1 Construction of Haar Measure

Definition 1.1. A family $\mathfrak{G}$ of linear transformations on a linear topological space $\mathfrak{X}$ is said to be **equicontinuous on a subset K** of $\mathfrak{X}$ if for every neighborhood V of the origin in $\mathfrak{X}$ there is a neighborhood U of the origin such that the following condition holds

$$\text{if } k_1, k_2 \in K \text{ and } k_1 - k_2 \in U, \text{ then } \mathfrak{G}(k_1 - k_2) \subseteq V$$

that is $T(k_1 - k_2) \in V$ for all $T \in \mathfrak{G}$.

Theorem 1.2 (Kakutani). *Let K be a compact, convex subset of a locally convex linear topological space $\mathfrak{X}$, and let $\mathfrak{G}$ be a group of linear mappings which is equicontinuous on K and such that $\mathfrak{G}(K) \subseteq K$. Then there exists a point $p \in K$ such that*

$$T(p) = p \quad \forall T \in \mathfrak{G}$$

Proof. By Zorn's lemma, K contains a minimal non-void compact convex subset K_1 such that $\mathfrak{G}(K_1) \subseteq K_1$. If K_1 contains just one point then the proof is complete. If this is not the case, the compact set $K_1 - K_1$ contains some point other than the origin.

Thus, there exists a neighborhood V of the origin such that $\bar{V} \not\supseteq K_1 - K_1$.

There is a convex neighborhood V_1 of the origin such that $\alpha V_1 \subseteq V$ for $|\alpha| \leq 1$.

By the equicontinuity of $\mathfrak{G}$ on the set K_1, there is a neighborhood U_1 of the origin such that if $k_1, k_2 \in K_1$ and $k_1 - k_2 \in U_1$ then $\mathfrak{G}(k_1 - k_2) \subseteq V_1$.

Because each $T \in \mathfrak{G}$ is invertible, T maps open sets to open sets (open mapping theorem) and $T(A \cap B) = TA \cap TB$ for any sets A, B.

Since T is linear,

$$T\text{convex-hull}(A) = \text{convex-hull}T(A)$$

for any set A.

Because $\mathfrak{G}$ is a group, $\mathfrak{G}(\mathfrak{G}A) = \mathfrak{G}A$ for any set A.

Thus

$$\begin{aligned} U_2 &:= \text{convex-hull}(\mathfrak{G}U_1 \cap (K_1 - K_1)) \\ &= \text{convex-hull}(\mathfrak{G}(U_1 \cap (K_1 - K_1))) \subseteq V_1 \end{aligned}$$

is relatively open in $K_1 - K_1$ and satisfies $\mathfrak{G}U_2 = U_2 \not\supseteq K_1 - K_1$. By continuity, $\mathfrak{G}\bar{U}_2 = \bar{U}_2$. Define

$$\infty > \delta := \inf\{a : a > 0, aU_2 \supseteq K_1 - K_1\} \geq 1$$

and $U := \delta U_2$. For each $0 < \epsilon < 1$,

$$(1+\epsilon)U \supseteq K_1 - K_1 \not\subseteq (1-\epsilon)\bar{U}.$$

The family of relatively open sets $\{2^{-1}U + k\}, k \in K_1$, is a covering of K_1. Let $\{2^{-1}U + k_1, \ldots, 2^{-1}U + k_n\}$ be a finite sub-covering and let $p = (k_1 + \ldots k_n)/n$. If k is any point in K_1, then $k_i - k \in 2^{-1}U$ for some $1 \leq i \leq n$. Since $k_i - k \in (1+\epsilon)U$ for all i and all $\epsilon > 0$, we have

$$p \in \frac{1}{n}\left(2^{-1}U + (n-1)\cdot(1+\epsilon)U\right) + k.$$

For $\epsilon = \frac{1}{4(n-1)}$, we have $p \in (1 - \frac{1}{4n})U + k$ for each $k \in K_1$. Let

$$K_2 = K_1 \cap \bigcap_{k \in K_1} \left((1 - \frac{1}{4n})\bar{U} + k\right) \neq \emptyset.$$

Because $(1-\frac{1}{4n})\bar{U} \not\supseteq K_1 - K_1$, we have $K_2 \neq K_1$. The closed set K_2 is clearly convex. Further since $T(a\bar{U}) \subseteq a\bar{U}$ for $T \in \mathfrak{G}$, we have

$$T(a\bar{U}+k) \subseteq a\bar{U} + Tk \text{ for all } T \in \mathfrak{G}, k \in K_1.$$

Recalling $TK_1 = K_1$ for $T \in \mathfrak{G}$, we find that $\mathfrak{G}K_2 \subseteq K_2$, which contradicts the minimality of K_1. $\square$

Theorem 1.3 (Haar Measure). *Let G be a compact group. Let $\mathcal{C}(G)$ be the space of continuous maps from G to $\mathbb{C}$. Then, there is a unique linear form*

$$m : \mathcal{C}(G) \longrightarrow \mathbb{C}$$

having the following properties:

1. $m(f) \geq 0$ *for* $f \geq 0$ *(m is positive).*
2. $m(\mathbb{1}) = 1$ *(m is normalized).*
3. $m({}_s f) = m(f)$ *where* ${}_s f$ *is defined as the function*

   $${}_s f(g) = f(s^{-1}g) \quad s, g \in G$$

   (m is left invariant).
4. $m(f_s) = m(f)$ *where* $f_s(g) = f(gs)$ *for* $s, g \in G$ *(m is right invariant).*

Proof. For $f \in \mathcal{C}(G)$, let $\mathcal{C}_f$ denote the convex hull of all left translates of f. The elements of $\mathcal{C}_f$ are finite sums of the form:

$$g(x) = \sum_{\texttt{finite}} a_i f(s_i x) \qquad a_i > 0, \sum_{\texttt{finite}} a_i = 1$$

Clearly

$$||g|| = \max\{|g(x)| : x \in G\} \leq ||f||$$

Thus all sets $\mathcal{C}_f(x) = \{g(x) : g \in \mathcal{C}_f\}$ are bounded and relatively compact in $\mathbb{C}$. Since G is compact, f is *uniformly continuous*, namely for all $\epsilon > 0$, $\exists$ a neighborhood $V = V_\epsilon$ of the identity element $e \in G$ such that:

$$y^{-1}x \in V \Rightarrow |f(x) - f(y)| < \epsilon$$

Since $(s^{-1}y)^{-1}s^{-1}x = y^{-1}x$, we also have

$$|{}_sf(y) - {}_sf(x)| < \epsilon \text{ whenever } y^{-1}x \in V$$

Since the functions g are convex combinations of functions of the form ${}_sf$,

$$|g(y) - g(x)| < \epsilon \text{ whenever } y^{-1}x \in V$$

Thus the set $\mathcal{C}_f$ is equicontinuous. By Ascoli's theorem $\mathcal{C}_f$ is relatively compact in $\mathcal{C}(G)$. Define the compact convex set $K_f = \bar{\mathcal{C}}_f$ in $\mathcal{C}(G)$. The compact group G acts by left translations (isometrically) on $\mathcal{C}(G)$ and leaves $\mathcal{C}_f$ and hence K_f invariant. By Kakutani's Theorem 1.2, there is a fixed point g of this action G in K_f. Such a fixed point satisfies by definition

$${}_sg = g \;\; (\forall s \in G) \;\Rightarrow\; g(s^{-1}) = {}_sg(e) = c \;\; (\forall s \in G)$$

for some constant c.

