Chapter 3
Schur Functions

Abstract As a warmup for our discussion of Macdonald polynomials, we review fundamental properties of Schur functions. We start here with two definitions of the Schur functions, one by combinatorics of semi-standard tableaux, and the other in terms of ratios of Vandermonde-type determinants. Then we establish the equivalence of the two definitions by means of the Cauchy formula. It should be noted that the theory of Macdonald polynomials is modeled in many respects on that of Schur functions.

3.1 Definitions of the Schur Functions

3.1.1 Two Definitions

We now move on to the *Schur functions* $s_\lambda(x)$ $(\lambda \in \mathcal{P}_n)$; they are a family of symmetric polynomials indexed by the same set $\mathcal{P}_n$ of partitions λ with $\ell(\lambda) \le n$ as in the case of $m_\lambda(x)$. Each $s_\lambda(x)$ is homogeneous of degree $|\lambda|$ and has the leading term x^λ with respect to the dominance order:

$$s_\lambda(x) = x^\lambda + \cdots = m_\lambda(x) + \cdots . \tag{3.1}$$

With this property, they also form a $\mathbb{C}$-basis of the ring of symmetric polynomials:

$$\mathbb{C}[x]^{\mathfrak{S}_n} = \bigoplus_{\lambda \in \mathcal{P}_n} \mathbb{C}\, s_\lambda(x). \tag{3.2}$$

As we will see below, $s_\lambda(x)$ are in fact symmetric polynomials with nonnegative integer coefficients, i.e. $s_\lambda(x) \in \mathbb{N}[x]^{\mathfrak{S}_n}$.

We give two definitions of the Schur functions here, denoting them by $s_\lambda^{\mathrm{comb}}(x)$ and $s_\lambda^{\mathrm{det}}(x)$ respectively, and show later that they in fact coincide.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

M. Noumi, *Macdonald Polynomials*, SpringerBriefs in Mathematical Physics,
https://doi.org/10.1007/978-981-99-4587-0_3

Definition 3.1 (*combinatorial*) For each $\lambda \in \mathcal{P}_n$, we define the Schur function $s_\lambda^{\mathrm{comb}}(x)$ as the sum

$$s_\lambda^{\mathrm{comb}}(x) = \sum_{T \in \mathrm{SSTab}_n(\lambda)} x^{\mathrm{wt}(T)} \tag{3.3}$$

of monomials $x^{\mathrm{wt}(T)}$ over the set $\mathrm{SSTab_n}(\lambda)$ of all *semi-standard tableaux* T of shape λ in letters $\{1, \dots, n\}$.

We explain below the precise meaning of a semi-standard tableau T and its weight $\mathrm{wt}(T)$. By definition we have $s_\lambda^{\mathrm{comb}}(x) \in \mathbb{N}[x]$, but it is not obvious why it should be symmetric since this definition depends strongly on the ordering of the indexing set $\{1, \dots, n\}$.

Definition 3.2 (*determinantal*) For each $\lambda \in \mathcal{P}_n$, we define the Schur function $s_\lambda^{\det}(x)$ as the ratio of two determinants of Vandermonde type:

$$s_\lambda^{\det}(x) = \frac{\det\left(x_i^{\lambda_j+n-j}\right)_{i,j=1}^n}{\det\left(x_i^{n-j}\right)_{i,j=1}^n} = \frac{\det\left(x_i^{\lambda_j+n-j}\right)_{i,j=1}^n}{\Delta(x)}, \tag{3.4}$$

where $\Delta(x) = \prod_{1 \le i < j \le n}(x_i - x_j)$ stands for the difference product.

Since the numerator $\det\left(x_i^{\lambda_j+n-j}\right)_{i,j=1}^n \in \mathbb{Z}[x]$ is an alternating polynomial in $\mathbb{Z}[x]^{\mathfrak{S}_n,\mathrm{sgn}}$, it is divisible by $\Delta(x)$ in the polynomial ring $\mathbb{Z}[x]$ with integer coefficients. Hence the resulting $s_\lambda^{\det}(x)$ is a symmetric polynomial with coefficients in $\mathbb{Z}$, i.e. $s_\lambda^{\det}(x) \in \mathbb{Z}[x]^{\mathfrak{S}_n}$ (see Remark 2.2). It is not obvious, however, why they should have coefficients in $\mathbb{N} = \mathbb{Z}_{\ge 0}$.

Theorem 3.1 *For any $\lambda \in \mathcal{P}_n$, we have $s_\lambda^{\mathrm{comb}}(x) = s_\lambda^{\det}(x)$.*

Namely, the two definitions of the Schur functions give the same polynomials, which we denote by $s_\lambda(x)$. An immediate consequence of this theorem is that the Schur functions are symmetric polynomials with coefficients in $\mathbb{N} = \mathbb{Z}_{\ge 0}$, i.e. $s_\lambda(x) \in \mathbb{N}[x]^{\mathfrak{S}_n}$. The equivalence of the two definitions will be established later in Sect. 3.5 on the basis of Cauchy's formula.

3.1.2 Combinatorial Definition

By a *semi-standard tableau* T of shape λ in letters $\{1, \dots, n\}$, we mean a mapping $T : D(\lambda) \to \{1, \dots, n\}$ such that the numbers $T(s)$ ($s \in D(\lambda)$) are weakly increasing along the rows and strictly increasing along the columns.[1] For example,

[1] T is called a *column strict tableau* in the terminology of Macdonald [20].

$$T = \begin{array}{|c|c|c|c|c|}\hline 1 & 1 & 2 & 3 & 4 \\ \hline 2 & 3 & 3 \\ \cline{1-3} 4 \\ \cline{1-1}\end{array} \qquad \begin{array}{c} a \le b \\ \wedge \\ c \end{array} \tag{3.5}$$

Namely, T should satisfy

$$\begin{aligned} &T(i,j) \le T(i,j+1) \quad (1 \le i \le \ell(\lambda), 1 \le j < \lambda_i), \\ &T(i,j) < T(i+1,j) \quad (1 \le j \le \lambda_1,\ 1 \le i < \lambda'_j). \end{aligned} \tag{3.6}$$

We denote by $\mathrm{SSTab}_n(\lambda)$ the set of all semi-standard tableaux of shape λ in letters $\{1,\ldots,n\}$. For each semi-standard tableau T, we denote by $\mathrm{wt}(T)$ the composition (multi-index)

$$\mu = (\mu_1,\ldots,\mu_n) \in \mathbb{N}^n, \quad \mu_i = \#\{s \in D(\lambda) \mid T(s) = i\} \quad (i = 1,\ldots,n) \tag{3.7}$$

obtained by counting the number of i's in the tableau T for each i; $\mathrm{wt}(T)$ is called the *weight* of T. In the example of T in (3.5), we have

$$\mathrm{wt}(T) = (2,2,3,2), \quad x^{\mathrm{wt}(T)} = x_1^2 x_2^2 x_3^3 x_4^2. \tag{3.8}$$

$s_\lambda^{\mathrm{comb}}(x)$ attached to columns and rows

(1) Single column $\lambda = (1^r)$: $(r = 0,1,2,\ldots)$

$$s_{(1^r)}^{\mathrm{comb}}(x) = \sum_{1 \le i_1 < \cdots < i_r \le n} x_{i_1} \cdots x_{i_r} = e_r(x). \tag{3.9}$$

(2) Single row $\lambda = (l)$: $(l = 0,1,2,\ldots)$

$$s_{(l)}^{\mathrm{comb}}(x) = \sum_{1 \le j_1 \le \cdots \le j_l \le n} x_{j_1} \cdots x_{j_r} = h_l(x). \tag{3.10}$$

Example of $s_\lambda^{\mathrm{comb}}(x)$: $n = 3$, $\lambda = (2,1,0)$

When $n = 3$ and $\lambda = (2,1,0)$, there are 8 semi-standard tableaux of shape λ.

$$\begin{array}{cccccccc} \begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1&1\\\hline 3\\\cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1&2\\\hline 2\\\cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1&3\\\hline 3\\\cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 2&2\\\hline 3\\\cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 2&3\\\hline 3\\\cline{1-1}\end{array} \\ x_1^2x_2 & x_1^2x_3 & x_1x_2^2 & x_1x_2x_3 & x_1x_2x_3 & x_1x_3^2 & x_2^2x_3 & x_2x_3^2 \end{array} \tag{3.11}$$

Hence we have

$$
\begin{aligned}
s^{\mathrm{comb}}_{(21)}(x) &= x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2 + 2x_1x_2x_3 \\
&= m_{(21)}(x) + 2m_{(1^3)}(x).
\end{aligned}
\tag{3.12}
$$

Note that the definition of $s^{\mathrm{comb}}_\lambda(x)$ strongly depends on the ordering of $1, 2, \ldots, n$. By definition we have $s^{\mathrm{comb}}_\lambda(x) \in \mathbb{N}[x]$, but why are they symmetric?

For each $\mu \in \mathbb{N}^n$ with $|\mu| = |\lambda|$, we set

$$
\mathrm{SSTab}_n(\lambda)_\mu = \{T \in \mathrm{SSTab}_n(\lambda) \mid \mathrm{wt}(T) = \mu\}. \tag{3.13}
$$

The number

$$
K_{\lambda,\mu} = \#\mathrm{SSTab}_n(\lambda)_\mu \in \mathbb{N} \tag{3.14}
$$

of semi-standard tableaux of shape λ with weight μ is called the *Kostka number*. Then we have

$$
s^{\mathrm{comb}}_\lambda(x) = \sum_{\mu\in\mathbb{N}^n} \left(\#\mathrm{SSTab}_n(\lambda)_\mu\right) x^\mu = \sum_{\mu\in\mathbb{N}^n} K_{\lambda,\mu}\, x^\mu. \tag{3.15}
$$

In fact we have

$$
s^{\mathrm{comb}}_\lambda(x) = x^\lambda + \sum_{\mu<\lambda} K_{\lambda,\mu}\, x^\mu, \tag{3.16}
$$

namely, $s^{\mathrm{comb}}_\lambda(x)$ has the leading term x^λ with respect to the dominance order.

Exercise 3.1 Let $\lambda \in \mathcal{P}_n$. Prove the following:
(1) If $T \in \mathrm{SSTab}_n(\lambda)$ and $\mathrm{wt}(T) = \mu$, then $\mu \le \lambda$.
(2) $K_{\lambda,\lambda} = 1$, and $K_{\lambda,\mu} > 0$ if and only if $\mu \le \lambda$.

