = \sum_{n\leq 0} t^n u^n + tu\sum_{n\geq 0} t^n u^n = \sum_{n\in\mathbb{Z}} t^n u^n = \varphi(tu),$$

we have

$$B^+(t)B^-(u) + tuB^-(u)B^+(t) = \varphi(tu)E^+(-u^{-1})H^+(t)H^-(u)E^-(-t^{-1})$$
$$\overset{(5.44)}{=} \varphi(tu)H^-(u)E^-(-t^{-1})$$
$$\overset{(5.44)}{=} \varphi(tu). \qquad \square$$

Corollary 5.5.10. *For all $r, s \in \mathbb{Z}$, we have*

$$B_r^+ B_s^+ + B_{s-1}^+ B_{r+1}^+ = 0, \tag{5.48}$$
$$B_r^- B_s^- + B_{s+1}^- B_{r-1}^- = 0, \tag{5.49}$$
$$B_r^+ B_s^- + B_{s-1}^- B_{r-1}^+ = \delta_{rs}. \tag{5.50}$$

Proof. These relations follow from comparing the coefficients of $t^r u^s$ in (5.45) to (5.47). $\square$

The relations (5.48) to (5.50) resemble the defining relations (5.10) of the Clifford algebra. However, the relations (5.48) to (5.50) involve some shifting of indices that does not occur in (5.10). Thus, in order to obtain operators giving an action of the Clifford algebra, it is necessary to shift the Bernstein operators.

Let T denote the map that replaces the indeterminate q by tq. So T^{-1} is the map that replaces q by $t^{-1}q$. Now define

$$\Psi^+(t) = \sum_{i\in\mathbb{Z}} \Psi_i^+ t^i = TqB^+(t) \qquad \text{and} \qquad \Psi^-(t) = \sum_{i\in\mathbb{Z}} \Psi_i^- t^{-i} = q^{-1}T^{-1}B^-(t^{-1}). \tag{5.51}$$

Restricting to $q^c\mathrm{Sym}_{\mathbb{Q}}$ and comparing powers of t, it follows that

$$\Psi_i^+|_{q^c\mathrm{Sym}_{\mathbb{Q}}} = qB_{i-c-1}^+|_{q^c\mathrm{Sym}_{\mathbb{Q}}} \qquad \text{and} \qquad \Psi_i^-|_{q^c\mathrm{Sym}_{\mathbb{Q}}} = q^{-1}B_{i-c}^-|_{q^c\mathrm{Sym}_{\mathbb{Q}}}. \tag{5.52}$$

Proposition 5.5.11. *For all $i, j \in \mathbb{Z}$, we have*

$$\Psi_i^+\Psi_j^+ + \Psi_j^+\Psi_i^+ = 0, \qquad \Psi_i^-\Psi_j^- + \Psi_j^-\Psi_i^- = 0, \qquad \Psi_i^+\Psi_j^- + \Psi_j^-\Psi_i^+ = \delta_{ij}. \tag{5.53}$$

Proof. Suppose $c \in \mathbb{Z}$ and $\lambda \in \mathrm{Par}$. We have, using (5.52),

$$\Psi_i^+\Psi_j^+(q^c s_\lambda) = \Psi_i^+\left(q^{c+1}B_{j-c-1}^+ s_\lambda\right) = q^{c+2}B_{i-c-2}^+ B_{j-c-1}^+ s_\lambda,$$
$$\Psi_j^+\Psi_i^+(q^c s_\lambda) = \Psi_j^+\left(q^{c+1}B_{i-c-1}^+ s_\lambda\right) = q^{c+2}B_{j-c-2}^+ B_{i-c-1}^+ s_\lambda.$$

Since the elements $q^c s_\lambda$ span $\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$, the first equation in (5.53) follows from (5.48). Similarly,

$$\Psi_i^-\Psi_j^-(q^c s_\lambda) = \Psi_i^-\left(q^{c-1}B_{j-c}^- s_\lambda\right) = q^{c-2}B_{i-c+1}^- B_{j-c}^- s_\lambda,$$
$$\Psi_j^-\Psi_i^-(q^c s_\lambda) = \Psi_j^-\left(q^{c-1}B_{i-c}^- s_\lambda\right) = q^{c-2}B_{j-c+1}^- B_{i-c}^- s_\lambda,$$

and so the second relation in (5.53) follows from (5.49). Finally, we have

$$\Psi_i^+\Psi_j^-(q^c s_\lambda) = \Psi_i^+\left(q^{c-1}B_{j-c}^- s_\lambda\right) = q^c B_{i-c}^+ B_{j-c}^- s_\lambda,$$
$$\Psi_j^-\Psi_i^+(q^c s_\lambda) = \Psi_j^-\left(q^{c+1}B_{i-c-1}^+ s_\lambda\right) = q^c B_{j-c-1}^- B_{i-c-1}^+ s_\lambda,$$

and so the last relation in (5.53) follows from (5.50). □

We are now ready to prove the main result of this section.

Theorem 5.5.12 (Fermionization)**.** *The map*

$$\mathrm{Cl} \to \mathrm{End}_{\mathbb{Q}}\left(\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}\right), \quad \psi_i^\pm \mapsto \Psi_i^\pm, \quad i \in \mathbb{Z}, \tag{5.54}$$

is a representation of Cl *and, under this action,* $\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$ *is isomorphic to fermionic Fock space. Furthermore, for all* $i \in \mathbb{Z}$*, the diagram*

$$\begin{array}{ccc} \mathcal{F} & \xrightarrow{\psi_i^\pm} & \mathcal{F} \\ {\scriptstyle\mathbf{A}}\downarrow & & \downarrow{\scriptstyle\mathbf{A}} \\ \mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]} & \xrightarrow{\Psi_i^\pm} & \mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]} \end{array} \tag{5.55}$$

commutes, where $\mathbf{A}$ *is the map from (5.32).*

Proof. It follows from Proposition 5.5.11 that (5.54) is a representation of Cl. We would like to use Theorem 5.3.8 to show that, under this action, $\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$ is isomorphic to fermionic Fock space. We will do this by showing that $1 \in \mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$ is a vacuum vector that generates $\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$.

First note that, for $i \leq 0$,

$$\Psi_i^+ 1 \overset{(5.52)}{=} qB_{i-1}^+ 1 \overset{(5.40)}{=} q\sum_{r\geq 0}(-1)^r h_{r+i-1}^+ e_r^- 1 = qh_{i-1}^+ 1 = 0.$$

Similarly, for $i > 0$,

$$\Psi_i^- 1 \overset{(5.52)}{=} q^{-1}B_i^- 1 \overset{(5.40)}{=} q^{-1}\sum_{r\geq 0}(-1)^r e_r^+ h_{i+r}^- 1 = 0.$$

Thus 1 is a vacuum vector.

For any $c \in \mathbb{Z}$, define

$$\Phi_c = \begin{cases} \Psi_c^+ \cdots \Psi_2^+\Psi_1^+ & \text{if } c > 0, \\ 1 & \text{if } c = 0, \\ \Psi_{c+1}^- \cdots \Psi_{-1}^-\Psi_0^- & \text{if } c < 0, \end{cases} \quad \text{and} \quad \phi_c = \begin{cases} \psi_c^+ \cdots \psi_2^+\psi_1^+ & \text{if } c > 0, \\ 1 & \text{if } c = 0, \\ \psi_{c+1}^- \cdots \psi_{-1}^-\psi_0^- & \text{if } c < 0. \end{cases}$$

Then, for $c > 0$, we have

$$\Phi_c 1 \overset{(5.52)}{=} q^c (B_0^+)^c 1 \overset{(5.40)}{=} q^c.$$

Similarly, for $c < 0$, we have

$$\Phi_c 1 \overset{(5.52)}{=} q^c (B_0^-)^{-c} 1 \overset{(5.40)}{=} q^c.$$

Since we clearly have $\Phi_0 1 = 1$, it follows that $\Phi_c 1 = q^c$ for all $c \in \mathbb{Z}$. We also have

$$\phi_c v_\varnothing^{(0)} = v_\varnothing^{(c)} \quad \text{for all } c \in \mathbb{Z}. \tag{5.56}$$

(See Exercise 5.5.5.)

Now, for $c \in \mathbb{Z}$ and $\lambda = (\lambda_1, \dots, \lambda_k) \in \mathrm{Par}$, we have

$$\begin{aligned}\Psi^+_{\lambda_1+c}\Psi^+_{\lambda_2+c-1}\cdots\Psi^+_{\lambda_k+c-k+1}\Phi_{c-k}1 &= \Psi^+_{\lambda_1+c}\Psi^+_{\lambda_2+c-1}\cdots\Psi^+_{\lambda_k+c-k+1}q^{c-k}\\ &\overset{(5.52)}{=} q^c B^+_{\lambda_1}B^+_{\lambda_2}\cdots B^+_{\lambda_{k-1}}B^+_{\lambda_k}1\\ &\overset{(5.42)}{=} q^c s_\lambda.\end{aligned}$$

Since $q^c s_\lambda$, $c \in \mathbb{Z}$, $\lambda \in \mathrm{Par}$, span $\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$, it follows that 1 generates $\mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$ as a Cl-module.

Theorem 5.3.8 now implies that we have an isomorphism of Cl-modules $\mathbf{A}' \colon \mathcal{F} \xrightarrow{\cong} \mathrm{Sym}_{\mathbb{Q}[q,q^{-1}]}$ sending $v_\varnothing^{(0)}$ to 1. Then we have

$$\begin{aligned}\mathbf{A}'\left(v_\lambda^{(c)}\right) &= \mathbf{A}'\left(\psi^+_{\lambda_1+c}\psi^+_{\lambda_2+c-1}\cdots\psi^+_{\lambda_k+c-k+1}\phi_{c-k}v_\varnothing^{(0)}\right)\\ &= \Psi^+_{\lambda_1+c}\Psi^+_{\lambda_2+c-1}\cdots\Psi^+_{\lambda_k+c-k+1}\Phi_{c-k}1 = q^c s_\lambda.\end{aligned}$$

Hence $\mathbf{A}' = \mathbf{A}$, and so (5.55) commutes. □

It follows from (2.38), (5.39), and (5.51) that

$$\Psi^+(t) = \sum_{i\in\mathbb{Z}} \Psi_i^+ t^i = Tq \exp\left(\sum_{n=1}^\infty \frac{t^n}{n} p_n^+\right) \exp\left(-\sum_{n=1}^\infty \frac{t^{-n}}{n} p_n^-\right) \quad \text{and} \tag{5.57}$$

$$\Psi^-(t) = \sum_{i\in\mathbb{Z}} \Psi_i^- t^{-i} = q^{-1}T^{-1} \exp\left(-\sum_{n=1}^\infty \frac{t^{-n}}{n} p_n^+\right) \exp\left(\sum_{n=1}^\infty \frac{t^n}{n} p_n^-\right). \tag{5.58}$$

These expressions are known as *vertex operators*. They give the Clifford algebra generators $\psi_i^\pm$ in terms of the standard Heisenberg algebra generators $p_n^\pm$.

Exercises.

5.5.1. Show that, for all $n \in \mathbb{Z}$, the operators B_n^+ and B_n^- are adjoint with respect to the Hall inner product. In other words, show that

$$\langle B_n^+ f, g\rangle = \langle f, B_n^- g\rangle \qquad \text{for all } f, g \in \mathrm{Sym}_{\mathbb{Q}}.$$

5.5.2. Prove Lemma 5.5.5

5.5.3. With $\varphi(tu)$ as in (5.43), show that

$$f(t)\varphi(tu) = f(u^{-1})\varphi(tu) \qquad \text{for all } f(t) \in \mathbb{Q}[t].$$

5.5.4. Prove (5.46).

5.5.5. Prove (5.56).

5.5.6. Recall the definition of the ring of formal Laurent series from Exercise 1.4.1. Show that $E^-(-t^{-1})H^+(t)$ is a well-defined map from Sym to $\mathrm{Sym}(\!(t)\!)$, but $H^-(-t^{-1})H^+(t)$ is not.

Chapter 6

Applications to representation theory

In this final chapter, we discuss important applications of the theory of symmetric functions to representation theory. We begin by giving a brief overview of the character theory of finite groups. We then give a precise relationship between the characters of the symmetric group and the ring of symmetric functions. This connection, together with our knowledge of symmetric functions, allows us to deduce representation theoretic results. In particular, we give a combinatorial formula for the character table for the symmetric groups and explicit constructions of all their irreducible representations. We conclude with a very brief discussion of how Schur functions also give characters for polynomial representations of the general linear group.