By the definition of the set K_f, given any $\epsilon > 0$ there exists a finite set $\{s_1, \ldots s_n\}$ in G and $a_i > 0$ such that

$$\sum_1^n a_i = 1 \quad \text{and} \quad \left| c - \sum_1^n a_i f(s_i x) \right| < \epsilon \ \ (\forall x \in G) \qquad (1.1)$$

We first show that there is **only one** constant function K_f. Start the same construction as above, only now using *right* translations of f (e.g. we can apply the preceding construction to the opposite group G' of G , or the function $f' = f(x^{-1})$), obtaining a relatively compact set $\mathcal{C}'_f$ with compact convex closure K'_f containing a constant function c'. It will be enough to show $c = c'$.(all constants c in K_f must be equal to *one chosen constant* c' of K'_f and conversely.)

There is certainly a finite combination of right translates which is close to c' namely

$$|c' - \sum b_j f(xt_j)| < \epsilon \quad (\text{ for some } t_j \in G, b_j > 0 \text{ with } \sum b_j = 1)$$

Let us multiply this inequality by a_i and put $x = s_i$ to get

$$|c' a_i - \sum a_i b_j f(s_i t_j)| < \epsilon a_i \tag{1.2}$$

Summing over i, we obtain

$$|c' \sum a_i - \sum_{i,j} a_i b_j f(s_i t_j)| < \epsilon \sum a_i = \epsilon \tag{1.3}$$

Operating symmetrically on Equation (1.1) (multiplying by b_j, putting $x = t_j$ and summing over j), we find:

$$|c - \sum_{i,j} a_i b_j f(s_i t_j)| < \epsilon \tag{1.4}$$

Subtracting (or adding) Equation (1.3) from (1.4) we get $|c - c'| < 2\epsilon$. Since ϵ was arbitrary this completes the proof.

From now on **the** constant c in K_f will be denoted by $m(f)$. It is the only constant function which can be approximated arbitrarily close with convex combinations of left or right translates of f.

The following properties are obvious:

- $m(\mathbb{1}) = 1$ since $K_f = \{1\}$ if $f = 1$.
- $m(f) \geq 0$ if $f \geq 0$.
- $m(af) = am(f)$ for any $a \in \mathbb{C}$ (since $K_{af} = K_f$).
- $m({}_sf) = m(f) = m(f_s)$ (by uniqueness)

The proof will be complete if we show that m is *additive* (hence *linear*). Let us take $f, g \in \mathcal{C}(G)$ and start with Equation (1.1) above with $c = m(f)$. Further let

$$h(x) = \sum a_i g(s_i x)$$

Since $h \in \mathcal{C}_g$, we certainly have $\mathcal{C}_h \subseteq \mathcal{C}_g$ whence $K_h \subseteq K_g$. But the set K_g contains only one constant: $m(h) = m(g)$.

We can write

$$|m(h) - \sum b_j h(t_j x)| < \epsilon$$

for finitely many suitable $t_j \in G$ and $b_j > 0$ with $\sum b_j = 1$. Using the definition of h and $m(h) = m(g)$, this implies

$$|m(g) - \sum_{i,j} a_i b_j g(s_i t_j x)| < \epsilon \tag{1.5}$$

However multiplying Equation (1.1) by b_j and replacing x by $t_j x$ and summing over j we find

$$|m(f) - \sum_{i,j} a_i b_j f(s_i t_j x)| < \epsilon \tag{1.6}$$

Adding Equation (1.5) and (1.6), this implies

$$|m(f) + m(g) - \sum_{i,j} a_i b_j (f+g)(s_i t_j x)| < 2\epsilon$$

Thus the constant $m(f) + m(g)$ is in K_{f+g}. However note that the only constant in this compact convex set is $m(f+g)$. This completes the proof. □

1.1 Exercises

Exercise 1.4. Let m be the normalized Haar measure of a compact group G. For $f \in \mathcal{C}(G)$ or $L^1(G)$ show that $m(f) = m(\tilde{f})$ where the function $\tilde{f}$ is defined as the function $\tilde{f}(x) = f(x^{-1})$. This equality is usually written as

$$\int_G f(x)dx = \int_G f(x^{-1})dx$$

Hint: Observe that $f \to m(\tilde{f})$ is a Haar measure on G and use the uniqueness part of the Theorem on Haar measures, Theorem 1.3

Before stating the next exercise we need a definition

Definition 1.5 (Semidirect products). Let L be a group and assume it contains a normal subgroup G and a subgroup H such that $GH = L$ and $G \cap H = \{e\}$. That is, suppose one can select exactly one element h from each coset of G so that $\{h\}$ forms a subgroup H. If H is also normal then L is isomorphic with the direct product $G \times H$. If H fails to be normal, we can still reconstruct L if we know how the inner automorphisms ρ_h behave on G. Namely for $x_j \in G$ and $h_j \in H$ $(j = 1, 2)$, we have:

$$(x_1 h_1)(x_2 h_2) = x_1 h_1 x_2 h_1^{-1} h_1 h_2 = (x_1 \rho_{h_1}(x_2))\, h_1 h_2$$

The construction just given can be cast in an abstract form. Let G and H be groups and suppose there is a homomorphism $h \to \tau_h$ which carries H onto a group of automorphisms of G, namely $\tau_h \circ \tau_{h'} = \tau_{hh'}$ for $h, h' \in H$. Let $G \circledS H$ denote the cartesian product of G and H. For (x, h) and (x', h') in $G \circledS H$, define:

$$(x, h)(x', h') = (x\,(\tau_h(x'))\,, hh')$$

Then $G \circledS H$ is a group; it is called a *semidirect product* of G and H. Its identity is (e_1, e_2) where e_1 and e_2 are the identities of G and H respectively. The inverse of (x, h) is $(\tau_{h^{-1}}(x^{-1}), h^{-1})$. Let

$$G_1 := \{(x, e_2) : x \in G\}$$

and

$$H_1 := \{(e_1, h) : h \in H\}$$

Then G_1 is a normal subgroup of $G \circledS H$ and H_1 is a subgroup. Since

$$(e_1, h) \cdot (x, e_2) \cdot (e_1, h)^{-1} = (\tau_h(x), e_2)$$

the inner automorphism $\rho_{(e_1,h)}$ for $(e_1, h) \in H_1$ reproduces the action τ_h on G. Thus every semidirect product is obtained by the process described in the previous paragraphs.