Remark 3.1 As we mentioned already, each $s^{\mathrm{comb}}_\lambda(x)$ is in fact a symmetric polynomial. This statement is equivalent to $K_{\lambda,\mu} = K_{\lambda,\sigma.\mu}$ ($\mu \in \mathbb{N}^n$) for any permutation $\sigma \in \mathfrak{S}_n$. We remark that, for each adjacent transposition $s_i = (i, i+1)$ ($i = 1, \ldots, n-1$), there is a bijection

$$
\mathrm{SSTab}_n(\lambda)_\mu \xrightarrow{\sim} \mathrm{SSTab}_n(\lambda)_{s_i.\mu} \tag{3.17}
$$

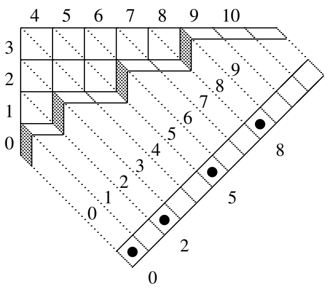
called the *Bender–Knuth involution*. It implies that $K_{\lambda,\mu} = K_{\lambda,s_i.\mu}$ ($\mu \in \mathbb{N}^n$) for $i = 1, \ldots, n-1$, and hence $K_{\lambda,\mu} = K_{\lambda,\sigma.\mu}$ ($\mu \in \mathbb{N}^n$) for any $\sigma \in \mathfrak{S}_n$. For a combinatorial proof of $\mathfrak{S}_n$-invariance of this sort, see Sagan's textbook [31, Proposition 4.4.2] for example.

3.1.3 Determinantal Definition

For each $\lambda \in \mathcal{P}_n$, we defined $s_\lambda^{\mathrm{det}}(x)$ as the ratio of two determinants in Definition 3.2. We denote by $\delta = (n-1, n-2, \ldots, 0)$ the *staircase* partition of $n-1$ parts so that $\delta_i = n - i$ $(i = 1, \ldots, n)$. Then the definition of $s_\lambda^{\mathrm{det}}(x)$ can be rewritten as

$$s_\lambda^{\mathrm{det}}(x) = \frac{\det\left(x_i^{(\lambda+\delta)_j}\right)_{i,j=1}^n}{\det\left(x_i^{\delta_j}\right)_{i,j=1}^n}. \tag{3.18}$$

We give here a remark on the strict partition $l = \lambda + \delta$ appearing in the numerator.[2] The sequence $l = (l_1, \ldots, l_n)$, $l_j = \lambda_j + n - j$ $(j = 1, \ldots, n)$, can be read off from the boundary of the Young diagram as shown below.



$$\begin{aligned} n &= 4 \\ \lambda &= (5, 3, 1, 0) \\ \delta &= (3, 2, 1, 0) \\ \lambda + \delta &= (8, 5, 2, 0) \end{aligned} \tag{3.19}$$

The subset $M = \{l_1, \ldots, l_n\} \subseteq \mathbb{N}$ is often called the *Maya diagram* attached to λ.

Example of $s_\lambda^{\mathrm{det}}(x)$: $n = 3$, $\lambda = (2, 1, 0)$

Since $\lambda + \delta = (4, 2, 0)$, we have

$$s_{(21)}^{\mathrm{det}}(x) = \det\begin{bmatrix} x_1^4 & x_1^2 & 1 \\ x_2^4 & x_2^2 & 1 \\ x_3^4 & x_3^2 & 1 \end{bmatrix} \Big/ \det\begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{bmatrix} = \frac{\Delta(x_1^2, x_2^2, x_3^2)}{\Delta(x_1, x_2, x_3)}. \tag{3.20}$$

Hence

$$\begin{aligned} s_{(21)}^{\mathrm{det}}(x) &= \frac{(x_1^2 - x_2^2)(x_1^2 - x_3^2)(x_2^2 - x_3^2)}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)} = (x_1 + x_2)(x_1 + x_3)(x_2 + x_3) \\ &= x_1^2 x_2 + x_1^2 x_3 + x_1 x_2^2 + x_1 x_3^2 + x_2^2 x_3 + x_2 x_3^2 + 2 x_1 x_2 x_3 \\ &= m_{(21)}(x) + 2 m_{(1^3)}(x). \end{aligned} \tag{3.21}$$

[2] A partition $\lambda = (\lambda_1, \lambda_2, \ldots)$ with $\ell(\lambda) = l$ is called *strict* if $\lambda_1 > \lambda_2 > \cdots > \lambda_l > 0$.

Exercise 3.2 Show that $s^{\det}_{(1^k)}(x) = e_k(x), \quad s^{\det}_{(k)}(x) = h_k(x) \quad (k = 0, 1, 2, \ldots)$.

A possible approach would be to use the following identities:

$$\prod_{j=1}^{n}(u - x_j) \cdot \Delta(x_1, \ldots, x_n) = \Delta(u, x_1, \ldots, x_n),$$
$$\frac{\Delta(x_1, \ldots, x_n)}{\prod_{i=1}^{n}(1 - x_i u)} = \det\left(\frac{x_i^{n-j}}{1 - x_i u}\right)_{i,j=1}^{n}. \tag{3.22}$$

Exercise 3.3 Prove that both $s^{\mathrm{comb}}_\lambda(x)$ and $s^{\det}_\lambda(x)$ carry the following properties.
(1) For any $\lambda \in \mathcal{P}_n$ and $k \in \mathbb{N}$, $s_{\lambda+(k^n)}(x) = (x_1 \ldots x_n)^k s_\lambda(x)$, where $(k^n) = (k, \ldots, k)$ denotes the $n \times k$ rectangle.
(2) Let $\lambda \in \mathcal{P}_n$ and $m < n$. Then we have

$$s_\lambda(x_1, \ldots, x_m, 0, \ldots, 0) = \begin{cases} s_\lambda(x_1, \ldots, x_m) & (\ell(\lambda) \leq m), \\ 0 & (\ell(\lambda) > m). \end{cases} \tag{3.23}$$

3.2 Principal Specialization and Self-duality

Before giving a proof of Theorem 3.1, we explain some consequences of the equivalence of the two definitions of Schur functions. From this section on, we set $s_\lambda(x) = s^{\det}_\lambda(x)$.

3.2.1 Principal Specialization: Evaluation at $x = t^\delta$

According to the combinatorial definition, the Schur function $s^{\mathrm{comb}}_\lambda(x)$ counts the semi-standard tableaux T of shape λ with weights $x^{\mathrm{wt}(T)}$. In particular, we have

$$s_\lambda(1, \ldots, 1) = s^{\mathrm{comb}}_\lambda(1, \ldots, 1) = \sum_{T \in \mathrm{SSTab}_n(\lambda)} 1 = \#\mathrm{SSTab}_n(\lambda). \tag{3.24}$$

In terms of the determinantal definition, the evaluation of $s_\lambda(x)$ at $x = (1, \ldots, 1)$ is a subtle question since the denominator $\Delta(x)$ vanishes at this point. In order to avoid this singularity, we first evaluate $s_\lambda(x)$ at $t^\delta = (t^{n-1}, t^{n-2}, \ldots, 1)$ and then take the limit as $t \to 1$.

Proposition 3.1 (Principal specialization) *For each $\lambda \in \mathcal{P}_n$, we have*

$$s_\lambda(t^\delta) = \frac{\Delta(t^{\lambda+\delta})}{\Delta(t^\delta)} = t^{n(\lambda)} \prod_{1 \leq i < j \leq n} \frac{1 - t^{\lambda_i - \lambda_j + j - i}}{1 - t^{j-i}}, \tag{3.25}$$

where $n(\lambda) = \sum_{i=1}^{n}(i - 1)\lambda_i$.

Proof In fact, we have

$$
\begin{aligned}
s_\lambda(t^\delta) &= \frac{\det\left(t^{\delta_i(\lambda+\delta)_j}\right)_{i,j=1}^n}{\Delta(t^\delta)} = \frac{\Delta(t^{\lambda+\delta})}{\Delta(t^\delta)} \\
&= \prod_{1\le i<j\le n} \frac{t^{\lambda_i+n-i} - t^{\lambda_j+n-j}}{t^{n-i} - t^{n-j}} = \prod_{1\le i<j\le n} t^{\lambda_j} \frac{1 - t^{\lambda_i-\lambda_j+j-i}}{1 - t^{j-i}}.
\end{aligned}
\tag{3.26}
$$

□

We are now allowed to take the limit as $t \to 1$ in (3.25), to obtain an explicit formula

$$
\#\mathrm{SSTab}_n(\lambda) = s_\lambda(1, \ldots, 1) = \frac{\Delta(\lambda+\delta)}{\Delta(\delta)} = \prod_{1\le i<j\le n} \frac{\lambda_i - \lambda_j + j - i}{j - i} \tag{3.27}
$$

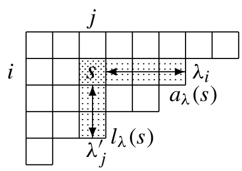
for the number of semi-standard tableaux of shape λ.

3.2.2 *Hook Length Formula*

Formulas (3.25) and (3.27) can be rewritten into a combinatorial expression of the Young diagram. For each square $s = (i, j) \in D(\lambda)$, we define the *content* $c_\lambda(s)$ and the *hook length* $h_\lambda(s)$ by

$$
c_\lambda(s) = j - i, \quad h_\lambda(s) = \lambda_i + \lambda'_j - i - j + 1. \tag{3.28}
$$

Note that, in terms of the *arm length* $a_\lambda(s) = \lambda_i - j$ and the *leg length* $l_\lambda(s) = \lambda'_j - i$, the hook length is expressed as $h_\lambda(s) = a_\lambda(s) + l_\lambda(s) + 1$.



$$
\begin{aligned}
a_\lambda(s) &= \lambda_i - j \\
l_\lambda(s) &= \lambda'_j - i \\
h_\lambda(s) &= a_\lambda(s) + l_\lambda(s) + 1
\end{aligned}
\tag{3.29}
$$

In terms of the Maya diagram $M = \{l_1, \ldots, l_n\}$, a square $s \in \lambda$ is in one-to-one correspondence with a pair (k, l) of nonnegative integers such that $k < l$, $k \notin M$, $l \in M$; the hook length is then interpreted as $h_\lambda(s) = l - k$.