6.1 Characters of finite groups

In this section we give a brief summary of the theory of characters of finite groups. In order to keep our treatment as concise as possible, we restrict our attention to working over a ground field of characteristic zero. More details on the material found here, including proofs, can be found, for instance, in [DF04, Ch. 19, 20].

Throughout this section, G is a finite group and $\Bbbk$ is an arbitrary field of characteristic zero unless otherwise specified. The *group algebra* of G, denoted $\Bbbk G$, is the $\Bbbk$-vector space with basis G, with multiplication given by

$$\left(\sum_{x\in G} a_x x\right)\left(\sum_{y\in G} b_y y\right) = \sum_{x,y} a_x b_y (xy), \qquad a_x, b_y \in \Bbbk.$$

A *representation* of G over $\Bbbk$ is a homomorphism of algebras

$$\rho\colon \Bbbk G \to \mathrm{End}_{\Bbbk}(V)$$

where V is a finite-dimensional vector space over $\Bbbk$. (The assumption of finite dimensionality is not necessary in general. However, we will assume through this chapter that all representation are finite dimensional.) Such a representation defines a G-module structure on V given by

$$x \cdot v := \rho(x)(v), \qquad x \in G,\ v \in V.$$

Conversely, every G-module V defines a representation of G by

$$\rho\left(\sum_{x\in G} a_x x\right)(v) := \sum_{x\in G} a_x(x\cdot v), \qquad a_x \in \Bbbk.$$

Note that, by definition, a G-module is the same as a $\Bbbk G$-module.

The *degree* of a representation is the dimension of the corresponding G-module V. A representation is said to be *irreducible* if the corresponding G-module is simple. Otherwise it is said to be *reducible*. Two representations are *equivalent*, or *isomorphic*, if the corresponding modules are isomorphic. Equivalently, representations $\rho_1\colon \Bbbk G \to \mathrm{End}_{\Bbbk}(V_1)$ and $\rho_2\colon \Bbbk G \to \mathrm{End}_{\Bbbk}(V_2)$ are equivalent if there exists an isomorphism of $\Bbbk$-vector spaces $f\colon V_1 \to V_2$ such that the diagram

$$\begin{array}{ccc} V_1 & \xrightarrow{\rho_1(x)} & V_1 \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ V_2 & \xrightarrow{\rho_2(x)} & V_2 \end{array}$$

commutes for all $x \in G$.

Proposition 6.1.1 (Maschke's theorem)**.** *Every G-module is a direct sum of simple modules.*

Proof. For a proof, see, for example, [Sav, Th. 4.6.3]. □

For the remainder of this section, let Z be an associative unital $\Bbbk$-algebra. For example, we can have $Z = \Bbbk$. Let $\mathcal{F}(G, Z)$ denote the ring of functions from G to Z with pointwise multiplication. An element $f \in \mathcal{F}(G, Z)$ is a *class function* if it is constant on conjugacy classes, that is, if

$$f\left(y^{-1}xy\right) = f(x) \qquad \text{for all } x, y \in G.$$

We let $\mathcal{C}(G, Z)$ denote the set of class functions on G with values in Z. It is straightforward to verify that this is a subring of $\mathcal{F}(G, Z)$ and that $\dim_{\Bbbk} \mathcal{C}(G, \Bbbk)$ is the number of conjugacy classes of G; see Exercise 6.1.3.

If $\rho\colon G \to \mathrm{End}_{\Bbbk}(V)$ is a representation, the *character* of ρ is the function

$$\chi_\rho = \chi_V\colon G \to \Bbbk, \qquad \chi_\rho(x) = \operatorname{tr}\rho(x),$$

where $\operatorname{tr}\rho(x)$ is the trace of the matrix $\rho(x)$ with respect to some basis of V. Every character is a class function (Exercise 6.1.1) and equivalent representations have the same character (Exercise 6.1.2). A character is said to be *irreducible* or *reducible* when the corresponding representation is irreducible or reducible, respectively. The *degree* of a character is the degree of the corresponding representation. Since $\rho(1)$ is the identity linear transformation of V and the trace of an $n \times n$ identity matrix is n, we see that

$$\chi_\rho(1) \text{ is the degree of } \rho.$$

For class functions $f, g \in \mathcal{C}(G, \mathbb{C})$, define

$$\langle f, g\rangle_G^{\mathbb{C}} := \frac{1}{|G|}\sum_{x\in G} f(x)\overline{g(x)}, \tag{6.1}$$

where $\bar{z}$ denotes the complex conjugate of $z \in \mathbb{C}$. We leave it as Exercise 6.1.4 to show that (6.1) is a Hermitian inner product on $\mathcal{C}(G, \mathbb{C})$. In other words, for all $a, b \in \mathbb{C}$, we have

- $\langle af_1 + bf_2, g\rangle_G^{\mathbb{C}} = a\langle f_1, g\rangle_G^{\mathbb{C}} + b\langle f_2, g\rangle_G^{\mathbb{C}}$,
- $\langle f, g\rangle_G^{\mathbb{C}} = \overline{\langle g, f\rangle_G^{\mathbb{C}}}$, and
- if $f \neq 0$, then $\langle f, f\rangle_G^{\mathbb{C}} > 0$.

We define the *norm* of a class function f to be

$$\|f\| := \sqrt{\langle f, f\rangle_G^{\mathbb{C}}}.$$

Proposition 6.1.2. *Assume* $\Bbbk = \mathbb{C}$.

(a) *Two representations are equivalent if and only if they have the same character.*

(b) *The irreducible characters are an orthonormal basis for* $\mathcal{C}(G, \mathbb{C})$.

(c) *If* χ_1 *and* χ_2 *are characters corresponding to* G*-modules* V_1 *and* V_2*, respectively, then* $\chi_1 + \chi_2$ *is the character of* $V_1 \oplus V_2$ *and* $\chi_1\chi_2$ *is the character of* $V_1 \otimes V_2$*. (Recall the definition of the* G*-module structure on* $V_1 \otimes V_2$ *from Example 4.3.10(a).)*

(d) *If* χ *is any character of* G*, then* $\chi(x)$ *is a sum of roots of unity in* $\mathbb{C}$ *and* $\chi(x^{-1}) = \overline{\chi(x)}$ *for all* $x \in G$.

It follows from Proposition 6.1.2(b) that the number of pairwise inequivalent irreducible characters of G is equal to the number of conjugacy class of G. Hence, by Proposition 6.1.2(a), the number of isomorphism classes of simple G-modules is also equal to the number of conjugacy classes of G.

The *character table* of G is the table whose rows are labelled by irreducible representations of G and whose columns are labelled by conjugacy classes of G. The entry in the row labelled by the irreducible representation ρ and the column labelled by the conjugacy class C is $\chi_\rho(x)$, where $x \in C$. Thus the character table of G completely describes the characters of G.

Characters give a method for computing the decomposition of a G-module into a sum of simple modules as follows. Suppose we are working over $\Bbbk = \mathbb{C}$, and let $\chi_1, \ldots, \chi_r$ be the irreducible characters of G, with corresponding modules $V_1, \ldots, V_r$. If V is a G-module, then we have

$$\chi_V = \sum_{i=1}^{r} \langle \chi_V, \chi_i\rangle_G^{\mathbb{C}} \chi_i \qquad \text{and} \qquad \langle \chi_V, \chi_i\rangle_G^{\mathbb{C}} \in \mathbb{N} \text{ for } 1 \leq i \leq n.$$

Then we have an isomorphism of G-modules

$$V \cong \bigoplus_{i=1}^{r} V_i^{\oplus \langle \chi_V, \chi_i\rangle_G^{\mathbb{C}}}, \qquad \text{where } V^{\oplus n} = \underbrace{V \oplus \cdots \oplus V}_{n \text{ summands}}$$

In other words, $\langle \chi_V, \chi_i\rangle_G^{\mathbb{C}}$ gives the multiplicity of V_i in the decomposition of V. In particular, a class function is the character of some representation if and only if it is a nonnegative integer linear combination of irreducible characters.

An *integer* linear combination of characters is called a *virtual character*. In other words, if $\chi_1, \ldots, \chi_r$ are the irreducible characters of G, then a virtual character is a class function of the form

$$f = \sum_{i=1}^{r} a_i \chi_i, \qquad a_1, \ldots, a_r \in \mathbb{Z}.$$

Since

$$\|f\| = \left(\sum_{i=1}^{r} a_i^2\right)^{1/2},$$

it follows that a virtual character has norm 1 if and only if it is an irreducible character up to sign. (We can then check the sign by computing $f(1)$, which is positive if f is a character.) Note that this is *not* true for arbitrary class functions.

Let $\mathcal{R}(G,\mathbb{C})$ denote the $\mathbb{Z}$-module of complex virtual characters. It follows from Proposition 6.1.2(c) that $\mathcal{R}(G,\mathbb{C})$ is closed under multiplication, and hence is a subring of $\mathcal{C}(G,\mathbb{C})$.

Lemma 6.1.3. *The irreducible complex characters are the unique orthonormal basis of* $\mathcal{R}(G,\mathbb{C})$ *up to sign and reordering.*

Proof. The proof of this lemma is almost identical to the proof of Proposition 3.3.8. □

It follows from Proposition 6.1.2(d) that

$$\langle f,g\rangle_G^{\mathbb{C}} = \frac{1}{|G|}\sum_{x\in G} f(x)g(x^{-1}), \qquad f,g\in\mathcal{R}(G,\mathbb{C}). \tag{6.2}$$

Motivated by this, for any associative unital $\Bbbk$-algebra Z, any functions $f,g\colon G\to Z$,

$$\langle f,g\rangle_G = \frac{1}{|G|}\sum_{x\in G} f(x)g(x^{-1}). \tag{6.3}$$

Thus $\langle f,g\rangle_G = \langle f,g\rangle_G^{\mathbb{C}}$ for $f,g\in\mathcal{R}(G,\mathbb{C})$, but these can differ for general class functions $f,g\in\mathcal{C}(G,\mathbb{C})$. Since we will be most interested in the subring $\mathcal{R}(G,\mathbb{C})$, we will work from now on with the inner product (6.3) since it has the advantage of being well defined over any $\Bbbk$-algebra Z.

Now suppose G_1 and G_2 are both finite groups. If $f_1\in\mathcal{R}(G_1,\mathbb{C})$ and $f_2\in\mathcal{R}(G_2,\mathbb{C})$, then

$$\begin{gathered} f_1\times f_2\colon G_1\times G_2\to\mathbb{C},\\ (f_1\times f_2)(x_1,x_2) = f_1(x_1)f_2(x_2), \quad x_1\in G_1,\ x_2\in G_2, \end{gathered} \tag{6.4}$$

is a class function on $G_1\times G_2$. Furthermore, if f_1 and f_2 are characters, then so is $f_1\times f_2$. More precisely, if f_i is the character of the G_i-module V_i, for $i\in\{1,2\}$, then $f_1\times f_2$ is the character of the $(G_1\times G_2)$-module $V_1\otimes V_2$, with action given by

$$(g_1,g_2)(v_1\otimes v_2) = g_1v_1\otimes g_2v_2, \qquad g_1\in G_1,\ g_2\in G_2,\ v_1\in V_1,\ v_2\in V_2,$$

extended by linearity. Thus, we have a bilinear map

$$\mathcal{R}(G_1,\mathbb{C})\times\mathcal{R}(G_2,\mathbb{C})\to\mathcal{R}(G_1\times G_2,\mathbb{C}).$$

If χ^1,ψ^1 are irreducible characters of G_1 and χ^2,ψ^2 are irreducible characters of G_2, then it follows from (6.3) that

$$\left\langle \chi^1\times\chi^2,\psi^1\times\psi^2\right\rangle_{G_1\times G_2} = \left\langle \chi^1,\psi^1\right\rangle_{G_1}\left\langle \chi^2,\psi^2\right\rangle_{G_2} = \delta_{\chi^1\psi^1}\delta_{\chi^2\psi^2}.$$

Hence the characters $\chi^1 \times \chi^2$ have norm 1, and so are irreducible, and they are pairwise distinct. Since the conjugacy classes of $G_1 \times G_2$ are precisely the products $C_1 \times C_2$ of conjugacy classes C_i of G_i, the number of such characters is equal to the number of conjugacy classes of $G_1 \times G_2$. So, by Proposition 6.1.2(b), they are all of the irreducible characters of $G_1 \times G_2$. (In other words, every simple $(G_1 \times G_2)$-module is isomorphic to $V_1 \otimes V_2$ for some simple G_i-modules V_i, $i \in \{1, 2\}$, which are unique up to isomorphism.) It follows that we have an isomorphism of algebras

$$\mathcal{R}(G_1, \mathbb{C}) \otimes \mathcal{R}(G_2, \mathbb{C}) \xrightarrow{\cong} \mathcal{R}(G_1 \times G_2, \mathbb{C}), \qquad \chi^1 \otimes \chi^2 \mapsto \chi^1 \times \chi^2, \tag{6.5}$$

for characters χ^1 of G_1 and χ^2 of G_2, extended by $\mathbb{Z}$-linearity.