Exercise 1.6. Let G and H be compact groups and let $G \circledS H$ be a semidirect product of G and H. Suppose also that the mapping $(x, h) \to \tau_h(x)$ is a continuous mapping of $G \times H$ onto G. In particular, each τ_h is a homeomorphism of G onto itself. Show that the semidirect product $G \circledS H$ with the product topology is a compact group. What is the Haar measure on $G \circledS H$ in terms of the Haar measures on G and H ?

Exercise 1.7. Let $O_n(\mathbb{R})$ be the group of $n \times n$ orthogonal matrices. Suppose that $Z_{ij}, 1 \leq i \leq j \leq n$ are *i.i.d.* standard normal random variables. Let U be the random orthogonal matrix with rows obtained by applying the Gram-Scmidt process to the vectors $(Z_{11}, \ldots, Z_{1n})$, ..., $(Z_{n1}, \ldots Z_{nn})$. Show that U is distributed according to the Haar measure on $O_n(\mathbb{R})$.

2 Representations, General Constructions

For E , a complex Banach space, let $Gl(E)$ denote the group of continuous isomorphisms of E onto itself. A **representation** $\boldsymbol{\pi}$ of a compact group G in E is a homomorphism $\boldsymbol{\pi}$:

$$\boldsymbol{\pi} : G \longrightarrow Gl(E)$$

for which all the maps $G \to E$ defined as $s \to \boldsymbol{\pi}(s)v \ (v \in E)$ are continuous. The space $E = E_{\boldsymbol{\pi}}$ in which the representation takes place is called the **representation space** of $\boldsymbol{\pi}$. A representation $\boldsymbol{\pi}$ of a group G in a vector space E canonically defines an **action** (also denoted by $\boldsymbol{\pi}$)

$$\begin{aligned} \boldsymbol{\pi} : G \times E \quad &\longrightarrow \quad E \\ (s, v) \quad &\longrightarrow \quad \boldsymbol{\pi}(s)v \end{aligned}$$

The definition requires this action to be *separately continuous*. The action is then automatically globally continuous.

We say that a representation $\boldsymbol{\pi}$ is **unitary** when $E = H$, is a Hilbert space and each operator $\boldsymbol{\pi}$(s) $(s \in G)$ is a unitary operator (i.e. each $\boldsymbol{\pi}$(s) is *isometric and surjective*). Thus $\boldsymbol{\pi}$ is unitary when $E = H$ is a Hilbert space and

$$\boldsymbol{\pi}(s)^* = \boldsymbol{\pi}(s)^{-1} = \boldsymbol{\pi}(s^{-1}) \quad (s \in G)$$

The representation $\boldsymbol{\pi}$ of G in E is said to be **irreducible** when E and $\{0\}$ are distinct and are the only two closed invariant subspaces under all operators $\boldsymbol{\pi}$(s) $(s \in G)$ (*topological irreducibility*).

Two representations $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ of the same group G are called **equivalent** when the two spaces over which they act are G *-isomorphic*, namely there exists a continuous isomorphism $A : E \to E'$ of their respective spaces with

$$A(\boldsymbol{\pi}(s)v) = \boldsymbol{\pi}'(s)Av \qquad (s \in G, v \in E)$$

More generally, continuous linear operators $A : E \to E'$ satisfying all commutation relations $A(\boldsymbol{\pi}(s)) = \boldsymbol{\pi}'(s)A$ for all $s \in G$ are called **intertwining operators** or G -morphisms (from $\boldsymbol{\pi}$ to $\boldsymbol{\pi}'$) and their set is a vector space denoted either by

$$\mathrm{Hom}_G(E, E') \text{ or } \mathrm{Hom}(\boldsymbol{\pi}, \boldsymbol{\pi}')$$

Proposition 2.8. *Let $\boldsymbol{\pi}$ be a unitary representation of G in the Hilbert space H . If H_1 is an invariant subspace of H (with respect to all operators $\boldsymbol{\pi}(s), s \in G$), then the orthogonal space $H_2 = H_1^{\perp}$ of H_1 in H is also invariant.*

Proof. We need to show that if $v \in H$, $v \perp H_1$ then $\boldsymbol{\pi}(s)v$ is also orthogonal to H_1 for all s in G . For any $x \in H_1$,

$$\langle x, \boldsymbol{\pi}(s)v \rangle = \langle \boldsymbol{\pi}(s)^* x, v \rangle = \langle \boldsymbol{\pi}(s^{-1})x, v \rangle = 0$$

since by assumption $\boldsymbol{\pi}(s^{-1})x$ also lies in H_1.

□

Proposition 2.9. *Let $\boldsymbol{\pi}$ be a representation of a compact group G in a Hilbert space H . Then there exists a positive definite hermitian form φ which is invariant under the G -action, and which defines the same topological structure on H .*

Proof. By continuity of the mappings $s \to \boldsymbol{\pi}(s)v$, the mappings

$$s \longrightarrow \langle \boldsymbol{\pi}(s)v, \boldsymbol{\pi}(s)w \rangle \quad (v, w \in H)$$

are also continuous (by continuity of scalar product in $H \times H$). We can thus define

$$\varphi(v, w) = \int_G \langle \boldsymbol{\pi}(s)v, \boldsymbol{\pi}(s)w \rangle ds$$

using the Haar integral.