Proposition 3.2 (Hook length formula) *For each $\lambda \in \mathcal{P}_n$, we have*

$$
s_\lambda(t^\delta) = t^{n(\lambda)} \prod_{s\in\lambda} \frac{1 - t^{n+c_\lambda(s)}}{1 - t^{h_\lambda(s)}}, \tag{3.30}
$$

and

$$s_\lambda(1,\ldots,1) = \#\mathrm{SSTab}_n(\lambda) = \prod_{s\in\lambda}\frac{n+c_\lambda(s)}{h_\lambda(s)}. \tag{3.31}$$

Proof We show

$$\prod_{1\le i<j\le n}\frac{\lambda_i-\lambda_j+j-i}{j-i} = \prod_{s\in\lambda}\frac{n+c_\lambda(s)}{h_\lambda(s)}. \tag{3.32}$$

Setting $l_i = \lambda_i + n - i$ $(i = 1,\ldots,n)$, consider the Maya digram $M = \{l_1,\ldots,l_n\}$ attached to λ. In terms of M, we see

$$\prod_{1\le i<j\le n}\frac{\lambda_i-\lambda_j+j-i}{j-i} = \prod_{1\le i<j\le n}\frac{l_i-l_j}{j-i} = \frac{\prod_{0\le k<l;k,l\in M}(l-k)}{\prod_{0\le k<l<n}(l-k)} \tag{3.33}$$

and

$$\prod_{s\in\lambda}h_\lambda(s) = \prod_{\substack{0\le k<l\\ k\notin M,l\in M}}(l-k). \tag{3.34}$$

Since

$$\prod_{\substack{0\le k<l\\ l\in M}}(l-k) = \prod_{\substack{0\le k<l\\ k,l\in M}}(l-k)\prod_{\substack{0\le k<l\\ k\notin M,l\in M}}(l-k), \tag{3.35}$$

we have

$$\prod_{1\le i<j\le n}\frac{\lambda_i-\lambda_j+j-i}{j-i}\cdot\prod_{s\in\lambda}h_\lambda(s) \tag{3.36}$$

$$\begin{aligned}
&= \frac{\prod_{0\le k<l;k,l\in M}(l-k)\prod_{0\le k<l;\ k\notin M,l\in M}(l-k)}{\prod_{0\le k<l<n}(l-k)}\\
&= \frac{\prod_{0\le k<l;l\in M}(l-k)}{\prod_{0\le k<l<n}(l-k)} = \frac{\prod_{i=1}^n(\lambda_i+n-i)!}{\prod_{i=1}^n(n-i)!}\\
&= \prod_{i=1}^n(n-i+1)_{\lambda_i} = \prod_{s\in\lambda}(n+c_\lambda(s)),
\end{aligned} \tag{3.37}$$

where we have used the notation of shifted factorials $(a)_k = a(a+1)\cdots(a+k-1)$ $(k = 0, 1, \ldots)$. The same proof applies to the formula for $s_\lambda(t^\delta)$ as well. □

Hook length formula

(1) $n = 3, \lambda = (2, 1, 0)$.

$$\prod_{s\in\lambda} \frac{n + c_\lambda(s)}{h_\lambda(s)} = \frac{\begin{array}{|c|c|}\hline 3 & 4 \\ \hline 2 \\ \cline{1-1} \end{array}}{\begin{array}{|c|c|}\hline 3 & 1 \\ \hline 1 \\ \cline{1-1} \end{array}} = 2 \cdot 4 = 8. \tag{3.38}$$

(2) $n = 4, \lambda = (5, 3, 1, 0)$.

$$\prod_{s\in\lambda} \frac{n + c_\lambda(s)}{h_\lambda(s)} = \frac{\begin{array}{|c|c|c|c|c|}\hline 4 & 5 & 6 & 7 & 8 \\ \hline 3 & 4 & 5 \\ \cline{1-3} 2 \\ \cline{1-1} \end{array}}{\begin{array}{|c|c|c|c|c|}\hline 7 & 5 & 4 & 2 & 1 \\ \hline 4 & 2 & 1 \\ \cline{1-3} 1 \\ \cline{1-1} \end{array}} = 360. \tag{3.39}$$

Exercise 3.4 Confirm that the hook length formula implies the following:

(1) Single column $\lambda = (1^r)$: $s_{(1^r)}(1, \ldots, 1) = \binom{n}{r}$ $(r \geq 0)$.

(2) Single row $\lambda = (l)$: $s_{(l)}(1, \ldots, 1) = \binom{n+l-1}{l}$ $(l \geq 0)$.

3.2.3 Self-duality

The values of $s_\lambda(x)$ at the discrete set $x = t^{\mu+\delta}$ $(\mu \in \mathcal{P}_n)$ have a remarkable duality property (evaluation symmetry).

Proposition 3.3 (Self-duality) *For any pair of partitions* $\lambda, \mu \in \mathcal{P}_n$, *we have*

$$\frac{s_\lambda(t^{\mu+\delta})}{s_\lambda(t^\delta)} = \frac{s_\mu(t^{\lambda+\delta})}{s_\mu(t^\delta)}. \tag{3.40}$$

Proof Since $s_\lambda(t^\delta) = \Delta(t^{\lambda+\delta})/\Delta(t^\delta)$, we have

$$\frac{s_\lambda(t^{\mu+\delta})}{s_\lambda(t^\delta)} = \frac{\Delta(t^\delta)\det\left(t^{(\mu+\delta)_i(\lambda+\delta)_j}\right)_{i,j=1}^n}{\Delta(t^{\lambda+\delta})\Delta(t^{\mu+\delta})}. \tag{3.41}$$

This formula is symmetric with respect to exchanging λ and μ. □

Regarding $x = t^\delta$ as a base point, we set

$$\widetilde{s}_\lambda(x) = \frac{s_\lambda(x)}{s_\lambda(t^\delta)} \tag{3.42}$$

so that $\widetilde{s}_\lambda(t^\delta) = 1$. Then Proposition 3.3 implies that $\widetilde{s}_\lambda(t^{\mu+\delta}) = \widetilde{s}_\mu(t^{\lambda+\delta})$ for any pair of partitions $\lambda, \mu \in \mathcal{P}_n$. Namely, regarded as a function of $(\lambda, \mu) \in \mathcal{P}_n \times \mathcal{P}_n$, $\widetilde{s}_\lambda(t^{\mu+\delta})$ is invariant under the exchange of the arguments λ and μ.

3.3 Cauchy Formula

In this section, we give a proof of the *Cauchy formula* for Schur functions; it will be used in Sect. 3.5 to establish the equivalence of two definitions of the Schur functions.

3.3.1 Cauchy Determinant

Lemma 3.1 (Cauchy) *For two sets of variables $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, we have*

$$\det\left(\frac{1}{x_i + y_j}\right)_{i,j=1}^{n} = \frac{\Delta(x)\Delta(y)}{\prod_{i,j=1}^{n}(x_i + y_j)}, \tag{3.43}$$

$$\det\left(\frac{1}{1 - x_i y_j}\right)_{i,j=1}^{n} = \frac{\Delta(x)\Delta(y)}{\prod_{i,j=1}^{n}(1 - x_i y_j)}. \tag{3.44}$$

□

The two formulas of Lemma 3.1 are equivalent to each other; the second formula is obtained from the first by change of variables $x_i \to -x_i^{-1}$ $(i = 1, \dots, n)$ and *vice versa*.

Exercise 3.5 Prove Cauchy's lemma (3.43) by means of the property of alternating polynomials.

Exercise 3.6 (1) For any $n \times n$ matrix $\left(a_{ij}\right)_{i,j=1}^{n}$ with $a_{nn} \neq 0$, its determinant is expressed as follows by a determinant of 2×2 minors (a variant of the *Dodgson condensation*):

$$\det\left(a_{ij}\right)_{i,j=1}^{n} = a_{nn}^{-n+2} \det\left(a_{ij}a_{nn} - a_{in}a_{nj}\right)_{i,j=1}^{n}. \tag{3.45}$$

(2) Use (3.45) to give an inductive proof of Cauchy's lemma.

Remark 3.2 Lemma 3.1 can be extended to the following family of determinant formulas involving an extra parameter u:

$$\det\left(\frac{u+x_i+y_j}{u(x_i+y_j)}\right)_{i,j=1}^{n}=\frac{u+\sum_{i=1}^{n}x_i+\sum_{j=1}^{n}y_j}{u}\frac{\Delta(x)\Delta(y)}{\prod_{i,j=1}^{n}(x_i+y_j)},\tag{3.46}$$

$$\det\left(\frac{1-ux_iy_j}{(1-u)(1-x_iy_j)}\right)_{i,j=1}^{n}=\frac{1-ux_1\cdots x_ny_1\cdots y_n}{1-u}\frac{\Delta(x)\Delta(y)}{\prod_{i,j=1}^{n}(1-x_iy_j)},\tag{3.47}$$

$$\det\left(\frac{\sigma(u+x_i+y_j)}{\sigma(u)\sigma(x_i+y_j)}\right)_{i,j=1}^{n}$$
$$=\frac{\sigma(u+\sum_{i=1}^{n}x_i+\sum_{j=1}^{n}y_j)}{\sigma(u)}\frac{\prod_{1\le i<j\le n}\sigma(x_i-x_j)\sigma(y_i-y_j)}{\prod_{i,j=1}^{n}\sigma(x_i+y_j)},\tag{3.48}$$

where $\sigma(z)=\sigma(z|\Omega)$ stands for the *Weierstrass sigma function* attached to a period lattice $\Omega=\mathbb{Z}\omega_1\oplus\mathbb{Z}\omega_2\subset\mathbb{C}$ of rank two ($\mathrm{Im}(\omega_2/\omega_1)>0$), defined by

$$\sigma(z|\Omega)=z\prod_{\omega\in\Omega,\ \omega\neq0}\left(1-\frac{z}{\omega}\right)e^{z/\omega+z^2/2\omega^2}\qquad(z\in\mathbb{C}).\tag{3.49}$$

These three variations (rational, trigonometric and elliptic) play crucial roles in various situations of integrable systems. Here, formula (3.47) is called trigonometric in the sense of additive variables θ_i such that $x_i=e^{\sqrt{-1}\theta_i}$.

3.3.2 Cauchy Formula for Schur Functions

In what follows, we use the notation of Schur functions $s_\lambda(x)$ for $s_\lambda^{\det}(x)$.