Exercises.

6.1.1. Prove that characters are class functions.

6.1.2. Prove that equivalent representations have the same character.

6.1.3. Prove that $\mathcal{C}(G, Z)$ is a subring of $\mathcal{F}(G, Z)$ and that $\dim_{\Bbbk} \mathcal{C}(G, \Bbbk)$ is the number of conjugacy classes of G.

6.1.4. Prove that (6.1) is a Hermitian inner product on the space of class functions.

6.2 Induction and restriction

In order to connect the representation theory of the symmetric groups to the ring of symmetric functions, it is necessary to consider all of the symmetric groups $\mathfrak{S}_n$, $n \in \mathbb{N}$, at once. The main tools for doing this are the operations of induction and restriction, which allow us to move between representations of different symmetric groups. In this section we introduce these operations. Further details, including proofs, can be found, for instance, in [DF04, §19.3].

Throughout this section we fix a field $\Bbbk$ of characteristic zero and an associative unital $\Bbbk$-algebra Z. (One important example will be $\Bbbk = Z = \mathbb{C}$.) We also fix a subgroup H of a finite group G.

If V is a G-module, then we can restrict the action to obtain an H-module, which we denote by $\mathrm{Res}^G_H(V)$. If χ is the character of V, then the *restricted character* $\mathrm{Res}^G_H(\chi)$ is the character of the restricted module $\mathrm{Res}^G_H(V)$. We can extend this to a ring homomorphism

$$\mathrm{Res}^G_H \colon \mathcal{C}(G, Z) \to \mathcal{C}(H, Z),$$
$$\mathrm{Res}^G_H(f)(x) = f(x), \quad f \in \mathcal{C}(G, Z),\ x \in H.$$

We would now like to define a natural map in the other direction: $\mathcal{C}(H,Z) \to \mathcal{C}(G,Z)$. Given $f \in \mathcal{C}(H,Z)$, we can extend it to a function $f\colon G \to Z$ (which continue to denote by the same symbol f) by declaring

$$f(x) = 0 \quad \text{if} \quad x \in G,\ x \notin H. \tag{6.6}$$

In general, this extended function will not be a class function. However, we can fix this by averaging over G, defining the *induced class function*

$$\mathrm{Ind}_H^G(f)(x) = \frac{1}{|H|}\sum_{y\in G} f(y^{-1}xy), \qquad f \in \mathcal{C}(H,Z),\ x \in G. \tag{6.7}$$

We leave it as Exercise 6.2.1 to show that this defines a class function on G. In fact, when $f \in \mathcal{R}(H,\mathbb{C})$, we would like to show that $\mathrm{Ind}_H^G(f) \in \mathcal{R}(G,\mathbb{C})$. This will follow from the fact that, if χ is a character, then so is the induced character $\mathrm{Ind}_H^G(\chi)$. Namely, it is the character of an induced module, which we now describe.

We begin with a more general construction. Let $\Bbbk$ be an arbitrary field, and let B be an associative unital $\Bbbk$-algebra. If M is a right B-module and V is a left B-module, then we define $M \otimes_B V$ to be the $\Bbbk$-vector space generated by the elements $M \times V$, modulo the subspace W generated by the elements

$$(m+m',v)-(m,v)-(m',v), \qquad (m,v+v')-(m,v)-(m,v'), \qquad (mb,v)-(m,bv),$$
$$(cm,v)-c(m,v), \qquad (m,cv)-c(m,v),$$

for $m,m' \in M$, $v,v' \in V$, $b \in B$, and $c \in \Bbbk$. We set

$$m \otimes v := (m,v) + W, \qquad m \in M, v \in V.$$

This notion of tensor product has a universal property analogous to the one of Proposition 4.1.2, as we now describe. Consider the $\Bbbk$-bilinear map

$$h\colon M \times V \to M \otimes_B V, \qquad h(m,v) = m \otimes v.$$

If W is a $\Bbbk$-vector space, we say that a map $\varphi\colon M \times V \to W$ is *B-balanced* if

$$\varphi(mb,v) = \varphi(m,bv) \qquad \text{for all } m \in M,\ b \in B,\ v \in V.$$

Proposition 6.2.1. *For every $\Bbbk$-module W and every $\Bbbk$-bilinear, B-balanced map $\varphi\colon M \times V \to W$, there exists a unique $\Bbbk$-linear map $\tilde{\varphi}\colon M \otimes_B V \to W$ such that $\tilde{\varphi} \circ h = \varphi$; that is, the diagram*

$$\begin{array}{ccc} M \times V & \xrightarrow{\ h\ } & M \otimes_B V \\ & {\scriptstyle\varphi}\searrow & \downarrow{\scriptstyle\tilde{\varphi}} \\ & & W \end{array}$$

commutes.

Proof. For a proof, see [Sav, Prop. 3.5.3]. □

Now suppose that A and B are both associative unital $\Bbbk$-algebras. An (A, B)-*bimodule* is a $\Bbbk$-vector space M that is a left A-module, a right B-module, and the two actions commute:

$$(am)b = a(mb) \qquad \text{for all } a \in A,\ b \in B,\ m \in M.$$

Then $M \otimes_B V$ is naturally an A-module, with action given by

$$a(m \otimes v) = (am) \otimes v, \qquad a \in A,\ m \in M,\ v \in V.$$

In this way, the bimodule M gives us a way of turning B-modules into A-modules.

Remark 6.2.2. Note that any $\Bbbk$-vector space V is a $(\Bbbk, \Bbbk)$-bimodule, where the right and left action are both given by scalar multiplication. Then, if $A = B = \Bbbk$, the above notion of tensor product agrees with the one we discussed in Section 4.1. The crucial property in this case is that $\Bbbk$ is commutative, which makes left and right $\Bbbk$-modules equivalent. △

We now consider a special case of the above construction. The group algebra $\Bbbk G$ is both a left and a right module over itself, via multiplication. These two actions commute, and so $\Bbbk G$ is a $(\Bbbk G, \Bbbk G)$-bimodule. We can restrict the action on the right, and consider $\Bbbk G$ as a $(\Bbbk G, \Bbbk H)$-bimodule. If V is an H-module, then we have the *induced module*

$$\operatorname{Ind}_G^H(V) := \Bbbk G \otimes_{\Bbbk H} V.$$

Recall that the *left cosets* of H in G are the sets

$$xH = \{xy : y \in H\}, \qquad x \in G.$$

Let $x_1, \ldots, x_m$ be a complete set of left coset representatives. Then $\Bbbk G$ is free right $\Bbbk H$-module of rank m:

$$\Bbbk G = x_1 \Bbbk H \oplus \cdots \oplus x_m \Bbbk H.$$

If V is an H-module with $\Bbbk$-basis $v_1, \ldots, v_n$, the induced module $\operatorname{Ind}_G^H(V)$ has $\Bbbk$-basis

$$x_i \otimes v_j, \qquad 1 \le i \le m,\ 1 \le j \le n.$$

In particular,

$$\dim_{\Bbbk} \operatorname{Ind}_H^G(V) = m \dim_{\Bbbk} V.$$

In addition, for $f \in \mathcal{C}(H, Z)$ and $x \in G$, we have

$$\operatorname{Ind}_H^G(f)(x) = \frac{1}{|H|} \sum_{y \in G} f\left(y^{-1}xy\right) = \sum_{i=1}^{m} f\left(x_i^{-1}xx_i\right), \tag{6.8}$$

(The first equality is the definition (6.7), and the second is left as Exercise 6.2.3.)

Proposition 6.2.3. *If H is a subgroup of a finite group G and V is an H-module, then*

$$\operatorname{Ind}_H^G(\chi_V) = \chi_{\operatorname{Ind}_H^G(V)}.$$

In other words, the induced character $\operatorname{Ind}_H^G(\chi_V)$ is the character of the induced module $\operatorname{Ind}_H^G(V)$.

It follows from Proposition 6.2.3 that induction yields a $\mathbb{Z}$-linear map

$$\mathrm{Ind}_H^G \colon \mathcal{R}(H, \mathbb{C}) \to \mathcal{R}(G, \mathbb{C}).$$

We conclude this section with an important result in representation theory that states that this map is adjoint to restriction. (In the language of category theory, they are adjoint functors.)

Proposition 6.2.4 (Frobenius reciprocity)**.** *Suppose H is a subgroup of the finite group G, Z is a $\Bbbk$-algebra, $f \in \mathcal{C}(H, Z)$, and $g \in \mathcal{C}(G, Z)$. Then*

$$\left\langle \mathrm{Ind}_H^G(f), g \right\rangle_G = \left\langle f, \mathrm{Res}_H^G(g) \right\rangle_H.$$

Proof. We have

$$\begin{aligned}
\left\langle \mathrm{Ind}_H^G(f), g \right\rangle_G &\overset{(6.3)}{=} \frac{1}{|G|} \sum_{x \in G} \mathrm{Ind}_H^G(f)(x) g(x^{-1}) \\
&\overset{(6.7)}{=} \frac{1}{|G||H|} \sum_{x,y \in G} f(y^{-1}xy) g(x^{-1}) \\
&= \frac{1}{|G||H|} \sum_{y \in G} \sum_{x \in yGy^{-1}} f(y^{-1}xy) g(x^{-1}) && \text{(since } yGy^{-1} = G\text{)} \\
&= \frac{1}{|G||H|} \sum_{x,y \in G} f(x) g(yx^{-1}y^{-1}) \\
&= \frac{1}{|G||H|} \sum_{x,y \in G} f(x) g(x^{-1}) && \text{(since } g \text{ is a class function)} \\
&= \frac{1}{|H|} \sum_{x \in G} f(x) g(x^{-1}) \\
&= \frac{1}{|H|} \sum_{x \in H} f(x) g(x^{-1}) && \text{(since } f(x) = 0 \text{ for } x \in G,\ x \notin H\text{)} \\
&\overset{(6.3)}{=} \left\langle f, \mathrm{Res}_H^G(g) \right\rangle_H.
\end{aligned}$$

□

Exercises.

6.2.1. Prove that the induced class function $\mathrm{Ind}_H^G(f)$, as defined in (6.7), is a class function directly from the definition.

6.2.2. Show that if H is a subgroup of an abelian group G and $f \in \mathcal{C}(H, \Bbbk)$, then $\mathrm{Ind}_H^G(f) = \frac{|G|}{|H|} f$. Here, on the right-hand side, f denotes the extended function (6.6).

6.2.3. Prove the second equality in (6.8).

6.3 Characters of symmetric groups

In this section we study the characters of the symmetric groups and describe a precise connection with the ring of symmetric functions. In particular, we will see that irreducible characters correspond to Schur functions. Once we have established this connection, our knowledge of symmetric functions allows us to deduce facts about the representation theory of the symmetric groups. Our approach is guided by the presentation of [Mac15, §I.7].

For $m, n \in \mathbb{N}$, we have an injective group homomorphism

$$\mathfrak{S}_m \times \mathfrak{S}_n \hookrightarrow \mathfrak{S}_{m+n}, \qquad (\pi, \sigma) \mapsto \pi \times \sigma,$$

where $\mathfrak{S}_m$ permutes the elements $\{1, \dots, m\}$ and $\mathfrak{S}_n$ permutes the elements $\{m+1, \dots, m+n\}$. Precisely, for $\pi \in \mathfrak{S}_m$ and $\sigma \in \mathfrak{S}_n$, we have

$$(\pi \times \sigma)(k) = \begin{cases} \pi(k) & \text{if } 1 \le k \le m, \\ \sigma(k-m) + m & \text{if } m < k \le m+n. \end{cases}$$

We will identify $\mathfrak{S}_m \times \mathfrak{S}_n$ with its image under the above map and, in this way, view $\mathfrak{S}_m \times \mathfrak{S}_n$ as a subgroup of $\mathfrak{S}_{m+n}$.