It is clear that φ is hermitian and positive. Let us show that it is non-degenerate and defines the same topology on H . Since G is compact, $\boldsymbol{\pi}(G)$ is also compact in $Gl(H)$ (with the strong topology). In particular, $\boldsymbol{\pi}(G)$ is *simply bounded* and thus *uniformly bounded* (uniform boundedness principle $\equiv$ Banach-Steinhaus theorem). Thus, there exists a positive constant $M > 0$ with

$$||\boldsymbol{\pi}(s)v|| \leq M||v|| \qquad (\forall s \in G, v \in H)$$

This implies

$$||v|| = ||\boldsymbol{\pi}(s^{-1})\boldsymbol{\pi}(s)v|| \leq M||\boldsymbol{\pi}(s)v|| \leq M^2||v||$$

Thus

$$M^{-1}||v|| \leq ||\boldsymbol{\pi}(s)v|| \leq M||v||$$

Squaring and Integrating over G , we find

$$M^{-2}||v||^2 \leq \varphi(v,v) \leq M^2||v||^2$$

Thus $\varphi(v,v) = 0$ implies $||v|| = 0$ and $v = 0$. Thus φ and $||\cdot||^2$ induce equivalent topologies (equivalent norms) on H . Invariance of φ comes from the invariance of the Haar measure.

$$\begin{aligned} \varphi(\boldsymbol{\pi}(t)v, \boldsymbol{\pi}(t)w) &= \int_G \langle \boldsymbol{\pi}(st)v, \boldsymbol{\pi}(st)w \rangle ds = \int_G f(st)ds \\ &= \int_G f_t(s)ds = \int_G f(s)ds = \varphi(v,w) \end{aligned}$$

This shows that $\boldsymbol{\pi}$ is φ-*unitary* as desired. □

These propositions imply any representation of a compact group in a Hilbert space is equivalent to a unitary one, and any finite dimensional representation (the dimension of a representation is the dimension of its rep. space) is completely reducible (direct sum of irreducible ones.)

Definition 2.10 (left translations). In any space of functions on G , define the left translations by

$$[l(s)f](x) = f(s^{-1}x)$$

(If we do not want to identify elements of $\mathbb{L}^p(G)$ with functions or classes of functions, we can simply extend translations from $\mathcal{C}(G)$ to $\mathbb{L}^p(G)$ by continuity).

Thus we have

$$l(s) \circ l(t) = l(st)$$

and we get homomorphisms

$$l : G \to Gl(E), \quad s \to l(s)$$

with any $E = \mathbb{L}^p(G), 1 \leq p < \infty$.

Exercise 2.11. Check that these homomorphisms are continuous in the representation sense.

The above were the **left regular representations** of G. The **right regular representations** of G in the Banach space $\mathbb{L}^p(G)$ are defined similarly with

$$[r(s)f](x) = f(xs) \qquad (f \in \mathbb{L}^p(G))$$

With this definition, one has $r(s) \circ r(t) = r(st)$.

One can also consider the **biregular representations** of $l \times r$ of $G \times G$ in $\mathbb{L}^p(G)$ defined as

$$[l \times r(s,t)f](x) = f(s^{-1}xt) \qquad (f \in \mathbb{L}^p(G))$$

and its restriction to the diagonal $G \to G \times G, s \to (s,s)$ which is the **adjoint representation** of G . It is defined as

$$[Ad(s)f](x) = f(s^{-1}xs) \qquad (f \in \mathbb{L}^p(G))$$

The regular representations are faithful, i.e $\boldsymbol{\pi}(s) = \mathbb{1} \Leftrightarrow s = e$

Let $\boldsymbol{\pi} : G \to Gl(E)$ and $\boldsymbol{\pi}' : G' \to Gl(E')$ be two representations. We can define the **external direct sum** representation of $G \times G'$ in $E \oplus E'$ by

$$\boldsymbol{\pi} \oplus \boldsymbol{\pi}'(s, s') = \boldsymbol{\pi}(s) \oplus \boldsymbol{\pi}'(s') \qquad (s \in G, s' \in G')$$

When $G = G'$, we can restrict this external direct sum to the diagonal G of $G \times G$, obtaining the usual direct sum of $\boldsymbol{\pi}$and $\boldsymbol{\pi}'$

$$\begin{aligned} \boldsymbol{\pi} \oplus \boldsymbol{\pi}' : G &\to Gl(E \oplus E') \\ s &\to \boldsymbol{\pi}(s) \oplus \boldsymbol{\pi}'(s) \end{aligned}$$

The **external tensor product** $\boldsymbol{\pi} \otimes \boldsymbol{\pi}'$ as a representation of $G \times G'$ in $E \times E'$is defined as

$$\boldsymbol{\pi} \otimes \boldsymbol{\pi}'(s, s') = \boldsymbol{\pi}(s) \otimes \boldsymbol{\pi}'(s') \qquad (s \in G, s' \in G')$$

We assume the two spaces E ,E' are finite dimensional, thus this algebraic tensor product is complete; in general some completion has to be devised.

The **usual tensor product** of two representations of the same group G is the restriction to the diagonal of the external tensor product $(G = G')$ and is given by

$$\boldsymbol{\pi} \otimes \boldsymbol{\pi}'(s) = \boldsymbol{\pi}(s) \otimes \boldsymbol{\pi}'(s) \qquad (s \in G)$$

For a given finite dimensional representation $\boldsymbol{\pi} : G \to Gl(E)$, define the **contragredient representation** $\check{\boldsymbol{\pi}}$. This representation acts in the dual E' of E (namely the space of linear forms on E) and

$$\check{\boldsymbol{\pi}}(s) = {}^t\boldsymbol{\pi}(s^{-1}) \qquad (s \in G)$$

Since transposition reverses the order of composition of mappings, namely ${}^t(AB) = {}^tB^tA$, it is necessary to reverse the operations by taking the inverse in the group. The above construction allows us to conclude that $\check{\boldsymbol{\pi}}(st) = \check{\boldsymbol{\pi}}(s)\check{\boldsymbol{\pi}}(t)$ as is required for a representation.

Conjugate representation $\boldsymbol{\pi}$: When $E = H$ is a Hilbert space the conjugate $\bar{\boldsymbol{\pi}}$ of $\boldsymbol{\pi}$ is a representation acting on the conjugate $\bar{H}$ of H . Recall that $\bar{H}$ has the same underlying additive group as H , but with the scalar multiplication in $\bar{H}$ twisted by complex conjugation, namely the external operation of scalars is given by

$$(a, v) \longrightarrow a \cdot v = \bar{a} v \quad \text{(we use a dot in } \bar{H} \text{)}$$

The inner product $\langle \cdot, \cdot \rangle^{-}$ of $\bar{H}$ is defined as

$$\langle v, w \rangle^{-} = \overline{\langle v, w \rangle} = \langle w, v \rangle$$

This suggests that an element $v \in H$ is written as $\bar{v}$ when we consider it as an element of the dual Hilbert space $\bar{H}$. With this notation we have:

$$\overline{av} = \overline{a} \cdot \overline{v} \;\; (a \in \mathbb{C}) \qquad \text{and} \qquad \langle \bar{v}, \bar{w} \rangle^{-} = \overline{\langle v, w \rangle}$$

The identity map $H \to \bar{H}$, $v \to \bar{v}$ is an *anti-isomorphism*. The conjugate of $\boldsymbol{\pi}$ is defined as $\bar{\boldsymbol{\pi}}(s) = \boldsymbol{\pi}(s)$ *in* $\bar{H}$. Since the (complex vector) subspaces of H and $\bar{H}$ are the *same* by definition, $\boldsymbol{\pi}$ and $\bar{\boldsymbol{\pi}}$ are reducible or irreducible simultaneously. However it is important to distinguish these two representations (in particular they are not always *equivalent*). Any orthonormal basis (e_i) of H is also an orthonormal basis of $\bar{H}$, but a decomposition $v = \sum v_i e_i$ in H gives rise to the decomposition

$$\bar{v} = \sum \bar{v}_i \bar{e}_i \qquad \text{(complex conjugate components in } \bar{H} \text{)}$$

Thus the *matrix representations* associated with $\boldsymbol{\pi}$ and $\bar{\boldsymbol{\pi}}$ are complex conjugate to one another.