Theorem 3.2 (Cauchy formula) *For two sets of variables $x=(x_1,\ldots,x_n)$ and $y=(y_1,\ldots,y_n)$, the following identity holds in the ring $\mathbb{C}[\![x,y]\!]$ of formal power series in x and y:*

$$\prod_{i,j=1}^{n}\frac{1}{1-x_iy_j}=\sum_{\lambda\in\mathcal{P}_n}s_\lambda(x)s_\lambda(y).\tag{3.50}$$

Proof We make use of the multiplicative version (3.44) of Cauchy's lemma.

$$\frac{\Delta(x)\Delta(y)}{\prod_{i,j=1}^{n}(1-x_iy_j)}=\det\left(\frac{1}{1-x_iy_j}\right)_{i,j=1}^{\infty}$$
$$=\sum_{\sigma\in\mathfrak{S}_n}\mathrm{sgn}(\sigma)\frac{1}{(1-x_{\sigma(1)}y_1)\cdots(1-x_{\sigma(n)}y_n)}$$
$$=\sum_{\sigma\in\mathfrak{S}_n}\mathrm{sgn}(\sigma)\sum_{k_1,\ldots,k_n\ge0}(x_{\sigma(1)}y_1)^{k_1}\cdots(x_{\sigma(n)}y_n)^{k_n}$$

$$
\begin{aligned}
&= \sum_{k_1,\ldots,k_n \geq 0} \left(\sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) x_{\sigma(1)}^{k_1} \cdots x_{\sigma(n)}^{k_n} \right) y_1^{k_1} \cdots y_n^{k_n} \\
&= \sum_{k_1,\ldots,k_n \geq 0} \Delta_{k_1,\ldots,k_n}(x)\, y_1^{k_1} \cdots y_n^{k_n}, \qquad (3.51)
\end{aligned}
$$

where we have used the notation of determinants

$$
\Delta_{k_1,\ldots,k_n}(x) = \det\left(x_i^{k_j}\right)_{i,j=1}^{n} \qquad (3.52)
$$

of Vandermonde type (alternating polynomials of monomial type $(k_1, \ldots, k_n)$). Note that $\Delta_{n-1,n-2,\ldots,0}(x) = \Delta(x)$. Since $\Delta_{k_1,\ldots,k_n}(x)$ is alternating in $(k_1, \ldots, k_n)$, we have only to consider the cases where $k_1, \ldots, k_n$ are mutually distinct. In such a case, there exists a unique sequence $(l_1, \ldots, l_n) \in \mathbb{N}^n$ and a permutation $\sigma \in \mathfrak{S}_n$ such that

$$
l_1 > \cdots > l_n \geq 0, \quad (k_1, \ldots, k_n) = (l_{\sigma(1)}, \ldots, l_{\sigma(n)}). \qquad (3.53)
$$

Then we have

$$
\Delta_{k_1,\ldots,k_n}(x) = \Delta_{l_{\sigma(1)},\ldots,l_{\sigma(n)}}(x) = \operatorname{sgn}(\sigma)\Delta_{l_1,\ldots,l_n}(x). \qquad (3.54)
$$

Hence,

$$
\begin{aligned}
\frac{\Delta(x)\Delta(y)}{\prod_{i,j=1}^{n}(1 - x_i y_j)} &= \sum_{l_1 > \ldots > l_n \geq 0} \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \Delta_{l_1,\ldots,l_n}(x)\, y_1^{l_{\sigma(1)}} \cdots y_n^{l_{\sigma(n)}} \\
&= \sum_{l_1 > \ldots > l_n \geq 0} \Delta_{l_1,\ldots,l_n}(x) \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) y_1^{l_{\sigma(1)}} \cdots y_n^{l_{\sigma(n)}} \\
&= \sum_{l_1 > \ldots > l_n \geq 0} \Delta_{l_1,\ldots,l_n}(x) \Delta_{l_1,\ldots,l_n}(y). \qquad (3.55)
\end{aligned}
$$

Each $l = (l_1, \ldots, l_n) \in \mathbb{N}^n$ with $l_1 > \cdots > l_n \geq 0$ is uniquely expressed in the form $l = \lambda + \delta$ with $\lambda \in \mathcal{P}_n$, and we have $\Delta_l(x) = \Delta_{\lambda+\delta}(x) = \Delta(x)s_\lambda(x)$ by the definition of $s_\lambda(x) = s_\lambda^{\det}(x)$. Hence we obtain

$$
\begin{aligned}
\frac{\Delta(x)\Delta(y)}{\prod_{i,j=1}^{n}(1 - x_i y_j)} &= \sum_{\lambda \in \mathcal{P}_n} \Delta_{\lambda+\delta}(x)\Delta_{\lambda+\delta}(y) \\
&= \Delta(x)\Delta(y) \sum_{\lambda \in \mathcal{P}_n} s_\lambda(x) s_\lambda(y), \qquad (3.56)
\end{aligned}
$$

as desired. □

It is convenient to introduce the *signed* version

$$f_{k_1,\ldots,k_n}(x) = \frac{\Delta_{k_1,\ldots,k_n}(x)}{\Delta(x)} \quad (k_1, \ldots, k_n \in \mathbb{N}) \tag{3.57}$$

of $s_\lambda(x)$ with alternating indices $(k_1, \ldots, k_n)$. Note that, if $l = (l_1, \ldots, l_n) \in \mathbb{N}^n$ is *strict* in the sense $l_1 > \ldots > l_n \geq 0$, then we have $f_{l_1,\ldots,l_n}(x) = s_\lambda(x)$ for the partition $\lambda \in \mathcal{P}_n$ such that $l = \lambda + \delta$. In terms of these functions, Cauchy's formula is written as

$$\begin{aligned} \frac{\Delta(y)}{\prod_{i,j=1}^{n}(1 - x_i y_j)} &= \sum_{\lambda \in \mathcal{P}_n} s_\lambda(x) \Delta_{\lambda+\delta}(y) \\ &= \sum_{k_1,\ldots,k_n \geq 0} f_{k_1,\ldots,k_n}(x)\, y_1^{k_1} \cdots y_n^{k_n}. \end{aligned} \tag{3.58}$$

This formula will be used in Sect. 3.5 to establish equivalence of the two definitions of Schur functions.

We also remark that Cauchy's formula can be generalized to the case of two sets of variables with unequal dimensions: For two sets of variables $x = (x_1, \ldots, x_m)$ and $y = (y_1, \ldots, y_n)$,

$$\prod_{i=1}^{m} \prod_{j=1}^{n} \frac{1}{1 - x_i y_j} = \sum_{\ell(\lambda) \leq \min\{m,n\}} s_\lambda(x_1, \ldots, x_m)\, s_\lambda(y_1, \ldots, y_n). \tag{3.59}$$

This formula is obtained from the $m = n$ case by setting unnecessary variables to zero, thanks to the stability property of Exercise 3.3 (2).

3.4 Recurrence on the Number of Variables

We recall the combinatorial definition of Schur functions:

$$s_\lambda^{\mathrm{comb}}(x) = \sum_{T \in \mathrm{SSTab}_n(\lambda)} x^{\mathrm{wt}(T)}. \tag{3.60}$$

Given a semi-standard tableau $T \in \mathrm{SSTab}_n(\lambda)$ of shape λ in letters $\{1, \ldots, n\}$, let T' be the sub-tableau of T consisting of letters in $\{1, \ldots, n-1\}$. Then by the condition of a semi-standard tableau, the shape $\mu = (\mu_1, \mu_2, \ldots)$ of T' is a partition satisfying the *interlacing property*

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \lambda_3 \geq \cdots . \tag{3.61}$$

A pair (λ, μ) of partitions in $\mathcal{P}$ with $\mu \subseteq \lambda$ (i.e. $\mu_i \leq \lambda_i$ for all $i \geq 1$) is referred to as a *skew diagram* λ/μ; we also use the notation $\lambda\backslash\mu$ for the set-theoretic difference $D(\lambda)\backslash D(\mu)$ of diagrams.

We say that a skew diagram λ/μ is a *horizontal strip* ("h-strip" for short) if the pair (λ, μ) satisfies the interlacing property (3.61). In terms of the Young diagrams, this condition is equivalent to saying that the difference $\lambda\backslash\mu$ has at most one square in each column. In this terminology, $s_\lambda^{\mathrm{comb}}(x)$ can be expanded in the form

$$s_\lambda^{\mathrm{comb}}(x) = \sum_{\lambda/\mu:\ \text{h-strip}} \ \sum_{T' \in \mathrm{SSTab}_{n-1}(\mu)} (x')^{\mathrm{wt}(T')}\, x_n^{|\lambda|-|\mu|} \tag{3.62}$$

$$= \sum_{\lambda/\mu:\ \text{h-strip}} s_\mu^{\mathrm{comb}}(x')\, x_n^{|\lambda|-|\mu|}, \tag{3.63}$$

where $x' = (x_1, \ldots, x_{n-1})$. Namely,

$$s_\lambda^{\mathrm{comb}}(x_1, \ldots, x_n) = \sum_{\lambda/\mu:\ \text{h-strip}} s_\mu^{\mathrm{comb}}(x_1, \ldots, x_{n-1})\, x_n^{|\lambda/\mu|}, \tag{3.64}$$

where $|\lambda/\mu| = |\lambda| - |\mu|$. The combinatorial Schur functions $s_\lambda^{\mathrm{comb}}(x)$ are completely determined by this recurrence formula with respect to the number of variables.

In order to establish the equivalence of the two definitions of Schur functions, we prove that $s_\lambda(x) = s_\lambda^{\det}(x)$ satisfy the same recurrence formula. Since

$$s_\lambda(x) = (x_1 \cdots x_n)^{\lambda_n} s_{\lambda-(\lambda_n^n)}(x), \quad s_\lambda^{\mathrm{comb}}(x) = (x_1 \cdots x_n)^{\lambda_n} s_{\lambda-(\lambda_n^n)}^{\mathrm{comb}}(x), \tag{3.65}$$

we have only to consider the case where $\lambda_n = 0$.

Theorem 3.3 *The Schur functions $s_\lambda(x)$ satisfy the following recurrence formula with respect to the number of variables n: For any $\lambda \in \mathcal{P}_n$,*

$$s_\lambda(x_1, \ldots, x_n) = \sum_{\substack{\mu \subseteq \lambda \\ \lambda/\mu:\ \text{h-strip}}} s_\mu(x_1, \ldots, x_{n-1})\, x_n^{|\lambda/\mu|}, \tag{3.66}$$

where the sum is over all partitions $\mu \subseteq \lambda$ such that λ/μ is a horizontal strip.

Recurrence formulas of this kind are called *branching formulas* as well. We give a proof of this theorem in Sect. 3.5.