Recall from Section 1.1 that every permutation $\pi \in \mathfrak{S}_n$ factors uniquely as a product of disjoint cycles. If the lengths of these cycles are $\rho_1, \rho_2, \dots$, ordered so that $\rho_1 \ge \rho_2 \ge \cdots$, then $\rho(\pi) := (\rho_1, \rho_2, \dots)$ is a partition of n called the *cycle type* of π. For $\rho \in \mathrm{Par}(n)$, the elements of $\mathfrak{S}_n$ of cycle type ρ form a conjugacy class of $\mathfrak{S}_n$. Thus, the conjugacy classes of $\mathfrak{S}_n$ are indexed by partitions of n. Let

$$C_\rho \subseteq \mathfrak{S}_n \text{ denote the conjugacy class of elements of cycle type } \rho.$$

By convention, $\mathfrak{S}_0$ is the trivial group.

Let

$$R = \bigoplus_{n=0}^{\infty} R^n, \qquad R^n = \mathcal{R}(\mathfrak{S}_n, \mathbb{C}).$$

Note that, since $\mathfrak{S}_0 = \{1\}$ is the trivial group, we have $R^0 = \mathbb{Z}$, spanned by the trivial character (that is, the character of the trivial representation of the trivial group). We can endow R with the structure of a ring as follows. If $f \in R^m$ and $g \in R^n$, then $f \times g$, as defined in (6.4), is a virtual character of $\mathfrak{S}_m \times \mathfrak{S}_n$. We then define

$$f \star g = \mathrm{Ind}_{\mathfrak{S}_m \times \mathfrak{S}_n}^{\mathfrak{S}_{m+n}} (f \times g) \in R^{m+n}. \tag{6.9}$$

In this way, we have defined a bilinear multiplication

$$\star \colon R^m \times R^n \to R^{m+n}.$$

We leave it as Exercise 6.3.1 to verify that, with this multiplication, R is a commutative, associative, unital, graded ring. Note that $\star$ is *not* the same as pointwise multiplication (denoted by juxtaposition), which gives a multiplication on each R^n.

We can also endow R with a scalar product. Recall the inner product (see (6.2))

$$R^n \times R^n \to \mathbb{C}, \qquad \langle f, g\rangle_{\mathfrak{S}_n} = \frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n} f(\pi)g(\pi^{-1}). \tag{6.10}$$

If $f, g \in R$, we define

$$\langle f, g\rangle = \sum_{n=0}^{\infty}\langle f_n, g_n\rangle_{\mathfrak{S}_n},$$

where $f = \sum_{n\geq 0} f_n$, $g = \sum_{n\geq 0} g_n$ with $f_n, g_n \in R^n$.

Let

$$\mathrm{Sym}_{\mathbb{C}} := \mathbb{C}\otimes_{\mathbb{Z}} \mathrm{Sym} \qquad \text{and} \qquad \mathrm{Sym}^n_{\mathbb{C}} := \mathbb{C}\otimes_{\mathbb{Z}} \mathrm{Sym}^n,$$

We will view Sym as a subring of $\mathrm{Sym}_{\mathbb{C}}$ via the inclusion

$$\mathrm{Sym} \hookrightarrow \mathrm{Sym}_{\mathbb{C}}, \qquad f \mapsto 1\otimes f.$$

Now define $\Psi \in \mathcal{C}(\mathfrak{S}_n, \mathrm{Sym}^n_{\mathbb{C}})$ by

$$\Psi\colon \mathfrak{S}_n \to \mathrm{Sym}^n_{\mathbb{C}}, \qquad \Psi(\pi) = p_{\rho(\pi)}.$$

If $\pi \in \mathfrak{S}_m$ and $\sigma \in \mathfrak{S}_n$, then $\rho(\pi\times\sigma) = \rho(\pi)\cup\rho(\sigma)$. Hence

$$\Psi(\pi\times\sigma) = \Psi(\pi)\Psi(\sigma). \tag{6.11}$$

Next we define a $\mathbb{Z}$-linear map

$$\begin{gathered}\mathrm{ch}\colon R \to \mathrm{Sym}_{\mathbb{C}},\\ \mathrm{ch}(f) = \langle f, \Psi\rangle_{\mathfrak{S}_n} = \frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n} f(\pi)\Psi(\pi), \quad f\in R^n,\end{gathered} \tag{6.12}$$

extended by linearity. (Note that $\Psi(\pi^{-1}) = \Psi(\pi)$ for all $\pi \in \mathfrak{S}_n$, which is why there is no inverse in (6.12), as opposed to (6.3).) We call $\mathrm{ch}(f)$ the *characteristic* of f and call ch the *characteristic map*.

Proposition 6.3.1. *The characteristic map (6.12) is a homomorphism of algebras.*

Proof. Suppose $f \in R^m$ and $g \in R^n$. Then

$$\begin{aligned}
\mathrm{ch}(f\star g) &\overset{\substack{(6.9)\\(6.12)}}{=} \left\langle \mathrm{Ind}^{\mathfrak{S}_{m+n}}_{\mathfrak{S}_m\times\mathfrak{S}_n}(f\times g), \Psi\right\rangle_{\mathfrak{S}_{m+n}}\\
&= \left\langle f\times g, \mathrm{Res}^{\mathfrak{S}_{m+n}}_{\mathfrak{S}_m\times\mathfrak{S}_n}(\Psi)\right\rangle_{\mathfrak{S}_m\times\mathfrak{S}_n} && \text{(Proposition 6.2.4)}\\
&\overset{(6.12)}{=} \frac{1}{m!n!}\sum_{\substack{\pi\in\mathfrak{S}_m\\ \sigma\in\mathfrak{S}_n}} (f\times g)(\pi\times\sigma)\Psi(\pi\times\sigma)\\
&\overset{\substack{(6.4)\\(6.11)}}{=} \frac{1}{m!n!}\sum_{\substack{\pi\in\mathfrak{S}_m\\ \sigma\in\mathfrak{S}_n}} f(\pi)g(\sigma)\Psi(\pi)\Psi(\sigma)\\
&= \langle f, \Psi\rangle_{\mathfrak{S}_m}\langle g, \Psi\rangle_{\mathfrak{S}_n}\\
&= \mathrm{ch}(f)\,\mathrm{ch}(g).
\end{aligned}$$

□

Lemma 6.3.2. *For $f \in R^n$,*

$$\mathrm{ch}(f) = \sum_{\rho \in \mathrm{Par}(n)} z_\rho^{-1} f_\rho p_\rho, \tag{6.13}$$

where f_ρ is the value of f at elements of cycle type ρ, and z_ρ is as defined in (2.47).

Proof. We have

$$\mathrm{ch}(f) = \frac{1}{n!} \sum_{\rho \in \mathrm{Par}(n)} \sum_{\pi \in C_\rho} f(\pi)\Psi(\pi) = \frac{1}{n!} \sum_{\rho \in \mathrm{Par}(n)} |C_\rho| f_\rho p_\rho.$$

It thus suffices to prove that

$$\frac{n!}{|C_\rho|} = z_\rho = \prod_{i \geq 1} i^{m_i(\rho)} m_i(\rho)!. \tag{6.14}$$

Fix $\sigma \in \mathfrak{S}_n$ of cycle type ρ and consider the action of $\mathfrak{S}_n$ on itself by conjugation, $\pi \cdot \tau = \pi\tau\pi^{-1}$. The orbit of σ is precisely the conjugacy class C_ρ, and the stabilizer of σ is the set X of elements of $\mathfrak{S}_n$ that commute with σ. Thus, by the orbit–stabilizer theorem, we have

$$|C_\rho| = \frac{|\mathfrak{S}_n|}{|X|} = \frac{n!}{z_\rho},$$

where, in the last equality, we used Exercise 1.1.2. This proves (6.14). □

Let η^n be the character of the trivial one-dimensional representation of $\mathfrak{S}_n$. Thus

$$\eta^n(\pi) = 1 \qquad \text{for all } \pi \in \mathfrak{S}_n.$$

Let ε^n be the character of the one-dimensional sign representation of $\mathfrak{S}_n$. Thus,

$$\varepsilon^n(\pi) = \mathrm{sgn}(\pi) \qquad \text{for all } \pi \in \mathfrak{S}_n.$$

Note that it follows from the definitions that $\eta^0 = \varepsilon^0 = 1$. We leave it as Exercise 6.3.2 to show that, if $\pi \in \mathfrak{S}_n$ has cycle type ρ, then

$$\varepsilon^n(\pi) = (-1)^{n-\ell(\rho)} \qquad \text{for all } \pi \in \mathfrak{S}_n. \tag{6.15}$$

Now, for $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathrm{Par}(n)$, let

$$\mathfrak{S}_\lambda := \mathfrak{S}_{\lambda_1} \times \cdots \times \mathfrak{S}_{\lambda_k}$$

and define the following elements of R^n:

$$\eta^\lambda := \eta^{\lambda_1} \star \cdots \star \eta^{\lambda_k} = \mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n}(\eta^{\lambda_1} \times \cdots \times \eta^{\lambda_k}), \tag{6.16}$$

$$\varepsilon^\lambda := \varepsilon^{\lambda_1} \star \cdots \star \varepsilon^{\lambda_k} = \mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n}(\varepsilon^{\lambda_1} \times \cdots \times \varepsilon^{\lambda_k}), \tag{6.17}$$

$$\chi^\lambda := \det\left(\eta^{\lambda_i - i + j}\right)_{1 \leq i,j \leq n}, \tag{6.18}$$

where, in computing the determinant, we use the product $\star$ and the convention that $\eta^r = 0$ for $r < 0$. Note that χ^λ is an integer linear combination of characters. In other words, $\chi^\lambda \in R$ is a virtual character. However, we will soon see (Theorem 6.3.6) that, in fact, it is an irreducible character.

Proposition 6.3.3. *For $\lambda \in \mathrm{Par}$,*

$$\mathrm{ch}(\eta^\lambda) = h_\lambda, \qquad \mathrm{ch}(\varepsilon^\lambda) = e_\lambda, \qquad \mathrm{ch}(\chi^\lambda) = s_\lambda. \tag{6.19}$$

Proof. First note that

$$\mathrm{ch}(\eta^n) \overset{(6.13)}{=} \sum_{\rho \in \mathrm{Par}(n)} z_\rho^{-1} p_\rho \overset{(2.49)}{=} h_n.$$

Thus

$$\mathrm{ch}(\eta^\lambda) = \mathrm{ch}(\eta^{\lambda_1}) \cdots \mathrm{ch}(\eta^{\lambda_k}) = h_{\lambda_1} \cdots h_{\lambda_k} = h_\lambda$$

Similarly,

$$\mathrm{ch}(\varepsilon^n) \underset{(6.15)}{\overset{(6.13)}{=}} \sum_{\rho \in \mathrm{Par}(n)} (-1)^{n-\ell(\rho)} z_\rho^{-1} p_\rho \overset{(2.49)}{=} e_n,$$

and so

$$\mathrm{ch}(\varepsilon^\lambda) = \mathrm{ch}(\varepsilon^{\lambda_1}) \cdots \mathrm{ch}(\varepsilon^{\lambda_k}) = e_{\lambda_1} \cdots e_{\lambda_k} = e_\lambda.$$

Finally, since ch is a ring homomorphism, we have

$$\mathrm{ch}\left(\chi^\lambda\right) = \det\left(\mathrm{ch}(\eta^{\lambda_i - i + j})\right) = \det\left(h_{\lambda_i - i + j}\right) \overset{(3.10)}{=} s_\lambda. \qquad \square$$

Theorem 6.3.4. *The characteristic map (6.12) is an isometric algebra isomorphism of R onto* Sym.

Proof. For $f, g \in R^n$, we have

$$\langle \mathrm{ch}(f), \mathrm{ch}(g) \rangle \overset{(6.13)}{=} \sum_{\rho, \rho' \in \mathrm{Par}(n)} z_\rho^{-1} z_{\rho'}^{-1} f_\rho g_{\rho'} \langle p_\rho, p_{\rho'} \rangle \overset{(3.29)}{=} \sum_{\rho \in \mathrm{Par}(n)} z_\rho^{-1} f_\rho g_\rho.$$

On the other hand,

$$\langle f, g \rangle_{\mathfrak{S}_n} = \frac{1}{n!} \sum_{\pi \in \mathfrak{S}_n} f(\pi) g(\pi^{-1}) = \frac{1}{n!} \sum_{\rho \in \mathrm{Par}(n)} |C_\rho| f_\rho g_\rho \overset{(6.14)}{=} \sum_{\rho \in \mathrm{Par}(n)} z_\rho^{-1} f_\rho g_\rho.$$

Thus ch is an isometry.