Exercise 2.12. Show that when $\boldsymbol{\pi}$ is unitary and finite dimensional, the contragredient $\check{\boldsymbol{\pi}}$ and the conjugate $\bar{\boldsymbol{\pi}}$ of $\boldsymbol{\pi}$ are equivalent.

2.1 Exercises

Exercise 2.13. Show that the left and right representations l and r of a group G (in any $\mathbb{L}^p(G)$ space) are equivalent.

Exercise 2.14. If $\boldsymbol{\pi}$and $\boldsymbol{\pi}'$ are two representations of the same group G (acting in respective Hilbert spaces H and H'), show that the matrix coefficients of $\boldsymbol{\pi} \otimes \boldsymbol{\pi}'$ (with respect to bases (e_i) in H and (e'_j) in H' and $e_i \otimes e'_j$ in $H \otimes H'$) are products of matrix coefficients of $\boldsymbol{\pi}$and $\boldsymbol{\pi}'$ (Kronecker product of matrices).

Exercise 2.15. Let $\mathbb{1}_n$ denote the identity representation of the group G in dimension n (the space of this identity representation is thus $\mathbb{C}^n$ and $\mathbb{1}_n(s) = id_{\mathbb{C}^n}$ for all $s \in G$). Show that for any representation $\boldsymbol{\pi}$of G ,

$$\boldsymbol{\pi} \otimes \mathbb{1}_n \text{ is equivalent to } \boldsymbol{\pi} \oplus \boldsymbol{\pi} \oplus \cdots \oplus \boldsymbol{\pi} \text{ (n terms)}$$

Exercise 2.16. (Schur's lemma) Let k be an algebraically closed field, V a finite dimensional vector space over k and Φ any irreducible set of operators in V (the only invariant subspaces, relatively to all operators belonging to Φ are V and $\{0\}$). Then, if an operator A commutes with all operators in Φ, A is a multiple of the identity operator (i.e. A is a scalar operator).

Hint: Take an eigenvalue a in the algebraically closed field k and consider $A - a \cdot I$, which still commutes with all operators of Φ. Show that the $Ker(A - a \cdot I)(\neq \{0\})$ is an invariant subspace.

3 Finite dimensional representations of compact groups (Peter-Weyl theorem)

Theorem 3.17 (Peter-Weyl). *Let G be a compact group. for any $s \neq e$ in G , there exists a finite dimensional irreducible representation $\boldsymbol{\pi}$ of G such that $\boldsymbol{\pi}(s) \neq \mathbb{1}$.*

Proof. We start with two Lemmas.

Lemma 3.18. *Let G be a compact group, $k : G \times G \to \mathbb{C}$ a continuous function and $K : \mathbb{L}^2 \to \mathcal{C}(G)$ the operator with kernel k, namely:*

$$(Kf)(x) = \int_G k(x, y) f(y) dy$$

Then K is a **compact** *operator. Moreover if $k(x, y) = \overline{k(y, x)}$ identically on $G \times G$, K is a Hermitian as an operator from $\mathbb{L}^2(G)$ to $\mathcal{C}(G)$.*

Lemma 3.19. *Let K be a compact Hermitian operator (in some Hilbert space H). Then the spectrum S of K consists of eigenvalues. Each eigenspace H_λ with respect to a non-zero eigenvalue $\lambda \in S$ is finite dimensional and the number of eigenvalues outside any neighborhood of 0 is finite. Moreover, $S \subseteq \mathbb{R}$ and*

$$||K|| = \sup\{|\lambda| : \lambda \in S\}$$

and the eigenspaces associated to distinct eigenvalues are orthogonal, i.e

$$H_\lambda \perp H_\mu \text{ for } \lambda \neq \mu \text{ in } S$$

Proof of Theorem 3.17: Assume that $s \neq e$ in G and take an open symmetric neighborhood $V = V^{-1}$ of e in G such that $s \notin V^2$. There exists a positive continuous function f such that

$$f(e) > 0 \ , \ f(x) = f(x^{-1}) = \check{f}(x) \ , \quad \text{Supp}(f) \subseteq V$$

where Supp(f) denotes the support of f, namely the complement of the largest open set on which f vanishes. Consider the function $\varphi = f * f$ defined as

$$\varphi(x) = \int_G f(y) f(y^{-1}x) dy$$

The support of φ is contained in V^2 and

$$\varphi(s) = 0 \ (s \notin V^2) \ , \ \varphi(e) = ||f||^2 > 0.$$

We also see that $l(s)\varphi \neq \varphi$. But the operator K with kernel $k(x,y) = f(y^{-1}x)$ is compact (see Lemma 3.18) and the convergence in quadratic mean of

$$f = f_0 + \sum f_i \quad , \quad f_i \in \mathsf{Ker}(K - \lambda_i) = H_i \ (\lambda_i \in \mathsf{Spec}(K))$$

implies that

$$\varphi = Kf = \sum K f_i = \sum \lambda_i f_i$$

where

$$f_i = \frac{1}{\lambda_i} K f_i \in \mathsf{Im}(K) \subseteq \mathcal{C}(G)$$

where we have uniform convergence holding in the series above. Since $l(s)\varphi \neq \varphi$, we must have $l(s) f_i \neq f_i$ for at least one index i. However the definition of the kernel k shows that

$$k(sx, sy) = k(x, y) = f(y^{-1}x) \qquad (s, x, y \in G)$$

The consequence of these identities is the translation invariance of all the eigenspaces H_i of K. The left regular representation restricted to a suitable finite dimensional subspace H_i (for any i, with $l(s)f_i \neq f_i$) will furnish an example of a finite dimensional representation $\boldsymbol{\pi}$ with $\boldsymbol{\pi}(s) \neq e$. □

The corollaries of this theorem are numerous and important.

Corollary 3.20. *A compact group is commutative if and only if all its finite dimensional irreducible representations have dimension 1.*

Proof. Exercise.