Applying this recurrence formula repeatedly, we obtain an alternative expression of the tableau representation of $s_\lambda(x)$:

$$s_\lambda(x) = \sum_{\substack{\emptyset=\lambda^{(0)} \subseteq \lambda^{(1)} \subseteq \cdots \subseteq \lambda^{(n)} = \lambda \\ |\lambda^{(i)}/\lambda^{(i-1)}|:\ \text{h-strip}}} \prod_{i=1}^{n} x_i^{|\lambda^{(i)}/\lambda^{(i-1)}|} \tag{3.67}$$

where the sum is taken over all weakly increasing sequences of partitions $\emptyset = \lambda^{(0)} \subseteq \lambda^{(1)} \subseteq \cdots \subseteq \lambda^{(n)} = \lambda$ connecting $\emptyset$ (empty diagram) and λ by n steps such that the successive skew diagrams $\lambda^{(i)}/\lambda^{(i-1)}$ are all horizontal strips. It is also convenient to display such a sequence of partitions $\lambda^{(1)}, \ldots, \lambda^{(n)}$ as a table

$$\begin{bmatrix} \lambda_1^{(n)} \ \lambda_2^{(n)} \ \lambda_3^{(n)} \ \ldots \ \lambda_n^{(n)} \\ \lambda_1^{(n-1)} \lambda_2^{(n-1)} \ldots \lambda_{n-1}^{(n-1)} \\ \cdots \\ \lambda_1^{(2)} \ \lambda_2^{(2)} \\ \lambda_1^{(1)} \end{bmatrix} \tag{3.68}$$

with interlacing property $\lambda_j^{(i)} \geq \lambda_j^{(i-1)} \geq \lambda_{j+1}^{(i)}$ for $1 \leq j < i \leq n$, called a *Gelfand–Tsetlin pattern*.

3.5 Equivalence of the Two Definitions

In this section, we give a proof of Theorem 3.3, thereby establishing the equivalence of two definitions of Schur functions.

The recurrence formula (3.66) for $s_\lambda(x)$ (with $\lambda_n = 0$) can be proved by means of Cauchy's formula (3.58) for $f_{l_1,\ldots,l_n}(x) = \Delta_{l_1,\ldots,l_n}(x)/\Delta(x)$:

$$\sum_{l_1,\ldots,l_n \geq 0} f_{l_1,\ldots,l_n}(x_1, \ldots, x_n)\, y_1^{l_1} \cdots y_n^{l_n} = \frac{\Delta(y_1, \ldots, y_n)}{\prod_{i,j=1}^n (1 - x_i y_j)}. \tag{3.69}$$

In this formula, we set $y_n = 0$ to obtain

$$\begin{aligned} &\sum_{l_1,\ldots,l_{n-1} \geq 0} f_{l_1,\ldots,l_{n-1},0}(x_1, \ldots, x_n)\, y_1^{l_1} \cdots y_{n-1}^{l_{n-1}} \\ &= \frac{\Delta(y_1, \ldots, y_{n-1})}{\prod_{i,j=1}^{n-1} (1 - x_i y_j)} \frac{y_1 \cdots y_{n-1}}{\prod_{j=1}^{n-1} (1 - x_n y_j)} \\ &= \left(\sum_{k_1,\ldots,k_{n-1} \geq 0} f_{k_1,\ldots,k_{n-1}}(x_1, \ldots, x_{n-1}) y_1^{k_1} \cdots y_{n-1}^{k_{n-1}} \right) \\ &\quad \cdot \left(\sum_{r_1,\ldots,r_{n-1} \geq 0} x_n^{\sum_j r_j} y_1^{r_1+1} \cdots y_{n-1}^{r_{n-1}+1} \right). \end{aligned} \tag{3.70}$$

We now look at the coefficient of $y_1^{l_1} \cdots y_{n-1}^{l_{n-1}}$ assuming that $l_1 > l_2 > \cdots > l_{n-1} \geq 0$:

$$f_{l_1,\ldots,l_{n-1},0}(x_1,\ldots,x_n) = \sum_{\substack{k_1,\ldots,k_{n-1}\geq 0 \\ 0\leq k_i<l_i}} f_{k_1,\ldots,k_{n-1}}(x_1,\ldots,x_{n-1})\, x_n^{\sum_j l_j - \sum_j k_j - (n-1)} \tag{3.71}$$

where the sum is taken over all $(k_1,\ldots,k_{n-1}) \in \mathbb{N}^{n-1}$ such that $0 \leq k_i < l_i$ ($i = 1,\ldots,n-1$), namely

$$(k_1,\ldots,k_{n-1}) \in [0,l_1) \times [0,l_2) \times \cdots \times [0,l_{n-1}), \tag{3.72}$$

where we have used the symbol $[a,b) = \{k \in \mathbb{Z} \mid a \leq k < b\}$ for an interval of integers. Notice that, in the expression

$$f_{l_1,\ldots,l_{n-1},0}(x_1,\ldots,x_n) = \sum_{\substack{(k_1,k_2,\ldots,k_{n-1}) \\ \in[0,l_1)\times[0,l_2)\times\cdots\times[0,l_{n-1}),}} f_{k_1,\ldots,k_{n-1}}(x_1,\ldots,x_{n-1})\, x_n^{\sum_j l_j - \sum_j k_j - (n-1)}, \tag{3.73}$$

the summand $f_{k_1,\ldots,k_{n-1}}(x_1,\ldots,x_{n-1})$ is alternating with respect to the permutation of $k_1,\ldots,k_{n-1}$. Thanks to this alternating property, the sum over the first two indices k_1, k_2 reduces as

$$\begin{aligned}&\sum_{(k_1,k_2)\in[0,l_1)\times[0,l_2)} f_{k_1,\ldots,k_{n-1}}(x_1,\ldots,x_{n-1})\, x_n^{\sum_j l_j - \sum_j k_j - (n-1)} \\ &= \sum_{(k_1,k_2)\in[l_2,l_1)\times[0,l_2)} f_{k_1,\ldots,k_{n-1}}(x_1,\ldots,x_{n-1})\, x_n^{\sum_j l_j - \sum_j k_j - (n-1)},\end{aligned} \tag{3.74}$$

since the sum of an alternating function over a symmetric region gives zero (Fig. 3.1).

Repeating this procedure with $(k_2, k_3) \in [0,l_2) \times [0,l_3)$ and so on, we finally obtain

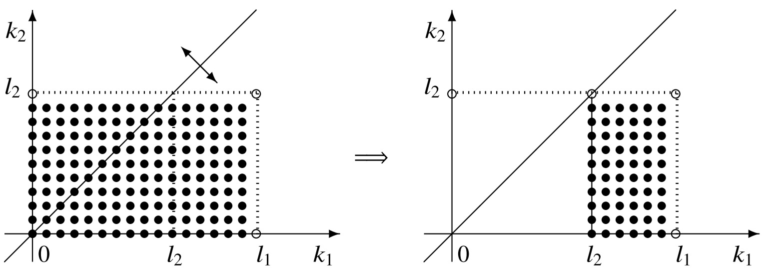


Fig. 3.1 Reducing the region of summation indices

$$
\begin{aligned}
&f_{l_1,\dots,l_{n-1},0}(x_1,\dots,x_n) \\
&= \sum_{\substack{(k_1,k_2,\dots,k_{n-1}) \\ \in [l_2,l_1)\times[l_3,l_2)\times\cdots\times[0,l_{n-1}),}} f_{k_1,\dots,k_{n-1}}(x_1,\dots,x_{n-1})\, x_n^{\sum_j l_j - \sum_j k_j - (n-1)}, \qquad (3.75)
\end{aligned}
$$

where the sum is taken over all $(k_1,\dots,k_{n-1})$ such that

$$l_1 > k_1 \geq l_2 > k_2 \geq l_3 > \cdots \geq l_{n-1} > k_{n-1} \geq 0. \qquad (3.76)$$

Then passing to the expressions by partitions $\lambda = (\lambda_1,\dots,\lambda_{n-1},0) \in \mathcal{P}_n$ and $\mu = (\mu_1,\dots,\mu_{n-1}) \in \mathcal{P}_{n-1}$ such that

$$l_i = \lambda_i + n - i, \quad k_i = \mu_i + n - i - 1 \qquad (i = 1,\dots,n-1), \qquad (3.77)$$

we obtain

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \lambda_{n-1} \geq \mu_{n-1} \geq 0, \qquad (3.78)$$

and hence

$$s_\lambda(x_1,\dots,x_n) = \sum_{\substack{\mu\subseteq\lambda \\ \lambda/\mu:\ \text{h-strip}}} s_\mu(x_1,\dots,x_{n-1})\, x_n^{|\lambda|-|\mu|}, \qquad (3.79)$$

as desired.

3.6 Dual Cauchy Formula

We propose two versions of the dual Cauchy formula for Schur functions.

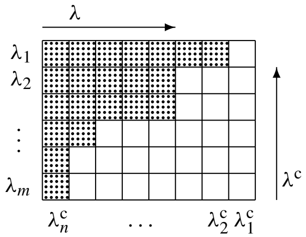
Theorem 3.4 (Dual Cauchy formulas) *For two sets of variables $x = (x_1,\dots,x_m)$ and $y = (y_1,\dots,y_n)$, we have*

$$\prod_{i=1}^{m}\prod_{j=1}^{n}(1 + x_i y_j) = \sum_{\lambda\subseteq(n^m)} s_\lambda(x) s_{\lambda'}(y), \qquad (3.80)$$

$$\prod_{i=1}^{m}\prod_{j=1}^{n}(x_i + y_j) = \sum_{\lambda\subseteq(n^m)} s_\lambda(x) s_{\lambda^{\mathrm{c}}}(y), \qquad (3.81)$$

where the sum is over all partitions λ contained in the $m \times n$ rectangle $(n^m) = (n,\dots,n)$; λ' denotes the conjugate partition of λ (see Sect. 2.4), and $\lambda^{\mathrm{c}} = (m - \lambda'_n, m - \lambda'_{n-1},\dots, m - \lambda'_1)$.

We call λ^{c} the *complementary partition* of λ in the $m \times n$ rectangle.



$$\begin{aligned}\lambda^c &= (\lambda_1^c, \dots, \lambda_n^c)\\ &= (m - \lambda'_n, \dots, m - \lambda'_1)\end{aligned} \tag{3.82}$$

For the proof of these formulas, we use a lemma on determinants. For an $N \times N$ matrix $Z = (z_{i,j})_{i,j=1}^N$, we denote by

$$\det Z^{i_1,\dots,i_r}_{j_1,\dots,j_r} = \det \left(z_{i_a,j_b}\right)_{a,b=1}^r \tag{3.83}$$

the $r \times r$ minor determinant of Z with row indices $i_1, \dots, i_r$ and column indices $j_1, \dots, j_r$. When $(i_1, \dots, i_r) = (1, \dots, r)$, we simply write $\det Z_{j_1,\dots,j_r}$ for $\det Z^{1,\dots,r}_{j_1,\dots,j_r}$. Also, for two subsets $I, J \subseteq \{1, \dots, N\}$ of indices with $|I| = |J| = r$, we use the notation $\det Z^I_J = \det Z^{i_1,\dots,i_r}_{j_1,\dots,j_r}$ and $\det Z_J = \det Z_{j_1,\dots,j_r}$ taking the increasing sequences $i_1 < \dots < i_r$ and $j_1 < \dots < j_r$ such that $I = \{i_1, \dots, i_r\}$ and $J = \{j_1, \dots, j_r\}$.