Now, since ch is an isometry, we have

$$\langle \chi^\lambda, \chi^\mu \rangle = \langle s_\lambda, s_\mu \rangle \overset{(3.30)}{=} \delta_{\lambda\mu} \quad \text{for all } \lambda, \mu \in \mathrm{Par}.$$

It follows from Lemma 6.1.3 that the χ^λ, $\lambda \in \mathrm{Par}(n)$, are, up to sign, irreducible characters of $\mathfrak{S}_n$. Since the number of conjugacy classes in $\mathfrak{S}_n$ is equal to the number of partitions of n, the χ^λ are, in fact, *all* of the irreducible characters of $\mathfrak{S}_n$, up to sign. Therefore the χ^λ, $\lambda \in \mathrm{Par}(n)$, form a basis of R^n. Hence ch is an isomorphism of R^n onto Sym^n for all $n \in \mathbb{N}$. It follows that ch is an isomorphism of R onto Sym. $\square$

Remark 6.3.5. Theorem 6.3.4 gives a representation theoretic motivation for Definition 3.3.1 of the Hall inner product. More precisely, the Hall inner product is precisely the inner product on Sym that corresponds to the natural inner product on virtual characters of the symmetric group. $\triangle$

Recall the definition of a *standard tableau* from Section 2.3.

Theorem 6.3.6. (a) *The irreducible characters of* $\mathfrak{S}_n$ *are* χ^λ, $\lambda \in \mathrm{Par}(n)$, *as defined in (6.18).*

(b) *The degree of* χ^λ *is* $K_{\lambda,(1^n)}$, *the number of standard tableaux of shape* λ.

Proof. We have shown in the proof of Theorem 6.3.4 that the irreducible characters of $\mathfrak{S}_n$ are χ^λ, $\lambda \in \mathrm{Par}(n)$, up to sign. It remains to prove that, for $\lambda \in \mathrm{Par}$, it is χ^λ, and not $-\chi^\lambda$, that is a character. To do this, it suffices to prove that $\chi^\lambda(1)$ is positive. We will show that $\chi^\lambda(1) = K_{\lambda,(1^n)}$, which also proves part (b).

Note that

$$s_\lambda \overset{(6.19)}{=} \mathrm{ch}(\chi^\lambda) \overset{(6.13)}{=} \sum_{\mu \in \mathrm{Par}(n)} z_\mu^{-1} \chi^\lambda_\mu p_\mu,$$

where χ^λ_μ is the value of χ^λ on elements of cycle type μ. Thus

$$\langle s_\lambda, p_\rho \rangle = \sum_{\mu \in \mathrm{Par}(n)} z_\mu^{-1} \chi^\lambda_\mu \langle p_\mu, p_\rho \rangle \overset{(3.29)}{=} \chi^\lambda_\rho. \tag{6.20}$$

Therefore,

$$\chi^\lambda(1) = \chi^\lambda_{(1^n)} = \langle s_\lambda, p_1^n \rangle.$$

Since the Schur functions s_λ, $\lambda \in \mathrm{Par}(n)$, form a basis of Sym^n, it follows that

$$h_1^n = p_1^n = \sum_{\lambda \in \mathrm{Par}(n)} \langle s_\lambda, p_1^n \rangle s_\lambda = \sum_{\lambda \in \mathrm{Par}(n)} \chi^\lambda(1) s_\lambda.$$

Thus, by Theorem 3.5.5 and Definition 3.4.14,

$$\chi^\lambda(1) = M_{(1^n),\lambda}(h, s) = (K^\intercal)_{(1^n),\lambda} = K_{\lambda,(1^n)}$$

is the number of standard tableaux of shape λ. □

Remark 6.3.7. It follows from the Littlewood–Richardson rule (Theorem 3.6.9) that a product of two Schur functions is equal to a linear combination of Schur functions with *nonnegative integer* coefficients. We now have a representation theoretic explanation for this fact. For $\mu \in \mathrm{Par}(m)$ and $\nu \in \mathrm{Par}(n)$, we have

$$s_\mu s_\nu \overset{(6.19)}{=} \mathrm{ch}(\chi^\mu)\,\mathrm{ch}(\chi^\nu) = \mathrm{ch}(\chi^\mu \star \chi^\nu).$$

By definition, the product $\chi^\mu \star \chi^\nu$ is the character of an induced representation. Thus, it is a positive integer linear combination of irreducible characters of $\mathfrak{S}_{m+n}$. It follows that $s_\mu s_\nu$ is a positive integer linear combination of Schur functions. △

Proposition 6.3.8. *The transition matrix* $M(p, s)$ *is the character table of* $\mathfrak{S}_n$. *Equivalently, for* $\rho \in \mathrm{Par}(n)$,

$$p_\rho = \sum_{\lambda \in \mathrm{Par}(n)} \chi^\lambda_\rho s_\lambda. \tag{6.21}$$

Thus χ^λ_ρ *is equal to the coefficient of* $x^{\lambda+\delta^n}$ *in* $a_{\delta^n} p_\rho$.

Proof. The equation (6.21) follows immediately from (6.20). Multiplying both sides by a_{δ^n}, and using (3.7), we have

$$a_{\delta^n} p_\rho = \sum_{\lambda \in \mathrm{Par}(n)} \chi^\lambda_\rho a_{\lambda+\delta^n}.$$

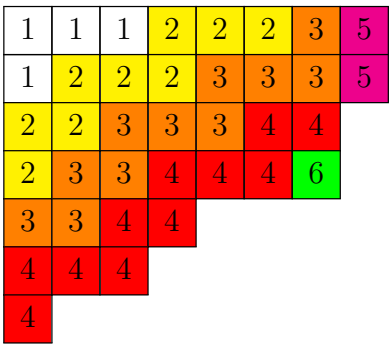
Then the final assertion in the statement of the proposition follows from the fact that the coefficient of $x^{\lambda+\delta^n}$ in $a_{\lambda+\delta^n}$ is 1. □

Combined with the Murnaghan–Nakayama rule (Proposition 3.7.1), the above results lead to a combinatorial method of computing the character tables of the symmetric groups that we now describe. Recall the definition of a border strip from Section 3.7. We define a *border-strip tableau* to be a tableau such that

- the entries in every row and column are weakly increasing, and
- for each i, the boxes with entry i form a border strip.

For example

(6.22)



is a border-strip tableau. For $\lambda, \rho \in \mathrm{Par}(n)$, $n \in \mathbb{N}$, we let $\mathrm{BST}(\lambda, \rho)$ denote the set of border-strip tableaux of shape λ and weight ρ. The *height* $\mathrm{ht}(T)$ of a border-strip tableau T is the sum of the heights of the border strips in T. For example, the height of border-strip tableau in (6.22) is

$$1 + 3 + 4 + 4 + 1 + 0 = 13.$$

The following result is sometimes also referred to as the Murnaghan–Nakayama rule.

Proposition 6.3.9. *For $\lambda, \rho \in \mathrm{Par}(n)$, we have*

$$\chi^\lambda_\rho = \sum_{T \in \mathrm{BST}(\lambda,\rho)} (-1)^{\mathrm{ht}(T)}$$

Proof. Repeated application of the Murnaghan–Nakayama rule (Proposition 3.7.1) shows that

$$p_\rho = p_{\rho_n} \cdots p_{\rho_1} s_\varnothing = \sum_{\mu \in \mathrm{Par}} \sum_{T \in \mathrm{BST}(\mu,\rho_)} (-1)^{\mathrm{ht}(T)} s_\mu.$$

Thus

$$\chi^\lambda_\rho \overset{(6.20)}{=} \langle s_\lambda, p_\rho \rangle = \sum_{\mu \in \mathrm{Par}} \sum_{T \in \mathrm{BST}(\mu,\rho_)} (-1)^{\mathrm{ht}(T)} \langle s_\lambda, s_\mu \rangle \overset{(3.30)}{=} \sum_{T \in \mathrm{BST}(\lambda,\rho)} (-1)^{\mathrm{ht}(T)}.$$

□

Remark 6.3.10. It follows from Proposition 6.3.9 that the characters of the symmetric groups are *integer* valued. In fact, we will prove something even stronger. Namely, the irreducible representations of the symmetric groups can be constructed over the integers; see Remark 6.4.4. This is *not* true for arbitrary groups; see Exercise 6.3.5. △

The next result shows that the involution ω of (2.31) corresponds to multiplication by the character of the sign representation.

Proposition 6.3.11. *For $f \in R^n$, we have*

$$\omega(\mathrm{ch}(f)) = \mathrm{ch}(\varepsilon^n f). \tag{6.23}$$

(Note that $\varepsilon^n f$ is the pointwise product, and not the $\star$ product.)

Proof. It suffices to prove the result when $f = \chi^\lambda$, $\lambda \in \mathrm{Par}(n)$, since these span R^n. First note that, for $\rho, \lambda \in \mathrm{Par}(n)$, we have

$$\chi^{\lambda'}_\rho \overset{(6.20)}{=} \langle s_{\lambda'}, p_\rho \rangle \overset{(3.31)}{=} \langle \omega(s_{\lambda'}), \omega(p_\rho) \rangle \underset{(2.45)}{\overset{(3.17)}{=}} (-1)^{n-\ell(\rho)} \langle s_\lambda, p_\rho \rangle \underset{(6.20)}{\overset{(6.15)}{=}} \varepsilon^n_\rho \chi^\lambda_\rho.$$

Thus $\chi^{\lambda'} = \varepsilon^n \chi^\lambda$ and so

$$\omega\left(\mathrm{ch}(\chi^\lambda)\right) = \omega(s_\lambda) = s_{\lambda'} = \mathrm{ch}\left(\chi^{\lambda'}\right) = \mathrm{ch}\left(\varepsilon^n \chi^\lambda\right),$$

as desired. □

We saw in Section 4.4 that the ring of symmetric functions can be naturally endowed with the structure of a Hopf algebra. Using Theorem 6.3.4, we can deduce the corresponding structure of a Hopf algebra on R, so that the characteristic map is also an isomorphism of Hopf algebras.

Define linear maps

$$\Delta\colon R \to R \otimes R, \qquad \Delta f = \sum_{k=0}^{n} \mathrm{Res}^{\mathfrak{S}_n}_{\mathfrak{S}_k \times \mathfrak{S}_{n-k}}(f), \tag{6.24}$$

$$\varepsilon\colon R \to \mathbb{Z}, \qquad \varepsilon(f) = \langle f, 1 \rangle, \tag{6.25}$$

$$S\colon R \to R, \qquad S(f) = (-1)^n \varepsilon^n f. \tag{6.26}$$

for $f \in R^n$, then extended by linearity. Note that we view Δf as an element of $R \otimes R$ using the isomorphism (6.5).

Proposition 6.3.12. *The maps Δ, ε, and S endow R with the structure of a Hopf algebra, and the characteristic map (6.12) is an isomorphism of Hopf algebras.*

Proof. Since we already know from Theorem 6.3.4 that the characteristic map is an isometric algebra isomorphism of R onto Sym, it suffices to show that the maps Δ, ε, and S, as defined in (6.24) to (6.26), correspond, under the characteristic map, to the comultiplication, counit, and antipode of Sym. (Note that $\varepsilon(f) = \langle f, 1 \rangle \in \mathbb{Z}$, since the inner product of two virtual characters is an integer.)