Corollary 3.21 (Peter-Weyl). *Any continuous function on a compact group is a uniform limit of (finite) linear combinations of coefficients of irreducible representations.*

Proof. Let $\boldsymbol{\pi}$ be a (finite dimensional) irreducible representation of the compact group G and take a basis in the representation space of $\boldsymbol{\pi}$ in order to be able to identify in $\boldsymbol{\pi} : G \to Gl_n(\mathbb{C})$, the coefficients of $\boldsymbol{\pi}$ being the continuous functions on G defined as

$$c^i_j : g \longrightarrow c^i_j(g) = \langle e_i, \boldsymbol{\pi}(g) e_j \rangle$$

More generally if u and v are elements of H, we can define the (function) coefficient c^u_v of $\boldsymbol{\pi}$ on G by

$$g \longrightarrow c^u_v(g) = \langle u, \boldsymbol{\pi}(g) v \rangle$$

These functions are obviously finite linear combinations of the previously defined matrix coefficients c^i_j. Introduce the subspace $V(\boldsymbol{\pi})$ of $\mathcal{C}(G)$ spanned by the c^i_j, or equivalently by all c^u_v for $v, u \in H_{\boldsymbol{\pi}}$. Observe that the subspaces of $\mathcal{C}(G)$ attached in this way to two *equivalent* representations $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ *coincide* namely, $V(\boldsymbol{\pi}) = V(\boldsymbol{\pi}')$. Thus we can form the algebraic sum (a priori this algebraic sum is not a direct sum)

$$A_G = \bigoplus V_{\boldsymbol{\pi}} \subseteq \mathcal{C}(G)$$

where the summation index $\boldsymbol{\pi}$ runs over all (classes of) finite dimensional irreducible representations of G . The corollary can be restated in the following form:

> A_G *is a dense subspace of the Banach space* $\mathcal{C}(G)$ *in the uniform norm*

But this algebraic sum A_G is a *subalgebra* of $\mathcal{C}(G)$ (the product of two continuous functions being the usual pointwise product). The product of the coefficients

$$c_v^u \text{ of } \boldsymbol{\pi} \quad \text{and} \quad \gamma_t^s \text{ of } \boldsymbol{\sigma}$$

is a coefficient of the representation $\boldsymbol{\pi} \otimes \boldsymbol{\sigma}$ (the coefficient of this representation with respect to the two vectors $u \otimes s$ and $v \otimes t$). Taking $\boldsymbol{\pi}$ and $\boldsymbol{\sigma}$ to be finite dimensional representations of G , $\boldsymbol{\pi} \otimes \boldsymbol{\sigma}$ will be finite dimensional, hence completely reducible and all its coefficients (in particular the product of c_u^v and γ_t^s) are finite linear combinations of coefficients of (finite dimensional) irreducible representations of G .

This subalgebra A_G of $\mathcal{C}(G)$ contains the constants, is stable under complex conjugation (because $\boldsymbol{\pi}$ is irreducible precisely when $\bar{\boldsymbol{\pi}}$ is irreducible) and separates points of G by the main Theorem 3.17. By the Stone-Weistrass theorem the proof is complete. $\square$

3.1 Exercises

Exercise 3.22. Let G be a compact totally discontinuous group. Show that A_G is the algebra of all locally constant functions on G . (Observe that a locally constant function on G is uniformly locally constant, hence can be identified with a function on a quotient G/H where H is some open subgroup of G . Conversely any finite dimensional representation of G must be trivial on an open subgroup H of G .)

Exercise 3.23. Let G be any compact group. Show that A_G consists of the continuous functions f on G for which the left and right translates of f generate a finite dimensional subspace of $\mathcal{C}(G)$. In particular if G_1 and G_2 are two compact groups, any continuous homomorphism $h : G_1 \to G_2$ has a transpose $t_h : A_2 \to A_1$ where $A_i = A_{G_i}$, defined by $t_h(f) = f \circ h$. A priori this transpose is a linear mapping $t_h : \mathcal{C}(G_2) \to \mathcal{C}(G_1)$.

Exercise 3.24. Let $G = U_n(\mathbb{C})$ with its canonical representation $\boldsymbol{\pi}$ in $V = \mathbb{C}^n$. Since $\boldsymbol{\pi}$ is unitary, we can identify $\bar{\boldsymbol{\pi}}$ with the contragredient of $\boldsymbol{\pi}$: it acts in the dual V^* of V.
(a) Let A_q^p denote the space of linear combinations of coefficients of the representation

$$\boldsymbol{\pi}_q^p = \bar{\boldsymbol{\pi}}^{\otimes p} \otimes \boldsymbol{\pi}^{\otimes q} \quad \text{in} \quad (V^*)^{\otimes p} \otimes V^{\otimes q} = T_q^p(V)$$

Prove that the sum of the subspaces A_q^p of $\mathcal{C}(G)$is an algebra A (show that $A_q^p A_s^r \subseteq A_{qs}^{pr}$), stable under conjugation (show that $\overline{A_q^p} = A_p^q$), which separates the points of G . Using the Stone-Weierstrass theorem, conclude that A is dense in $\mathcal{C}(G)$.
(b) Show that $A = A_G$. (use part(a) to prove that any irreducible representation of G appears as a subrepresentation of some $\boldsymbol{\pi}_q^p$, or in other words can be realized on a space of mixed tensors.)

Exercise 3.25. Let G be a closed subgroup of $U_n(\mathbb{C})$. Using the fact that any finite dimensional representation of G appears in the restriction of some finite dimensional representation of $U_n(\mathbb{C})$ (this is a consequence of the theory of induced representations), show that G is a real algebraic subvariety of $U_n(\mathbb{C})$. (The transpose of the embedding $G \hookrightarrow U_n(\mathbb{C})$ is the operation of restriction on polynomial functions and is surjective. Hence A_G is a quotient of the polynomial algebra A of $U_n(\mathbb{C})$. By the exercise 3.24, A is generated by the coordinate functions

$$U_n(\mathbb{C}) \longrightarrow \mathbb{C} \ , \ \ x = (x^i_j) \longmapsto x^i_j$$

and their conjugates.)

4 Decomposition of the regular representation

Lemma 4.1. *Let* $V \subseteq \mathbb{L}^2(G)$ *be a finite dimensional subspace which is invariant under the right regular representation of* G. *Then* V *consists of (classes of) continuous functions and each* $f \in V$ *can be written as*

$$f(x) = \mathit{Tr}(A\boldsymbol{\pi}(x)) \qquad \textit{for some } A \in \mathit{End}(V)$$

Here $\boldsymbol{\pi}$ *denotes the restriction of the right regular representation to* V.