Lemma 3.2 *Setting $N = m + n$, let $X = (x_{i,j})_{1\le i\le m, 1\le j\le N}$ be an $m \times N$ matrix, and $Y = (y_{i,j})_{1\le i\le n, 1\le j\le N}$ an $n \times N$ matrix. Define the $N \times N$ matrix $Z = (z_{i,j})_{1\le i,j\le N}$ by*

$$z_{ij} = x_{i,j} \quad (1 \le i \le n), \qquad z_{m+i,j} = y_{i,j} \quad (1 \le i \le n) \tag{3.84}$$

for all $j = 1, \dots, N$. Then the determinant of Z is expressed as

$$\det Z = \sum_{\substack{K \sqcup L = \{1,\dots,N\}\\ |K|=m,\ |L|=n}} \epsilon(K; L) \det X_K \det Y_L \tag{3.85}$$

in terms of minor determinants of X and Y, where the sum is over all pairs of subsets $K, L \subseteq \{1, \dots, N\}$ such that $|K| = m$, $|L| = n$ and $K \sqcup L = \{1, \dots, N\}$, and $\epsilon(K; L)$ denotes the sign defined by

$$\epsilon(K; L) = (-1)^{\ell(K;L)}, \quad \ell(K; L) = \#\{(k,l) \in K \times L \mid k > l\}. \tag{3.86}$$

For the proof of this lemma, we refer the reader to [25], for example. □

Proof **(of Theorem 3.4)** Taking the variables $x = (x_1, \dots, x_m)$ and $y = (y_1, \dots, y_n)$ as in Theorem 3.4, we apply this lemma to the matrices

$$X = \left(x_i^{j-1}\right)_{1\le i\le m; 1\le j\le N}, \quad Y = \left(y_i^{j-1}\right)_{1\le i\le n; 1\le j\le N}, \quad N = m + n. \tag{3.87}$$

Then we have

$$(-1)^{\binom{N}{2}} \det Z = \Delta(x, y) = \Delta(x)\Delta(y) \prod_{i=1}^{m} \prod_{j=1}^{n} (x_i - y_j). \tag{3.88}$$

On the other hand, by Lemma 3.2 we have

$$\det Z = \sum_{\substack{K \sqcup L = \{1,\dots,N\} \\ |K|=m, |L|=n}} \epsilon(K; L) \det X_K \det Y_L. \tag{3.89}$$

Hence we obtain

$$(-1)^{\binom{N}{2}} \prod_{i=1}^{m} \prod_{j=1}^{n} (x_i - y_j) = \sum_{\substack{K \sqcup L = \{1,\dots,N\} \\ |K|=m, |L|=n}} \frac{\det X_K}{\Delta(x)} \frac{\det Y_L}{\Delta(y)}. \tag{3.90}$$

Taking two sequences $k_1 > k_2 > \ldots > k_m \geq 0$ and $l_1 > l_2 > \ldots > l_n \geq 0$ such that $K = \{k_m + 1, \ldots, k_1 + 1\}$ and $L = \{l_n + 1, \ldots, l_1 + 1\}$. Then we have

$$\begin{aligned} \det X_K &= (-1)^{\binom{m}{2}} \det \left(x_i^{k_j}\right)_{i,j=1}^{m} = (-1)^{\binom{m}{2}} \Delta_{k_1,\dots,k_m}(x) \\ \det Y_L &= (-1)^{\binom{n}{2}} \det \left(y_i^{l_j}\right)_{i,j=1}^{n} = (-1)^{\binom{n}{2}} \Delta_{l_1,\dots,l_n}(x). \end{aligned} \tag{3.91}$$

For each pair (K, L), we take two partitions $\mu \in \mathcal{P}_m$ and $\nu \in \mathcal{P}_n$ such that $k_i = \mu_i + m - i$ $(i = 1, \ldots, m)$ and $l_i = \nu_i + n - i$ $(i = 1, \ldots, n)$. Then one can show that $\nu = (m - \mu'_n, \ldots, m - \mu'_1) = \mu^{\mathrm{c}}$ and $\epsilon(K; L) = (-1)^{|\mu|}$. Hence, we can rewrite (3.90) as

$$(-1)^{mn} \prod_{i=1}^{m} \prod_{j=1}^{n} (x_i - y_j) = \sum_{\mu \in (n^m)} (-1)^{|\mu|} s_\mu(x) s_{\mu^{\mathrm{c}}}(y). \tag{3.92}$$

Replacing y_j by $-y_j$, we obtain the dual Cauchy formula (3.81). Formula (3.80) is obtained from (3.81) by the relation

$$(y_1 \cdots y_n)^m s_{\lambda'}(y^{-1}) = s_{\lambda^{\mathrm{c}}}(y), \tag{3.93}$$

which can be verified directly from the determinantal definition of the Schur function. □

3.7 Jacobi–Trudi Formula

From the Cauchy and the dual Cauchy formulas, one can read off various properties of Schur functions. For example, one can derive a determinant formula, called the *Jacobi–Trudi formula*, which represents a general Schur function $s_\lambda(x)$ in terms of complete homogeneous symmetric functions $h_k(x)$ or elementary symmetric functions $e_k(x)$

Theorem 3.5 (Jacobi–Trudi formula) *Let $\lambda \in \mathcal{P}_n$ and $\ell(\lambda') \leq m$. Then we have*

$$(1) \qquad s_\lambda(x) = \det \left(h_{\lambda_i+j-i}(x)\right)_{i,j=1}^{n}. \tag{3.94}$$

$$(2) \qquad s_\lambda(x) = \det \left(e_{\lambda'_i+j-i}(x)\right)_{i,j=1}^{m}. \tag{3.95}$$

In these formulas, we understand $h_k(x) = 0$, $e_k(x)$ for $k < 0$. Explicitly,

$$s_\lambda = \det \begin{bmatrix} h_{\lambda_1} & h_{\lambda_1+1} & \dots & h_{\lambda_1+n-1} \\ h_{\lambda_2-1} & h_{\lambda_2} & \dots & h_{\lambda_2+n-2} \\ \vdots & & \ddots & \vdots \\ h_{\lambda_n-n+1} & h_{\lambda_n-n+2} & \dots & h_{\lambda_n} \end{bmatrix} = \begin{bmatrix} e_{\lambda'_1} & e_{\lambda'_1+1} & \dots & e_{\lambda'_1+n-1} \\ e_{\lambda'_2-1} & e_{\lambda'_2} & \dots & e_{\lambda'_2+n-2} \\ \vdots & & \ddots & \vdots \\ e_{\lambda'_n-n+1} & e_{\lambda'_n-n+2} & \dots & e_{\lambda'_n} \end{bmatrix}. \tag{3.96}$$

Note that the size of the determinant can be reduced as

$$s_\lambda(x) = \det \left(h_{\lambda_i+j-i}(x)\right)_{i,j=1}^{\ell(\lambda)}, \quad s_\lambda(x) = \det \left(e_{\lambda'_i+j-i}(x)\right)_{i,j=1}^{\ell(\lambda')}, \tag{3.97}$$

since the (i, j) entries of the matrix vanish for $i > \ell(\lambda)$ (or $i > \ell(\lambda')$) and $j < i$.

Proof (1) We rewrite the Cauchy formula (3.50) as

$$\Delta(x) \prod_{i,j=1}^{n} \frac{1}{1 - x_i y_j} = \sum_{\lambda \in \mathcal{P}_n} \Delta_{\lambda+\delta}(x) s_\lambda(y). \tag{3.98}$$

Then $s_\lambda(y)$ is the coefficient of $x^{\lambda+\delta}$ in the right-hand side. On the other hand,

$$\Delta(x) \prod_{i,j=1}^{n} \frac{1}{1 - x_i y_j} = \Delta(x) \prod_{i=1}^{n} (1 + x_i h_1(y) + x_i^2 h_2(y) + \cdots) \tag{3.99}$$

$$= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) x^{\sigma.\delta} \sum_{\mu \in \mathbb{N}^n} x^\mu h_\mu(y), \tag{3.100}$$

where $h_\mu(y) = h_{\mu_1}(y) \cdots h_{\mu_n}(y)$. Taking the coefficient of $x^{\lambda+\delta}$, we obtain

$$\begin{aligned} s_\lambda(y) &= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) h_{\lambda+\delta-\sigma.\delta}(y) \\ &= \det\left(h_{\lambda_i+\delta_i-\delta_j}\right)_{i,j=1}^n = \det\left(h_{\lambda_i+j-i}(y)\right)_{i,j=1}^n, \end{aligned} \tag{3.101}$$

which proves (3.94).

(2) We rewrite the dual Cauchy formula (3.80) as

$$\Delta(x) \prod_{i=1}^m \prod_{j=1}^n (1 + x_i y_j) = \sum_{\lambda \subseteq (n^m)} \Delta_{\lambda+\delta}(x)\, s_{\lambda'}(y). \tag{3.102}$$

Then $s_{\lambda'}(y)$ is the coefficient of $x^{\lambda+\delta}$ in the right-hand side. On the other hand,

$$\begin{aligned} \Delta(x) \prod_{i=1}^m \prod_{j=1}^n (1 + x_i y_j) &= \Delta(x) \prod_{i=1}^m (1 + x_i e_1(y) + \cdots + x_i^n e_n(y)) \\ &= \sum_{\sigma \in \mathfrak{S}_m} \operatorname{sgn}(\sigma) x^{\sigma.\delta} \sum_{\mu \in \mathbb{N}^m} x^\mu e_\mu(y), \end{aligned} \tag{3.103}$$

where $e_\mu(y) = e_{\mu_1}(y) \cdots e_{\mu_n}(y)$. Taking the coefficient of $x^{\lambda+\delta}$ in this formula, we obtain

$$\begin{aligned} s_{\lambda'}(y) &= \sum_{\sigma \in \mathfrak{S}_m} \operatorname{sgn}(\sigma) e_{\lambda+\delta-\delta.\sigma}(y) \\ &= \det\left(e_{\lambda_i+\delta_i-\delta_j}\right)_{i,j=1}^n = \det\left(e_{\lambda_i+j-i}\right)_{i,j=1}^n, \end{aligned} \tag{3.104}$$

as desired. □

3.8 q-Difference and Differential Equations

For each $i = 1, \ldots, n$, we define the *q-shift operator* T_{q,x_i} in x_i by

$$T_{q,x_i} \varphi(x_1, \ldots, x_i, \ldots, x_n) = \varphi(x_1, \ldots, q x_i, \ldots, x_n) \quad (i = 1, \ldots, n) \tag{3.105}$$

leaving x_j for $j \neq i$ unchanged. For $r = 0, 1, \ldots, n$, we define the q-difference operators $D_x^{(r)}$ by

$$
\begin{aligned}
D_x^{(r)} &= \sum_{\substack{I\subseteq\{1,\ldots,n\}\\ |I|=r}} \frac{T_{q,x}^I(\Delta(x))}{\Delta(x)} T_{q,x}^I \\
&= \sum_{\substack{I\subseteq\{1,\ldots,n\}\\ |I|=r}} q^{\binom{r}{2}} \prod_{i\in I;\, j\notin J} \frac{qx_i - x_j}{x_i - x_j} \prod_{i\in I} T_{q,x_i},
\end{aligned}
\tag{3.106}
$$

where $T_{q,x}^I = \prod_{i\in I} T_{q,x_i}$. As we will see below, the q-difference operators $D_x^{(r)}$ $(r = 1, \ldots, n)$ commutes with each other. We remark that these q-difference operators $D_x^{(r)}$ are the special case of *Macdonald–Ruijsenaars operators* with $q = t$ to be discussed from the next chapter on.