As in (4.6), we define an inner product on $R \otimes R$ by

$$\langle f_1 \otimes f_2, g_1 \otimes g_2 \rangle = \langle f_1, g_1 \rangle \langle f_2, g_2 \rangle.$$

Then, for $f \in R^n$, $g \in R^k$, $h \in R^{n-k}$, we have

$$\begin{aligned}
&\langle (\mathrm{ch} \otimes \mathrm{ch}) \circ \Delta f, \mathrm{ch}(g) \otimes \mathrm{ch}(h) \rangle \\
&\qquad = \langle \Delta f, g \otimes h \rangle && \text{(since ch is an isometry)} \\
&\qquad \overset{(6.24)}{=} \left\langle \mathrm{Res}^{\mathfrak{S}_n}_{\mathfrak{S}_k \times \mathfrak{S}_{n-k}}(f), g \otimes h \right\rangle \\
&\qquad = \left\langle f, \mathrm{Ind}^{\mathfrak{S}_n}_{\mathfrak{S}_k \times \mathfrak{S}_{n-k}}(g \otimes h) \right\rangle && \text{(by Proposition 6.2.4)} \\
&\qquad \overset{(6.9)}{=} \langle f, g \star h \rangle \\
&\qquad = \langle \mathrm{ch}(f), \mathrm{ch}(g \star h) \rangle && \text{(since ch is an isometry)} \\
&\qquad = \langle \mathrm{ch}(f), \mathrm{ch}(g)\,\mathrm{ch}(h) \rangle && \text{(since ch is a ring homomorphism)} \\
&\qquad \overset{(4.33)}{=} \langle \Delta \circ \mathrm{ch}(f), \mathrm{ch}(g) \otimes \mathrm{ch}(h) \rangle.
\end{aligned}$$

Thus

$$(\mathrm{ch} \otimes \mathrm{ch}) \circ \Delta f = \Delta \circ \mathrm{ch}(f).$$

(Here we use that, by Theorem 6.3.4, the elements $\mathrm{ch}(g) \otimes \mathrm{ch}(h)$, $g \in R^k$, $h \in R^{n-k}$, span $\mathrm{Sym}^k \otimes \mathrm{Sym}^{n-k}$, on which the inner product (4.6) is nondegenerate, since the $s_\lambda \otimes s_\mu$ form an orthonormal basis.) We also have

$$\begin{aligned}
\varepsilon \circ \mathrm{ch}(f) &= \langle \mathrm{ch}(f), 1 \rangle && \text{(since } \langle \mathrm{ch}(f), 1 \rangle \text{ is the constant term of } \mathrm{ch}(f)) \\
&= \langle f, 1 \rangle && \text{(since ch is an isometry)} \\
&\overset{(6.25)}{=} \varepsilon(f) \\
&= \mathrm{ch} \circ \varepsilon(f) && \text{(since ch is an algebra homomorphism and } \varepsilon(f) \in \mathbb{Z}).
\end{aligned}$$

Finally, for $f \in R^n$,

$$S \circ \mathrm{ch}(f) = (-1)^n \omega(\mathrm{ch}(f)) \overset{(6.23)}{=} (-1)^n \,\mathrm{ch}(\varepsilon^n f) \overset{(6.26)}{=} \mathrm{ch} \circ S(f),$$

where, in the first equality, we used Exercise 4.4.3. □

Exercises.

6.3.1. Prove that, with the multiplication $\star$ of (6.9), R is a commutative, associative, unital, graded ring.

Hint: Suppose $K \subseteq H \subseteq G$ and G' are finite groups, $f \in \mathcal{C}(K, Z)$, and $f' \in \mathcal{C}(G', Z)$. First show, directly from (6.7), that

(a) $\mathrm{Ind}_H^G\left(\mathrm{Ind}_K^H(f)\right) = \mathrm{Ind}_K^G(f)$ (transitivity of induction);

(b) $\mathrm{Ind}_K^H(f) \times f' = \mathrm{Ind}_{K\times G'}^{H\times G'}(f \times f')$, and similarly in the other factor;

(c) if $\sigma \in G$ and ${}^\sigma f \in \mathcal{C}(\sigma K \sigma^{-1}, Z)$ is defined by ${}^\sigma f(x) = f(\sigma^{-1} x \sigma)$, then $\mathrm{Ind}_{\sigma K \sigma^{-1}}^G({}^\sigma f) = \mathrm{Ind}_K^G(f)$.

Then show that, for $f \in R^a$, $g \in R^b$, and $h \in R^c$, both $(f \star g) \star h$ and $f \star (g \star h)$ are equal to $\mathrm{Ind}_{\mathfrak{S}_a \times \mathfrak{S}_b \times \mathfrak{S}_c}^{\mathfrak{S}_{a+b+c}}(f \times g \times h)$, and deduce commutativity from the third fact, applied to the permutation that interchanges the two blocks. See also [Gri26, Prop. 2.3].

6.3.2. Prove that if $\pi \in \mathfrak{S}_n$ has cycle type ρ, then $\varepsilon_n(\pi) = (-1)^{n-\ell(\rho)}$, were ε_n is the character of the sign representation of $\mathfrak{S}_n$.

6.3.3. Compute the character table of $\mathfrak{S}_3$.

6.3.4. Compute the character table of $\mathfrak{S}_4$.

6.3.5. Let G_n be the cyclic group of order n.

(a) What are the conjugacy classes of G_n?

(b) Describe all the irreducible representations of G_n.

(c) Show that, for $n > 2$, the characters of G_n are not all real-valued.

6.4 Specht modules

It follows from Theorem 6.3.6 that the irreducible $\mathfrak{S}_n$-module are enumerated by partitions of n. Our goal in this section is to construct these modules explicitly. We follow the procedure outlined in [Mac15, Example 15, p. 124]. A priori, we should work over the complex numbers to find all irreducible representations of a finite group. However, as we will see, the symmetric groups have the remarkable property that we can construct all of their irreducible representations over the integers. We first work over $\mathbb{Q}$, and find enough irreducible representations to know that we have found them all. We then show that one can, in fact, work over $\mathbb{Z}$.

It follows from our knowledge of transition matrices (Theorem 3.5.5, Lemma 3.5.2, and Proposition 3.5.6(a)) that

$$h_\lambda = s_\lambda + \sum_{\mu>\lambda} K_{\mu\lambda} s_\mu \qquad \text{and} \qquad e_{\lambda'} = s_\lambda + \sum_{\mu<\lambda} K_{\mu'\lambda'} s_\mu.$$

Applying the inverse of the characteristic map ch from (6.12) and using (6.19), we have

$$\eta^\lambda = \chi^\lambda + \sum_{\mu>\lambda} K_{\mu\lambda} \chi^\mu \qquad \text{and} \qquad \varepsilon^{\lambda'} = \chi^\lambda + \sum_{\mu<\lambda} K_{\mu'\lambda'} \chi^\mu.$$

These equations give the decomposition of the induced modules

$$H_\lambda := \mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n}(\mathrm{triv}_\lambda) \qquad \text{and} \qquad E_{\lambda'} := \mathrm{Ind}_{\mathfrak{S}_{\lambda'}}^{\mathfrak{S}_n}(\mathrm{sign}_{\lambda'})$$

into irreducible modules. Here, for $\mu = (\mu_1, \dots, \mu_k) \in \mathrm{Par}$, triv_μ denotes the trivial $\mathfrak{S}_\mu$-module and sign_μ denotes the sign $\mathfrak{S}_\mu$-module, that is, the one-dimensional $\mathfrak{S}_\mu$-module with character

$$\mathfrak{S}_\mu = \mathfrak{S}_{\mu_1} \times \cdots \times \mathfrak{S}_{\mu_k} \to \mathbb{C}, \qquad (\pi_1, \dots, \pi_k) \mapsto \prod_{i=1}^{k} \mathrm{sgn}(\pi_i).$$

It follows that the irreducible $\mathfrak{S}_n$-module with character χ^λ is the unique common irreducible component of H_λ and $E_{\lambda'}$. We would like to describe this module explicitly.

If A is a ring and $x, y \in A$, then Ay is an A-module with action given by left multiplication, and Axy is a submodule of Ay. In particular, Axy is the image of Ax under the A-module homomorphism

$$A \to Ay, \quad z \mapsto zy, \qquad z \in A.$$

Thus Axy is also a quotient of Ax.

Let $\lambda \in \mathrm{Par}(n)$, and let T be any standard tableau of shape λ. Thus the entries of T are $1, 2, \dots, n$. Let $\mathfrak{R}$ (respectively $\mathfrak{C}$) be the subgroup of $\mathfrak{S}_n$ that stabilizes each row (respectively, column) of T. (We say that $\pi \in \mathfrak{S}_n$ *stabilizes* a subset $X \subseteq \{1, 2, \dots, n\}$ if $\pi(X) = X$.) Thus we have group isomorphisms

$$\mathfrak{R} \cong \mathfrak{S}_\lambda \qquad \text{and} \qquad \mathfrak{C} \cong \mathfrak{S}_{\lambda'}.$$

Let $A = \mathbb{Q}\mathfrak{S}_n$ be the group algebra of $\mathfrak{S}_n$ and define $s, a \in A$ by

$$s = \sum_{\pi \in \mathfrak{R}} \pi \qquad \text{and} \qquad a = \sum_{\pi \in \mathfrak{C}} \mathrm{sgn}(\pi)\pi. \tag{6.27}$$

Then $(\mathbb{Q}\mathfrak{R})s = \mathbb{Q}s$ is the trivial $\mathbb{Q}\mathfrak{R}$-module, and we leave it as Exercise 6.4.2 to show that we have an isomorphism of $\mathfrak{S}_n$-modules

$$H_\lambda = A \otimes_{\mathbb{Q}\mathfrak{R}} \mathbb{Q}s \to As, \quad x \otimes cs \mapsto xcs, \quad x \in A,\ c \in \mathbb{Q}. \tag{6.28}$$

Similarly, we leave it as Exercise 6.4.3 to show that we have an isomorphism of $\mathfrak{S}_n$-modules

$$E_{\lambda'} = A \otimes_{\mathbb{Q}\mathfrak{C}} \mathbb{Q}a \to Aa, \quad x \otimes ca \mapsto xca, \quad x \in A,\ c \in \mathbb{Q}. \tag{6.29}$$

Now let

$$e = as \in A.$$

Any element of $\mathfrak{R} \cap \mathfrak{C}$ stabilizes each row *and* column of T and thus must be the identity. In other words $\mathfrak{R} \cap \mathfrak{C} = \{1\}$ is the trivial subgroup of $\mathfrak{S}_n$. It follows that the products $\sigma\pi$, $\sigma \in \mathfrak{C}$, $\pi \in \mathfrak{R}$, are all distinct (Exercise 6.4.4). Thus

$$e = \sum_{\substack{\sigma \in \mathfrak{C} \\ \pi \in \mathfrak{R}}} \mathrm{sgn}(\sigma)\sigma\pi \neq 0.$$

It then follows from the above discussion that

$$M_\lambda := Ae$$

is a nonzero submodule of $As \cong H_\lambda$ that is isomorphic to a quotient of $Aa \cong E_{\lambda'}$. Thus M_λ is a common irreducible component of H_λ and $E_{\lambda'}$ and so it is the irreducible $\mathfrak{S}_n$-module with character χ^λ.

Our goal is to give an explicit construction of the irreducible modules M_λ, $\lambda \in \mathrm{Par}$, by identifying M_λ with a space of polynomials. For a tableau T of size n and weight (1^n), let

$$m_T := x_1^{d(1)} \cdots x_n^{d(n)} \in \mathbb{Q}[x_1, \ldots, x_n],$$

where $d(i) = r - 1$ if the entry i lies in the r-th row of T. Then define

$$f_T := \prod_{(i,j)} (x_i - x_j),$$

where the product is taken over all pairs (i, j) such that j lies above, and in the same column as, i in the tableau T. In other words, f_T is the product of the Vandermonde determinants (see (3.5)) corresponding to the columns of T. It then follows as in Lemma 3.1.3 and Exercise 3.1.1 that

$$f_T = \sum_{\pi \in \mathfrak{C}} \mathrm{sgn}(\pi) \pi m_T = a m_T. \tag{6.30}$$

Recall from Section 2.3 that $\mathrm{ST}(\lambda)$ denotes the set of standard tableaux of shape λ.

Theorem 6.4.1. *For $\lambda \in \mathrm{Par}(n)$, the subspace*

$$\mathrm{Span}_{\mathbb{Q}}\{f_T : T \text{ is a tableau of shape } \lambda \text{ and weight } (1^n)\} \subseteq \mathrm{Pol}_n \tag{6.31}$$

is an $\mathfrak{S}_n$-submodule of Pol_n isomorphic to the irreducible module M_λ. Furthermore, the polynomials f_T, $T \in \mathrm{ST}(\lambda)$, form a basis of this $\mathfrak{S}_n$-module.