Proof. Take an orthonormal basis (χ_i) of V and recall the coefficients c^i_j of $\boldsymbol{\pi}$ defined as

$$\boldsymbol{\pi}(x)\chi_i = \sum_j c^j_i(x)\chi_j, \qquad x \in G$$

For $f = \sum_i a^i \chi_i$, we have

$$\boldsymbol{\pi}(x)f = \sum_i a^i \boldsymbol{\pi}(x)\chi_i = \sum_{i,j} a^i c^j_i(x)\chi_j$$

Hence

$$f(x) = [r(x)f](e) = \sum_{i,j} c^j_i(x) a^i_j \qquad \text{(with } a^i_j = a^i \chi_j(e))$$

Thus,

$$f(x) = \mathrm{Tr}(A\boldsymbol{\pi}(x))$$

as claimed.

Let $(\boldsymbol{\pi}, V)$ be any finite dimensional representation of the compact group G. For any endomorphism $A \in$ End(V), we define the corresponding coefficient c_A of $\boldsymbol{\pi}$ by $c_A(x) = Tr(A \cdot \boldsymbol{\pi}(x))$. The right translates of these coefficients are easily identified as

$$\begin{aligned} [r(s)c_A](x) &= c_A(xs) = Tr(A\boldsymbol{\pi}(x)\boldsymbol{\pi}(s)) \\ &= Tr(\boldsymbol{\pi}(s) \cdot A\boldsymbol{\pi}(x)) = Tr(B\boldsymbol{\pi}(x)) = c_B(x) \end{aligned}$$

where $B = \boldsymbol{\pi}(s) \cdot A$.

Consider the representation of G in $\text{End}(V)$ defined by

$$l_{\boldsymbol{\pi}}(s)A = \boldsymbol{\pi}(s) \cdot A \qquad (s \in G, A \in \text{End}(V))$$

The above computations show that $A \to c_A$ is a G-morphism

$$c : \text{End}(V) \longrightarrow \mathcal{C}(G) \subseteq \mathbb{L}^2(G)$$

(intertwining $l_{\boldsymbol{\pi}}$ and r.)

Suppose now that $(\boldsymbol{\pi}, V)$ is an irreducible finite dimensional representation of the compact group G.

Write $\mathbb{L}^2(G, \boldsymbol{\pi})$ for the image of End(V) under the map c. Note that the vector space $\mathbb{L}^2(G, \boldsymbol{\pi})$ only depends on the equivalence class of π.

The representation $(l_{\boldsymbol{\pi}}, \text{End}(V))$ is equivalent to $\boldsymbol{\pi} \oplus \cdots \oplus \boldsymbol{\pi}$ (n times, where $n = \dim(V)$) and $\mathbb{L}^2(G, \boldsymbol{\pi})$ is r-invariant, so the restriction of r to $\mathbb{L}^2(G, \boldsymbol{\pi})$ is equivalent to $\boldsymbol{\pi} \oplus \cdots \oplus \boldsymbol{\pi}$ (m times for some $m \leq n$). Thus, $\mathbb{L}^2(G, \boldsymbol{\pi})$ has dimension $mn \leq n^2$.

If V' is a r-invariant subspace of $\mathbb{L}^2(G)$ such that the restriction of r to V' is equivalent to $\boldsymbol{\pi}$, then V' is a subspace of $\mathbb{L}^2(G, \boldsymbol{\pi})$ by Lemma 4.1.

Hence, $\mathbb{L}^2(G, \boldsymbol{\pi})$ is the sum of all subrepresentations of $(\mathbb{L}^2(G), r)$ which are equivalent to $\boldsymbol{\pi}$.

Definition 4.2. Let $\boldsymbol{\pi}$ be a finite dimensional irreducible representation of a compact group G. The space $\mathbb{L}^2(G, \boldsymbol{\pi})$ consisting of the sum of all subspaces of the right regular representation which are equivalent to $\boldsymbol{\pi}$ is called the *isotypical component* of $\boldsymbol{\pi}$ in $\mathbb{L}^2(G)$.

Note that a function $f \in \mathbb{L}^2(G)$ belongs to $\mathbb{L}^2(G, \boldsymbol{\pi})$ precisely when the right translates of f generate an invariant subspace (of the right regular representation) equivalent to a finite multiple of $\boldsymbol{\pi}$ (that is, a finite direct sum of subrepresentations equivalent to $\boldsymbol{\pi}$).

We shall now prove that the dimension of an isotypical component $\mathbb{L}^2(G, \boldsymbol{\pi})$ is exactly $(\dim \boldsymbol{\pi})^2$.

Recall the G-morphism c : End(V) $\to \mathbb{L}^2(G)$, $A \mapsto c_A := Tr(A\boldsymbol{\pi})$. The fact that $c_A \neq 0$ for $A \neq 0$ will be deduced from a computation of the quadratic norm of these coefficient functions. It is easier to start with the case of rank ≤ 1 linear mappings. We use the isomorphisms

$$\check{V} \otimes V \longrightarrow \text{End}(V) \qquad\qquad (\check{V} = \text{ dual of } V)$$

defined as follows: If $u \in \check{V}$ and $v \in V$, the operator corresponding to $u \otimes v$ is

$$u \otimes v : x \to u(x)v = \langle u, x \rangle v$$

The image of $u \otimes v$ consists of multiples of v, and $u \otimes v$ has rank 1 when u and v are non-zero (quite generally, decomposable tensors corresponding to operators of rank ≤ 1). The coefficient c_A with respect to the operator $A = u \otimes v$ coincides with the previously defined coefficient

$$c_v^u = \langle u, \boldsymbol{\pi}(x)v \rangle = c_{u \otimes v}(x)$$

Lemma 4.3. *Let $\boldsymbol{\pi}$ and $\boldsymbol{\sigma}$ be two representations of a compact group G and $A : V_{\boldsymbol{\pi}} \to V_{\boldsymbol{\sigma}}$ be a linear mapping. Then*

$$A^\natural = \int_G \boldsymbol{\sigma}(g) A \boldsymbol{\pi}(g)^{-1} dg$$

is a G-morphism from $V_{\boldsymbol{\pi}}$ to $V_{\boldsymbol{\sigma}}$, namely $A^\natural \in Hom_G(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}})$.