Theorem 3.6 *For each $\lambda \in \mathcal{P}_n$, the Schur function $s_\lambda(x)$ satisfies the system of q-difference equations*

$$
D_x^{(r)} s_\lambda(x) = e_r(q^{\lambda+\delta}) s_\lambda(x) \quad (r = 0, 1, \ldots, n), \tag{3.107}
$$

where the eigenvalues $e_r(q^{\lambda+\delta})$ are the elementary symmetric functions of q^{λ_i+n-i} $(i = 1, \ldots, n)$.

In fact, the q-shift operator T_{q,x_i} acts on monomials in $x = (x_1, \ldots, x_n)$ by

$$
T_{q,x_i}(x^\mu) = q^{\mu_i} x^\mu, \quad \mu = (\mu_1, \ldots, \mu_n) \in \mathbb{N}^n \tag{3.108}
$$

for $i = 1, \ldots, n$. Hence, for each polynomials $f(\xi) \in \mathbb{C}[\xi]$ in $\xi = (\xi_1, \ldots, \xi_n)$, the q-difference operator $f(T_{q,x}) = f(T_{q,x_1}, \ldots, T_{q,x_n})$ acts on monomials by

$$
f(T_{q,x}) x^\mu = f(q^\mu) x^\mu \quad (\mu \in \mathbb{N}^n). \tag{3.109}
$$

If $f(\xi)$ is $\mathfrak{S}_n$-invarinat, then $f(T_{q,x})$ acts on monomial symmetric functions $m_\lambda(x) = \sum_{\mu\in\mathfrak{S}_n.\lambda} x^\mu$ by

$$
f(T_{q,x}) m_\lambda(x) = f(q^\lambda) m_\lambda(x) \qquad (\lambda \in \mathcal{P}_n), \tag{3.110}
$$

since $f(q^\mu) = f(q^{\sigma.\lambda}) = f(q^\lambda)$ for $\mu = \sigma.\lambda$, $\sigma \in \mathfrak{S}_n$. Taking elementary symmetric functions $e_r(\xi)$ for $f(\xi)$, we obtain

$$
e_r(T_{q,x}) m_\lambda(x) = e_r(q^\lambda) m_\lambda(x) \quad (\lambda \in \mathcal{P}_n, r = 0, 1, \ldots, n). \tag{3.111}
$$

Similarly, the alternating polynomials $\Delta_{\lambda+\delta}(x) = \sum_{\mu\in\mathfrak{S}_n.\lambda} \mathrm{sgn}(\sigma) x^{\sigma.(\lambda+\delta)}$ $(\lambda \in \mathcal{P}_n)$ satisfy

$$
f(T_{q,x}) \Delta_{\lambda+\delta}(x) = f(q^{\lambda+\delta}) \Delta_{\lambda+\delta}(x) \qquad (\lambda \in \mathcal{P}_n) \tag{3.112}
$$

for all $f(\xi) \in \mathbb{C}[\xi]^{\mathfrak{S}_n}$. By conjugation, we introduce the $q-$difference operators

$$D_x^f = \Delta(x) f(T_{q,x}) \Delta(x)^{-1}. \tag{3.113}$$

Then, we see that the Schur functions $s_\lambda(x) = \Delta_{\lambda+\delta}(x)/\Delta(x)$ satisfy the q-difference equations

$$D_x^f s_\lambda(x) = f(q^{\lambda+\delta}) s_\lambda(x) \qquad (\lambda \in \mathcal{P}_n) \tag{3.114}$$

for all symmetric polynomials $f(\xi) = \mathbb{C}[\xi]^{\mathfrak{S}_n}$. The q-difference operators $D_x^{(r)}$ of (3.106) are the special cases of D_x^f, where $f = e_r$ $(r = 0, 1, \dots, n)$. We also remark that the q-difference operators D_x^f for all $f(\xi) \in \mathbb{C}[\xi]^{\mathfrak{S}_n}$ pairwise commute since they are conjugations of $f(T_{q,x})$ by $\Delta(x)$.

The differential operators $x_i \partial_{x_i} = x_i \partial/\partial x_i$ acts on monomials in $x = (x_1, \dots, x_n)$ by

$$x_i \partial_{x_i} x^\mu = \mu_i x^\mu, \quad \mu = (\mu_1, \dots, \mu_n) \in \mathbb{N}^n \tag{3.115}$$

for $i = 1, \dots, n$. Hence for any polynomial $f(\xi) \in \mathbb{C}[\xi]$ in $\xi = (\xi_1, \dots, \xi_n)$, we have

$$f(x\partial_x) x^\mu = f(\mu) x^\mu \quad (\mu \in \mathbb{N}^n). \tag{3.116}$$

Hence for all $f(\xi) \in \mathbb{C}[\xi]^{\mathfrak{S}_n}$, we have

$$f(x\partial_x) m_\lambda(x) = f(\lambda) m_\lambda(x), \quad f(x\partial_x) \Delta_{\lambda+\delta}(x) = f(\lambda+\delta) \Delta_{\lambda+\delta}(x). \tag{3.117}$$

By conjugation, we introduce the differential operator

$$L_x^f = \Delta(x) f(x\partial_x) \Delta(x)^{-1}. \tag{3.118}$$

Then, we see that the Schur functions satisfy the differential equations

$$L_x^f s_\lambda(x) = f(\lambda+\delta) s_\lambda(x) \qquad (\lambda \in \mathcal{P}_n) \tag{3.119}$$

for all $f(\xi) = \mathbb{C}[\xi]^{\mathfrak{S}_n}$. In particular, for $L_x^{(r)} = L_x^{e_r}$ we have

$$L_x^{(r)} s_\lambda(x) = e_r(\lambda+\delta) s_\lambda(x) \quad (\lambda \in \mathcal{P}_n, r = 0, 1, \dots, n), \tag{3.120}$$

where

$$L_x^{(r)} = \sum_{|K|=r} \frac{1}{\Delta(x)} (x\partial_x)^K \Delta(x) = \sum_{|K|=r} \sum_{I \sqcup J = K} \frac{(x\partial_x)^I (\Delta(x))}{\Delta(x)} (x\partial_x)^J \tag{3.121}$$

with the notation $(x\partial_x)^I = \prod_{i \in I} x_i \partial_{x_i}$.

3.9 Link to the Representation Theory of GL_n (Overview)

In this section, we outline how the Schur functions arise, and how their properties are interpreted, in the context of representation theory of general linear groups. For the detail, see Goodman–Wallach [9] for example.

By a *representation* of a group G, we mean a $\mathbb{C}$-vector space M endowed with a group homomorphism $\pi_M : G \to \mathrm{GL}_{\mathbb{C}}(M)$, where $\mathrm{GL}_{\mathbb{C}}(M)$ denotes the group of invertible $\mathbb{C}$-linear transformations of M. In this situation, we also say that M is a *G-module*, and use the notation of the left action $g.v = \pi_M(g)(v)$ of $g \in G$ on $v \in M$. Suppose that M is finite-dimensional, and fix a $\mathbb{C}$-basis $v_1, \ldots, v_N$ of M. For each $g \in G$, we take the matrix representation $\Phi(g) = \big(\varphi_{ij}(g)\big)_{i,j=1}^N$ of $\pi_M(g) : M \to M$ with respect to the basis $(v_1, \ldots, v_N)$:

$$g.v_j = \pi_M(g)(v_j) = \sum_{i=1}^{N} v_i \varphi_{ij}(g) \quad (i = 1, \ldots, N). \tag{3.122}$$

Then we obtain an $N \times N$ matrix $\Phi_M(g) = \Phi(g)$ whose entries are functions on G satisfying the condition

$$\Phi(1_G) = I_N, \quad \Phi(g_1 g_2) = \Phi(g_1)\Phi(g_2), \quad \Phi(g^{-1}) = \Phi(g)^{-1}. \tag{3.123}$$

3.9.1 Polynomial Representations of GL_n

We consider the case of the general linear group $\mathrm{GL}_n = \mathrm{GL}_n(\mathbb{C})$ of degree n. Expressing a general element of GL_n as $g = (g_{ij})_{i,j=1}^n$, we regard g_{ij} $(1 \le i, j \le n)$ as the canonical coordinates of GL_n. A representation M of GL_n is called a *polynomial representation* if the matrix elements $\varphi_{ij}(g)$ are all polynomials of the coordinates g_{ij} $(1 \le i, j \le n)$. It is known that any polynomial representation is completely reducible, and the isomorphism classes of irreducible representations are parametrized by the partitions $\lambda \in \mathcal{P}$ with $\ell(\lambda) \le n$. Namely, for each $\lambda \in \mathcal{P}_n$, there exists an irreducible polynomial representation $V(\lambda) = V_n(\lambda)$ (with highest weight λ), uniquely determined up to isomorphism, such that $V(\lambda) \not\simeq V(\mu)$ if $\lambda \ne \mu$, and that any polynomial representation M is decomposed into a direct sum of the form

$$M \simeq \bigoplus_{\lambda \in \mathcal{P}_n} V(\lambda)^{\oplus m_\lambda} \tag{3.124}$$

with some multiplicities $m_\lambda \in \mathbb{N}$. We remark that $V(1^r) = V(\varpi_r)$, $\varpi_r = \varepsilon_1 + \cdots + \varepsilon_r$ (fundamental weights) $(r = 0, 1, \ldots, n)$ attached to single columns are the alternating tensor representation $\Lambda^r(V)$ of the vector space $V = \mathbb{C}$ on which GL_n is defined, and $V((l)) = V(l\varepsilon_1)$ $(l = 0, 1, 2, \ldots)$ attached to single rows are the symmetric tensor representation $S^l(V)$.