Proof. Consider the map

$$A \to \mathbb{Q}[x_1, \ldots, x_n], \qquad u \mapsto u m_T. \tag{6.32}$$

Since $d(i) = d(j)$ if and only if i and j lie in the same row of T, it follows that $s m_T = |\mathfrak{R}| m_T$. Restriction of (6.32) to Ae thus yields a surjective homomorphism of A-modules

$$\theta \colon M_\lambda = Aas \to Aasm_T = Aam_T \overset{(6.30)}{=} Af_T.$$

Since M_λ is a simple A-module, $\ker(\theta)$ is an A-submodule of M_λ, and θ is not the zero map, it follows that $\ker(\theta) = 0$. Thus θ is an isomorphism of $\mathfrak{S}_n$-modules.

It remains to prove that the f_T, $T \in \mathrm{ST}(\lambda)$, form a basis of (6.31). We know from Theorem 6.3.6(b) that the dimension of M_λ is the number of standard tableaux of shape λ. So it suffices to show that the f_T, $T \in \mathrm{ST}(\lambda)$, are linearly independent.

Define an ordering on monomials x^α, $\alpha \in \mathbb{N}^n$, by declaring that $x^\alpha < x^\beta$ if α precedes β in the lexicographic order on $\mathbb{N}^n$. That is,

$$x^\alpha < x^\beta \iff \exists\, i \in \{1, \ldots, n\} \text{ such that } \alpha_1 = \beta_1, \ldots,\ \alpha_{i-1} = \beta_{i-1},\ \alpha_i < \beta_i.$$

Now suppose that T is a standard tableaux of shape λ and $\pi \in \mathfrak{C}$. Let $i \in \{1, 2, \ldots, n\}$ be the smallest element such that $\pi(i) \neq i$. Then $i < \pi^{-1}(i)$, and so $d(i) < d(\pi^{-1}(i))$, since T is standard. Thus

$$m_T = x_1^{d(1)} \cdots x_n^{d(n)} < x_1^{d(1)} \cdots x_{i-1}^{d(i-1)} x_i^{d(\pi^{-1}(i))} \cdots x_n^{d(\pi^{-1}(n))} = \pi m_T.$$

It follows that

$$f_T \overset{(6.30)}{=} \sum_{\pi \in \mathfrak{C}} \operatorname{sgn}(\pi) \pi m_T = m_T + \sum_{\alpha} d_\alpha x^\alpha \tag{6.33}$$

where $d_\alpha \in \mathbb{Q}$ and the sum is over α such that $m_T < x^\alpha$.

Let $T_1, \ldots, T_r$ be the standard tableaux of shape λ. Since the monomials $m_{T_1}, \ldots, m_{T_r}$ are all distinct, we may choose the order of the T_i such that

$$m_{T_1} < m_{T_2} < \cdots < m_{T_r}.$$

Suppose we have a nontrivial linear dependence relation

$$\sum_{i=1}^{r} c_i f_{T_i} = 0. \tag{6.34}$$

Let i be the smallest index such that $c_i \neq 0$. It follows from (6.33) that the coefficient of m_{T_i} in the left-hand side of (6.34) is equal to c_i. But then the left-hand side cannot be equal to zero. This contradiction shows that the $f_{T_1}, \ldots, f_{T_r}$ are linearly independent, as desired. □

The realization of M_λ given in Theorem 6.4.1 is called the *Specht module* corresponding to the partition λ. From now on, we use the notation M_λ to denote this specific realization.

Remark 6.4.2. Although we worked in Theorem 6.4.1 over the field $\mathbb{Q}$, the last part of its proof, showing that the f_T, $T \in \mathrm{ST}(\lambda)$, are linearly independent, remains valid over *any* field. Indeed, the f_T have integer coefficients and, by (6.33), m_T appears in f_T with coefficient 1, while all other monomials appearing in f_T are larger than m_T. Thus, if $\Bbbk$ is any field, then the f_T, $T \in \mathrm{ST}(\lambda)$, are linearly independent elements of $\Bbbk[x_1, \ldots, x_n]$. (By contrast, the first part of the proof of Theorem 6.4.1 uses the simplicity of M_λ, which relies on working in characteristic zero.) △

Proposition 6.4.3. *Let $\lambda \in \mathrm{Par}(n)$. For each $\pi \in \mathfrak{S}_n$, the matrix representing the action of π on M_λ in the basis f_T, $T \in \mathrm{ST}(\lambda)$, has integer entries.*

Proof. As in the proof of Theorem 6.4.1, let $T_1, \ldots, T_r$ denote the standard tableaux of shape λ. It follows from Theorem 6.4.1 that, for all tableaux of shape λ, f_T is a linear combination of the $f_{T_1}, \ldots, f_{T_r}$. Suppose

$$f_T = \sum_{i=1}^{r} c_i f_{T_i}, \qquad c_1, \ldots, c_r \in \mathbb{Q}.$$

Let m be the lowest common denominator of the c_i. Then, for each $i \in \{1, \dots, r\}$, we have $c_i = m_i/m$ for some $m_i \in \mathbb{Z}$, and so

$$mf_T = \sum_{i=1}^{r} m_i f_{T_i}. \tag{6.35}$$

If $m > 1$, then choose a prime number p dividing m and reduce (6.35) modulo p. Since not all the m_i are divisible by p, it follows that the f_{T_i} are linearly dependent over the field $\mathbb{Z}/p\mathbb{Z}$. But this contradicts Remark 6.4.2.

It follows from the above that, if T is any tableau of shape λ, then f_T is an integer linear combination of $f_{T_1}, \dots, f_{T_r}$. Since $\pi f_{T_i} = f_T$, where T is obtained from T_i by permuting its entries by π, the proposition follows. □

Remark 6.4.4. It follows from Remark 6.4.2 that the f_T, $T \in \mathrm{ST}(\lambda)$, are linearly independent over $\mathbb{Z}$ (since any linear dependence relation over $\mathbb{Z}$ is a linear dependence relation over the field $\mathbb{Q}$). Thus, for $\lambda \in \mathrm{Par}$,

$$\mathrm{Span}_{\mathbb{Z}}\{f_T : T \in \mathrm{ST}(\lambda)\}$$

is a free $\mathbb{Z}$-module of rank equal to the number of standard tableaux of shape λ. By Proposition 6.4.3, the ring $\mathbb{Z}\mathfrak{S}_n$ acts on this space. Hence, for any field $\Bbbk$, the f_T, $T \in \mathrm{ST}(\lambda)$, form a basis of a $\Bbbk\mathfrak{S}_n$-submodule of $\Bbbk[x_1, \dots, x_n]$. In this way, the Specht modules are defined over the integers. We can then reduce modulo any prime p to obtain modules for$(\mathbb{Z}/p\mathbb{Z})\mathfrak{S}_n$.

Unlike the situation in characteristic zero, the Specht modules can be reducible in positive characteristic. For a prime number p, we say that a partition λ is *p-regular* if it does not have p nonzero parts of the same size. If λ is a p-regular partition, then the Specht module M_λ, defined over $\mathbb{Z}/p\mathbb{Z}$, has a unique irreducible quotient, and these irreducible quotients form a complete set of irreducible representations. However, there remain many unanswered questions about the representation theory of the symmetric group in positive characteristic and this is an active area of research. For a short discussion of the positive characteristic case, see [Wil]. △

Exercises.

6.4.1. With s and a defined as in (6.27), show that

$$s^2 = |\mathfrak{R}|s \qquad \text{and} \qquad a^2 = |\mathfrak{C}|a.$$

6.4.2. Prove that (6.28) is an isomorphism of $\mathfrak{S}_n$-modules.

6.4.3. Prove that (6.29) is an isomorphism of $\mathfrak{S}_n$-modules.

6.4.4. Suppose that H and K are subgroups of a group G such that $H \cap K = \{1\}$. Show that the products xy, $x \in H$, $y \in K$, are all distinct.

6.5 Representations of general linear groups

In this final section, we explain how Schur functions appear as characters of simple modules for the general linear Lie algebra and general linear Lie group. Our treatment is very brief. We refer the reader to the many books on Lie theory (for example, [Hal15]) for further details.

Fix $n \in \mathbb{N}$. Throughout this section we work over the field $\mathbb{C}$ of complex numbers. The *general linear Lie algebra* $\mathfrak{gl}_n(\mathbb{C})$ is vector space of complex $n \times n$ matrices, together with a binary operation, called the *Lie bracket*, given by

$$[X, Y] = XY - YX, \qquad X, Y \in \mathfrak{gl}_n(\mathbb{C}).$$

A *representation* of $\mathfrak{gl}_n(\mathbb{C})$ is a complex vector space V, together with a linear map

$$\rho \colon \mathfrak{gl}_n(\mathbb{C}) \to \mathrm{End}_{\mathbb{C}}(V)$$

such that

$$\rho([X, Y]) = \rho(X)\rho(Y) - \rho(Y)\rho(X), \qquad X, Y \in \mathfrak{gl}_n(\mathbb{C}).$$

We define

$$Xv := \rho(X)(v), \qquad X \in \mathfrak{gl}_n(\mathbb{C}),\ v \in V.$$

Then V, together with this action, is called a $\mathfrak{gl}_n(\mathbb{C})$*-module*.

The *Cartan subalgebra* $\mathfrak{h}$ of $\mathfrak{gl}_n(\mathbb{C})$ is the Lie subalgebra consisting of diagonal matrices. (Note that the Lie bracket of any two diagonal matrices is zero. We say that $\mathfrak{h}$ is an *abelian* Lie algebra.) For $1 \leq i \leq n$, let

$$\varepsilon_i \colon \mathfrak{h} \to \mathbb{C}, \quad X = (X_{ij}) \mapsto X_{ii}$$

be the linear map that returns the (i, i)-entry of a diagonal matrix. Then define

$$\Lambda := \bigoplus_{i=1}^{n} \mathbb{Z}\varepsilon_i \subseteq \mathfrak{h}^*.$$

We identify Λ with $\mathbb{Z}^n$ by identifying $\sum_{i=1}^n \lambda_i \varepsilon_i$ with $(\lambda_1, \ldots, \lambda_n)$. In particular, $\mathrm{Par}_n \subseteq \Lambda$. If V is a $\mathfrak{gl}_n(\mathbb{C})$-module and $\lambda \in \mathfrak{h}^*$, then the λ*-weight space* of V is

$$V_\lambda := \{v \in V : Xv = \lambda(X)v \text{ for all } X \in \mathfrak{h}\}.$$

We say that V is a *weight module* if

$$V = \bigoplus_{\lambda \in \Lambda} V_\lambda.$$

An element $v \in V$ is a *highest-weight vector of weight* $\lambda \in \Lambda$ if $v \in V_\lambda$ and v is annihilated by all strictly upper triangular matrices. The module V is a *highest-weight module of highest weight* λ if it is generated by a highest weight vector of weight λ.

Every finite-dimensional weight $\mathfrak{gl}_n(\mathbb{C})$-module is a direct sum of simple modules (modules containing no nontrivial submodules). Furthermore, for each

$$\lambda \in \Lambda^+ := \{\lambda \in \Lambda : \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n\},$$

there exists a unique (up to isomorphism) simple highest-weight module V^λ of highest weight λ, and it is a finite-dimensional weight module. The V^λ, $\lambda \in \Lambda^+$, are a complete list of finite-dimensional simple weight $\mathfrak{gl}_n(\mathbb{C})$-modules, up to isomorphism. Thus, to understand the finite-dimensional weight modules of $\mathfrak{gl}_n(\mathbb{C})$, one needs only understand the V^λ, $\lambda \in \Lambda^+$.

Warning 6.5.1. The hypothesis that V is a weight module cannot be omitted. For $n = 1$, the Lie algebra $\mathfrak{gl}_1(\mathbb{C}) = \mathbb{C}$ is abelian, so every linear map $\mathbb{C} \to \mathrm{End}_{\mathbb{C}}(V)$ is a representation. The representation $z \mapsto \left(\begin{smallmatrix} 0 & z \\ 0 & 0 \end{smallmatrix}\right)$ of $\mathfrak{gl}_1(\mathbb{C})$ on $\mathbb{C}^2$ is not a direct sum of simple modules, and the one-dimensional representation $z \mapsto z/2$ is simple with highest weight $\frac{1}{2}\varepsilon_1 \notin \Lambda$.

If V is a finite-dimensional weight module, we define its *character* to be

$$\mathrm{ch}(V) = \sum_{\lambda \in \Lambda} (\dim_{\mathbb{C}} V_\lambda) x^\lambda \in \mathbb{Z}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}], \qquad \text{where } x^\lambda := x_1^{\lambda_1} \cdots x_n^{\lambda_n}.$$

Thus, the character of V records the dimensions of the weight spaces of V. We extend the definition (3.1) of a_α to all $\alpha \in \mathbb{Z}^n$, so that $a_\alpha \in \mathbb{Z}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$. The *Weyl character formula* for $\mathfrak{gl}_n(\mathbb{C})$ states that, for $\lambda \in \Lambda^+$,

$$\mathrm{ch}\left(V^\lambda\right) = \frac{\sum_{\pi \in \mathfrak{S}_n} \mathrm{sgn}(\pi) x^{(\lambda+\delta^n)\pi}}{\sum_{\pi \in \mathfrak{S}_n} \mathrm{sgn}(\pi) x^{(\delta^n)\pi}} = \frac{a_{\lambda+\delta^n}}{a_{\delta^n}} = (x_1 \cdots x_n)^{\lambda_n} s_{\lambda - \lambda_n(1^n)}(x_1, \ldots, x_n), \tag{6.36}$$

where the last equality follows from (3.7), since $\lambda - \lambda_n(1^n) = (\lambda_1 - \lambda_n, \ldots, \lambda_{n-1} - \lambda_n, 0) \in \mathrm{Par}_n$ and $a_{\alpha + c(1^n)} = (x_1 \cdots x_n)^c a_\alpha$ for all $\alpha \in \mathbb{Z}^n$ and $c \in \mathbb{Z}$. In particular, if $\lambda \in \mathrm{Par}_n$, then the character of the simple module with highest weight λ is precisely the Schur polynomial $s_\lambda(x_1, \ldots, x_n)$.

Remark 6.5.2. The Weyl character formula is a much more general result that holds for other Lie algebras. We have presented here the special case of this formula for $\mathfrak{gl}_n(\mathbb{C})$. In fact, we have presented a form that uses the *Weyl denominator formula*. In Lie theory, one usually uses the Weyl vector $\rho = \frac{1}{2}\sum_{i<j}(\varepsilon_i - \varepsilon_j) = \left(\frac{n-1}{2}, \frac{n-3}{2}, \ldots, \frac{1-n}{2}\right)$ in place of δ^n. Since $\delta^n = \rho + \frac{n-1}{2}(1^n)$, this multiplies the numerator and denominator of the quotient in (6.36) by the same factor $(x_1 \cdots x_n)^{-(n-1)/2}$, and so does not change it. △

The *general linear group* $\mathrm{GL}_n(\mathbb{C})$ is the group of complex $n \times n$ invertible matrices. We have a surjective *exponential map*

$$\exp\colon \mathfrak{gl}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C}), \quad \exp(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!}.$$

This map allows one to move between representations of $\mathfrak{gl}_n(\mathbb{C})$ and $\mathrm{GL}_n(\mathbb{C})$, as follows. If $\tilde{\rho}\colon \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}(V)$ is a finite-dimensional holomorphic representation, then differentiating gives a representation ρ of $\mathfrak{gl}_n(\mathbb{C})$ on V such that $\tilde{\rho}(\exp X) = \exp(\rho(X))$ for all

$X \in \mathfrak{gl}_n(\mathbb{C})$; since exp is surjective, $\tilde{\rho}$ is determined by ρ. The $\mathfrak{gl}_n(\mathbb{C})$-modules that arise in this way are precisely the finite-dimensional weight modules. In particular, the simple holomorphic representations of $\mathrm{GL}_n(\mathbb{C})$ are indexed by Λ^+. A finite-dimensional representation $\rho\colon \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}(V)$ is *polynomial* if, after choosing a basis of V, every matrix entry of $\rho(g)$ is a polynomial in the entries g_{ij}. The simple polynomial representations of $\mathrm{GL}_n(\mathbb{C})$ are those corresponding to the V^λ with $\lambda \in \Lambda^+ \cap \mathbb{N}^n = \mathrm{Par}_n$; see [Mac15, App. A, §8]. If

$$\rho_\lambda\colon \mathfrak{gl}_n(\mathbb{C}) \to \mathrm{End}_{\mathbb{C}}(V^\lambda)$$

is the representation of $\mathfrak{gl}_n(\mathbb{C})$ corresponding to $\lambda \in \mathrm{Par}_n$, and $H = \mathrm{diag}(h_1, \ldots, h_n) \in \mathfrak{gl}_n(\mathbb{C})$ is a diagonal matrix, then $\rho_\lambda(H)$ acts on the weight space V^λ_μ as the scalar $\mu(H) = \sum_{i=1}^n \mu_i h_i$. Hence it follows from (6.36) that

$$\mathrm{tr}\left(e^{\rho_\lambda(H)}\right) = \sum_{\mu\in\Lambda} \left(\dim_{\mathbb{C}} V^\lambda_\mu\right) e^{\mu(H)} = s_\lambda\left(e^{h_1}, \ldots, e^{h_n}\right). \tag{6.37}$$

If $e^{h_1}, \ldots, e^{h_n}$ are pairwise distinct, then $\sum_{\pi\in\mathfrak{S}_n} \mathrm{sgn}(\pi) e^{(\delta^n)^\pi(H)} = a_{\delta^n}(e^{h_1}, \ldots, e^{h_n})$ is nonzero by (3.5), and

$$\mathrm{tr}\left(e^{\rho_\lambda(H)}\right) = \frac{\sum_{\pi\in\mathfrak{S}_n} \mathrm{sgn}(\pi) e^{(\lambda+\delta^n)^\pi(H)}}{\sum_{\pi\in\mathfrak{S}_n} \mathrm{sgn}(\pi) e^{(\delta^n)^\pi(H)}}.$$

This expression is sometimes also referred to as the character of ρ_λ. (Compare to the definition of character from Section 6.1.)

Index of notation

$\cup$, 69
$\frac{\partial}{\partial e_n}$, 88
$\frac{\partial}{\partial h_n}$, 88
$\frac{\partial}{\partial p_n}$, 88
$\lambda \supseteq \mu$, 58
$\langle\ ,\ \rangle$, 51, 113, 135
$\langle\ ,\ \rangle_G^{\mathbb{C}}$, 127
$\leq$, 8, 9
$\leq_R$, 69
$\otimes$, 81
$\star$, 134
$\vdash$, 6

$\mathbf{A}$, 116
a_α, 43
$\mathbf{A}^{(c)}$, 115
α^π, 8
Alt_n, 43

$B^\pm(t)$, 119
$B_n^\pm$, 119
$\mathrm{BST}(\lambda, \rho)$, 139

$\mathfrak{C}$, 143
$\mathcal{C}(G, Z)$, 127
ch, 135
χ^λ, 136
χ_ρ, 127
Cl, 108
C_ρ, 134

Δ, 92, 140
δ^n, 44

E_λ, 142
ε, 70, 92
ε^λ, 136
ε^n, 136
ε, 140
e_r, 29
$e_r(x_1, \ldots, x_n)$, 29
$E(t)$, 30
$E^\pm(t)$, 117
η, 91
η^λ, 136
η^n, 136

$\mathcal{F}$, 109
$f^\perp$, 86
$\mathcal{F}^{(c)}$, 109
$\mathcal{F}(G, Z)$, 127
f_λ, 35
πf, 21
f_ρ, 136
f_T, 144

$\mathrm{GL}_n(\mathbb{C})$, 148
$\mathfrak{gl}_n(\mathbb{C})$, 147
G^{op}, 10

Heis, 102
H_λ, 142
h_λ, 35
h_r, 33
$h_r(x_1, \ldots, x_n)$, 34
$H^\pm(t)$, 117
$\mathrm{ht}(T)$, 139
$\mathrm{ht}(\lambda)$, 78

I, 109
$I^{(c)}$, 109
Ind_H^G, 131, 132
$\mathrm{inv}(\pi)$, 42

J, 65, 112

K, 67
$K_{\lambda,\mu}$, 63

$K_{\lambda/\mu,\nu}$, 63
$\Bbbk G$, 126

L, 69
λ', 7
$|\lambda|$, 6
λ/μ, 58
$\ell(\lambda)$, 7

$M^{-\intercal}$, 66
$M_{\lambda,\mu}(e,m)$, 30
$M_{\lambda,\mu}(h,m)$, 36
$M_{\lambda,\mu}(p,m)$, 40
$m_i(\lambda)$, 7
M_λ, 143
m_λ, 25
$m_\lambda(x_1,\ldots,x_n)$, 21
m_T, 144
$M(u,v)$, 66

$\mathbb{N}$, 5
$[n]$, 5
∇, 91

Par, 6
Par_k, 7
$\mathrm{Par}_k(n)$, 7
$\mathrm{Par}(k)$, 6
$p(n)$, 16
$p(n;k)$, 16
Pol_n, 20
Pol_n^k, 20
P_r, 114
p_r, 37
Ψ, 135
$\Psi^{\pm}(t)$, 122
$\psi_i^{\pm}$, 108
$\Psi_i^{\pm}$, 122
$P(t)$, 38

R, 134
$R^\times$, 15
$\mathfrak{R}$, 143
Res_H^G, 130
$\mathcal{R}(G,\mathbb{C})$, 129
$\rho_{m,n}$, 23
ρ_n, 25
ρ_n^k, 24
R^n, 134
RowStrict, 28
RowWeak, 28
$R[\![t]\!]$, 14
R^X, 13
$R[\![x_1,x_2,\ldots]\!]$, 17

S, 93, 140
$\mathrm{shape}(T)$, 28
$\mathfrak{S}_\infty$, 6
$\mathfrak{S}_\lambda$, 136
s_λ, 47
$s_{\lambda/\mu}$, 56
$s_\lambda(x_1,\ldots,x_n)$, 46
$\mathfrak{S}_n$, 5
$\mathrm{sgn}(\pi)$, 42
$\mathrm{SST}(\lambda)$, 28
$\mathrm{SST}(\lambda,n)$, 28
$\mathrm{ST}(\lambda)$, 28
Sym, 25
Sym^+, 102
Sym^-, 102
Sym^k, 24
Sym_n, 21
Sym_n^k, 21
$\mathrm{Sym}_{\mathbb{Q}}$, 27

T, 122

UT_n, 65

$v_\lambda^{(c)}$, 110
$v_{\mathbf{i}}$, 109

ω, 34
WComp, 8
$\mathrm{WComp}(k)$, 8
WComp_n, 8
$\mathrm{WComp}_n(k)$, 8
$\mathrm{word}(T)$, 76
$\mathrm{wt}(T)$, 28

x^α, 16, 20

z, 70
z_λ, 40

Bibliography

[DF04] D. S. Dummit and R. M. Foote. *Abstract algebra*. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2004.

[DNR01] S. Dăscălescu, C. Năstăsescu, and Ş. Raianu. *Hopf algebras: An introduction*, volume 235 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 2001.

[Egg19] E. S. Egge. *An introduction to symmetric functions and their combinatorics*, volume 91 of *Student Mathematical Library*. American Mathematical Society, Providence, RI, 2019. doi:10.1090/stml/091.

[GAE02] Z. Gong, M. Aldeen, and L. Elsner. A note on a generalized Cramer's rule. *Linear Algebra Appl.*, 340:253–254, 2002. doi:10.1016/S0024-3795(01)00469-4.

[Gri26] Darij Grinberg. The representation theory of somewhere-to-below shuffles, 2026. Version 4, September 23, 2026. URL: https://www.cip.ifi.lmu.de/~grinberg/algebra/s2b3.pdf.

[Hal15] B. Hall. *Lie groups, Lie algebras, and representations*, volume 222 of *Graduate Texts in Mathematics*. Springer, Cham, second edition, 2015. An elementary introduction. doi:10.1007/978-3-319-13467-3.

[Mac15] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Classic Texts in the Physical Sciences. The Clarendon Press, Oxford University Press, New York, second edition, 2015. With contribution by A. V. Zelevinsky and a foreword by Richard Stanley, Reprint of the 2008 paperback edition [MR1354144].

[Sav] A. Savage. MAT 5141 – Algebra I, Lecture notes. URL: https://alistairsavage.ca/mat5141/notes/MAT5141-Algebra_I.pdf.

[Ste02] J. R. Stembridge. A concise proof of the Littlewood-Richardson rule. *Electron. J. Combin.*, 9(1):Note 5, 4, 2002. URL: http://www.combinatorics.org/Volume_9/Abstracts/v9i1n5.html.

[Wil] M. Wildon. Representation theory of the symmetric group. URL: http://www.ma.rhul.ac.uk/~uvah099/Maths/Sym/SymGroup2014.pdf.