We denote by $H_n \subseteq \mathrm{GL}_n$ the diagonal subgroup of GL_n. Expressing a general element of H_n as $g_x = \mathrm{diag}(x_1, \ldots, x_n)$, we regard $x = (x_1, \ldots, x_n)$ as coordinates of H_n, and identify H_n with $(\mathbb{C}^*)^n$, $\mathbb{C}^* = \mathbb{C}\setminus\{0\}$. For each polynomial representation M of GL_n, we define the function $\mathrm{ch}_M(x)$ of $x = (x_1, \ldots, x_n) \in H_n = (\mathbb{C}^*)^n$ by

$$\mathrm{ch}_M(x) = \mathrm{tr}(\pi_M(g_x) : M \to M) = \mathrm{tr}\,\Phi_M(g_x), \tag{3.125}$$

and call it the *character* of the representation M. For each $\mu \in \mathbb{N}^n$, we denote by

$$M_\mu = \{v \in M \mid g_x.v = x^\mu v \ \ (x \in H_n)\} \subseteq M \tag{3.126}$$

the subspace of *weight* μ. Since M decomposes into the direct sum $M = \bigoplus_{\mu\in\mathbb{N}^n} M_\mu$ of weight subspaces, we have

$$\mathrm{ch}_M(x) = \sum_{\mu\in\mathbb{N}^n} (\dim_{\mathbb{C}} M_\mu) x^\mu \in \mathbb{C}[x]. \tag{3.127}$$

In this sense, the character $\mathrm{ch}_M(x)$ provides the generating function for counting the weight multiplicities in M. Note that $\mathrm{ch}_M(1) = \dim_{\mathbb{C}} M$. Also, for two polynomial representations M, N, the character of the tensor product representation $M \otimes N$ is given by the multiplication of the two characters as functions on H_n, namely $\mathrm{ch}_{M\otimes N}(x) = \mathrm{ch}_M(x)\mathrm{ch}_N(x)$.

A fundamental fact in the representation theory of GL_n is that the Schur function $s_\lambda(x)$ attached to each $\lambda \in \mathcal{P}_n$ appears as the character of the irreducible polynomial representation $V(\lambda)$, namely, $\mathrm{ch}_{V(\lambda)}(x) = s_\lambda(x)$.

3.9.2 *Weyl Character Formula and Branching Rules*

In the context of representation theory, the determinant representation

$$s_\lambda(x) = \frac{\det(x_i^{\lambda_j+n-j})_{i,j=1}^n}{\det(x_i^{n-j})_{i,j=1}^n} = \frac{\Delta_{\lambda+\delta}(x)}{\Delta(x)} \tag{3.128}$$

is called the *Weyl character formula*. The combinatorial description of $s_\lambda(x)$ in terms of semi-standard tableaux arises from the *branching rule* for the restriction of $V(\lambda)$ along the tower of subgroups

$$\mathrm{GL}_n \supset \mathrm{GL}_{n-1} \supset \cdots \supset \mathrm{GL}_1. \tag{3.129}$$

In fact, if we restrict the representation $V(\lambda) = V_n(\lambda)$ ($\lambda \in \mathcal{P}_n$) from GL_n to GL_{n-1}, it decomposes into the direct sum

$$V_n(\lambda) \simeq \bigoplus_{\mu \in \mathcal{P}_{n-1},\ \lambda/\mu:\ \text{h-strip}} V_{n-1}(\mu) \qquad (\lambda \in \mathcal{P}_n) \tag{3.130}$$

of irreducible GL_{n-1}-modules. Passing to the level of characters, this multiplicity-free decomposition of $V_n(\lambda)$ gives rise to the recurrence formula for Schur functions of Theorem 3.3 with respect to the number of variables. Repeating this restriction procedure, we find that $V(\lambda) = V_n(\lambda)$ for each $\lambda \in \mathcal{P}_n$ has a $\mathbb{C}$-basis v_T ($T \in \mathrm{SSTab}_n(\lambda)$) parameterized by the semi-standard tableaux of shape λ such that $g_x.v_T = x^{\mathrm{wt}(T)} v_T$:

$$V(\lambda) = \bigoplus_{\mu \in \mathbb{N}^n} V(\lambda)_\mu, \quad V(\lambda)_\mu = \bigoplus_{T \in \mathrm{SSTab}_n(\lambda)_\mu} \mathbb{C}\, v_T. \tag{3.131}$$

This gives rise to the tableau representation

$$s_\lambda(x) = \sum_{T \in \mathrm{SSTab}_n(\lambda)} x^{\mathrm{wt}(T)} = \sum_{\mu \in \mathbb{N}^n} K_{\lambda,\mu}\, x^\mu \tag{3.132}$$

of the character $s_\lambda(x)$; in particular, the Kostka numbers count the multiplicities of weights in $V(\lambda)$, i.e. $K_{\lambda,\mu} = \dim_{\mathbb{C}} V(\lambda)_\mu$. In the language of representation theory, we have shown in this chapter that, from the Weyl character formula, one can derive the combinatorial description of the weight subspaces of irreducible representations $V(\lambda)$ ($\lambda \in \mathcal{P}_n$)

3.9.3 $(\mathbf{GL}_m, \mathbf{GL}_n)$ *Duality*

We also give a remark on the Cauchy formula for Schur functions. We consider the space $\mathrm{Mat}_{m,n} = \mathrm{Mat}_{m,n}(\mathbb{C})$ of all $m \times n$ matrices. Denoting a general element of $\mathrm{Mat}_{m,n}$ as $T = \big(t_{ij}\big)_{1 \le i \le m;\, 1 \le j \le n}$, we regard t_{ij} as the canonical coordinates of $\mathrm{Mat}_{m,n}$. Then the coordinate ring of $\mathrm{Mat}_{m,n}$ is identified with the ring of polynomials in t_{ij}, i.e. $\mathcal{A}(\mathrm{Mat}_{m,n}) = \mathbb{C}[t_{ij}\ (1 \le i \le m,\ 1 \le j \le n)]$. We regard $\mathcal{A}(\mathrm{Mat}_{m,n})$ as a representation of the product group $\mathrm{GL}_m \times \mathrm{GL}_n$ through the action of $(g, h) \in \mathrm{GL}_m \times \mathrm{GL}_n$ defined by

$$((g, h).\varphi)(T) = \varphi(g^{\mathrm{t}}\, T\, h) \qquad (\varphi \in \mathcal{A}(\mathrm{Mat}_{m,n}),\ T \in \mathrm{Mat}_{m,n}). \tag{3.133}$$

Then it turns out that $\mathcal{A}(\mathrm{Mat}_{m,n})$ has the irreducible decomposition

$$\mathcal{A}(\mathrm{Mat}_{m,n}) \simeq \bigoplus_{\ell(\lambda) \le \min\{m,n\}} V_m(\lambda) \otimes V_n(\lambda), \tag{3.134}$$

where the sum is over all partitions λ with $\ell(\lambda) \le \min\{m, n\}$. From this $(\mathrm{GL}_m, \mathrm{GL}_n)$ *duality*, we obtain the identity

$$\mathrm{ch}_{\mathcal{A}(\mathrm{Mat}_{m,n})}(x, y) = \sum_{\ell(\lambda)\le\min\{m,n\}} \mathrm{ch}_{V_m(\lambda)}(x)\, \mathrm{ch}_{V_n(\lambda)}(y) \tag{3.135}$$

for the (formal) character of the $\mathrm{GL}_m \times \mathrm{GL}_n$-module $\mathcal{A}(\mathrm{Mat}_{m,n})$, which is precisely the Cauchy formula (3.59) for Schur functions. In fact, for each $(x, y) \in H_m \times H_n$, the action of $(g_x, g_y) \in \mathrm{GL}_m \times \mathrm{GL}_n$ on the coordinates t_{ij} is given by

$$(g_x, g_y).t_{ij} = x_i t_{ij} y_j \quad (1 \le i \le m,\ 1 \le j \le n). \tag{3.136}$$

Hence, (g_x, g_y) acts on the monomials $t^A = \prod_{i=1}^m \prod_{j=1}^n t_{ij}^{a_{ij}}$ attached to $A = (a_{ij})_{ij} \in \mathrm{Mat}_{m,n}(\mathbb{N})$ by

$$(g_x, g_y).t^A = \prod_{i=1}^m \prod_{j=1}^n (x_i t_{ij} y_j)^{a_{ij}} = x^{\mu(A)} t^A y^{\nu(A)}, \tag{3.137}$$

where the weights $\mu(A) \in \mathbb{N}^m$ and $\nu(A) \in \mathbb{N}^n$ are the row sum and the column sum of A respectively, i.e. $\mu(A)_i = \sum_{j=1}^n a_{ij}$, $\nu(A)_j = \sum_{i=1}^m a_{ij}$. Noting that

$$\mathcal{A}(\mathrm{Mat}_{m,n}) = \bigoplus_{A\in\mathrm{Mat}_{m,n}(\mathbb{N})} \mathbb{C}\, t^A, \tag{3.138}$$

we obtain

$$\begin{aligned} \mathrm{ch}_{\mathcal{A}(\mathrm{Mat}_{m,n})}(x, y) &= \sum_{A\in\mathrm{Mat}_{m,n}(\mathbb{N})} x^{\mu(A)} y^{\nu(A)} \\ &= \sum_{A=(a_{ij})} \prod_{i=1}^m \prod_{j=1}^n (x_i y_j)^{a_{ij}} = \prod_{i=1}^m \prod_{j=1}^n \frac{1}{1 - x_i y_j}. \end{aligned} \tag{3.139}$$

Since $\mathrm{ch}_{V_m(\lambda)}(x) = s_\lambda(x)$ and $\mathrm{ch}_{V_n(\lambda)}(y) = s_\lambda(y)$, formula (3.134) implies the Cauchy formula

$$\prod_{i=1}^m \prod_{j=1}^n \frac{1}{1 - x_i y_j} = \sum_{\ell(\lambda)\le\min\{m,n\}} s_\lambda(x)\, s_\lambda(y). \tag{3.140}$$