Proof. We have

$$A^\natural \boldsymbol{\pi}(s) = \int \boldsymbol{\sigma}(g) A \boldsymbol{\pi}(g)^{-1} \boldsymbol{\pi}(s) dg = \int \boldsymbol{\sigma}(g) A \boldsymbol{\pi}(s^{-1}g)^{-1} dg$$

and replacing g by sg (i.e. $s^{-1}g$ by g)

$$A^\natural \boldsymbol{\pi}(s) = \int \boldsymbol{\sigma}(sg) A \boldsymbol{\pi}(g)^{-1} dg = \boldsymbol{\sigma}(s) A^\natural$$

□

Thus the averaging operation (given by the Haar integral) of Lemma 4.3 leads to a **projector**

$$\natural : \mathrm{Hom}(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}}) \to \mathrm{Hom}_G(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}}) \quad , \quad A \to A^\natural$$

In particular when $\boldsymbol{\pi}$ and $\boldsymbol{\sigma}$ are *disjoint*, i.e $\mathrm{Hom}_G(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}}) = 0$, we must have $A^\natural = 0$. This is certainly the case when $\boldsymbol{\pi}$ and $\boldsymbol{\sigma}$ are non-equivalent irreducible representations (Schur's lemma). Another case of special interest is $\boldsymbol{\pi} = \boldsymbol{\sigma}$, finite dimensional and irreducible. Schur's lemma gives $\mathrm{Hom}_G(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}}) = \mathbb{C}$ and thus $A = \lambda_A \mathbf{id}$ is a scalar operator.

Proposition 4.4. *If $\boldsymbol{\pi}$ is a finite dimensional irreducible representation of the compact group G in V, the projector*

$$\mathit{End}(V) \to \mathit{End}_G(V) = \mathbb{C}\,\mathbf{id}, \quad A \to A^\natural = \lambda_A\,\mathbf{id},$$

is given explicitly by the following formula:

$$A^\natural = \int_G \boldsymbol{\pi}(g) A \boldsymbol{\pi}(g)^{-1} dg = \frac{Tr(A)}{\mathit{dim}V}\mathbf{id}_V$$

Proof. Since we know a priori that the operator $A^\natural$ is a scalar operator $\lambda_A\mathbf{id}$, we can determine the value of the scalar by taking traces in the defining equalities

$$\begin{aligned} \lambda_A Tr(\mathbf{id}_V) &= Tr\left(\int_G \boldsymbol{\pi}(g) A \boldsymbol{\pi}(g)^{-1} dg\right) = \int_G Tr\left(\boldsymbol{\pi}(g) A \boldsymbol{\pi}(g)^{-1}\right) dg \\ &= \int_G Tr(A) dg = Tr(A) \end{aligned}$$

□

Theorem 4.5 (Schur's orthogonality relations). *Let G be a compact group and $\boldsymbol{\pi}, \boldsymbol{\sigma}$ be a two finite dimensional irreducible representations of G. Assume that $\boldsymbol{\pi}$ and $\boldsymbol{\sigma}$ are unitary. Then*
(a) If $\boldsymbol{\pi}$ and $\boldsymbol{\sigma}$ are non-equivalent, $\mathbb{L}^2(G, \boldsymbol{\pi})$ and $\mathbb{L}^2(G, \sigma)$ are orthogonal in $\mathbb{L}^2(G)$.
(b) If $\boldsymbol{\pi}$ and $\boldsymbol{\sigma}$ are equivalent, $\mathbb{L}^2(G, \boldsymbol{\pi}) = \mathbb{L}^2(G, \boldsymbol{\sigma})$ and the inner product of the two coefficients of this space is given by

$$\langle c_v^u, c_y^x \rangle = \int_G \overline{\langle u, \boldsymbol{\pi}(g) v \rangle} \langle x, \boldsymbol{\pi}(g) y \rangle dg = \overline{\langle u, x \rangle} \langle v, y \rangle / \mathit{dim} V$$

(c) More generally in the case $\boldsymbol{\pi} = \boldsymbol{\sigma}$, the inner product of general coefficients is given by

$$\langle c_A, c_B \rangle = \int_G \overline{Tr(A\boldsymbol{\pi}(g))} Tr(B\boldsymbol{\pi}(g)) dg = Tr(A^* B) / \mathit{dim} V$$

Proof. (a) follows from Lemma 4.3 and (b) follows similarly from Proposition 4.4. It will be enough to show how (b) is derived. For this purpose, we consider the particular operators $\bar{v} \otimes y$ ($\bar{v} \in \bar{V}_{\boldsymbol{\pi}}$, $y \in V_{\boldsymbol{\pi}}$) and apply the result of the proposition

$$\int_G \boldsymbol{\pi}(g)(\bar{v} \otimes y)\boldsymbol{\pi}(g)^{-1} dg = \frac{Tr(\bar{v} \otimes y)}{\text{dim } V}\mathbf{id}_V = \frac{\langle v, y\rangle}{\text{dim}V}\mathbf{id}_V$$

Let us apply this operator to the vector u and take the inner product with the vector x

$$\langle x, \int_G \boldsymbol{\pi}(g)(\bar{v} \otimes y)\boldsymbol{\pi}(g)^{-1} u \; dg\rangle = \frac{\langle v, y\rangle}{\text{dim}V}\langle x, u\rangle = \frac{\overline{\langle u, x\rangle}\langle v, y\rangle}{\text{dim}V}$$

But we have

$$\begin{aligned}\boldsymbol{\pi}(g)(\bar{v}\otimes y)\boldsymbol{\pi}(g)^{-1}u &= \boldsymbol{\pi}(g)\langle v, \boldsymbol{\pi}(g^{-1})u\rangle y \\ &= \langle \boldsymbol{\pi}(g)v, u\rangle \boldsymbol{\pi}(g)y = \overline{\langle u, \boldsymbol{\pi}(g)v\rangle}\boldsymbol{\pi}(g)y\end{aligned}$$

hence

$$\langle x, \int_G \boldsymbol{\pi}(g)(\bar{v}\otimes y)\boldsymbol{\pi}(g)^{-1}u\ dg\rangle = \int_G \overline{\langle u, \boldsymbol{\pi}(g)v\rangle}\langle x, \boldsymbol{\pi}(g)y\rangle\ dg = \langle c_v^u, c_y^x\rangle$$

as expected. Finally (c) follows from (b) by linearity since the operators $\bar{v}\otimes y$ (of rank ≤ 1) generate End(V).

$\square$

In particular we see that if $0 \neq A \in \text{End}(V)$,

$$||c_A||^2 = Tr(A^*A)/\ \text{dim}\ V \neq 0$$

and the mapping $c : \text{End}(V) \to \mathbb{L}^2(G, \boldsymbol{\pi})$ is one-to-one and onto. The dimension of this isotypical component is thus (dim $V)^2$.

Corollary 4.6. *The Hilbert space* $\mathbb{L}^2(G)$ *is the Hilbert sum of all isotypical components*

$$\mathbb{L}^2(G) = \widehat{\bigoplus} \mathbb{L}^2(G, \boldsymbol{\pi})$$

The summation index $\boldsymbol{\pi}$ *runs over* **equivalence classes** *of finite dimensional irreducible representations of the compact group* G*.*

Proof. We have already seen that the isotypical subspaces $\mathbb{L}^2(G, \boldsymbol{\pi})$ are mutually orthogonal to each other. Thus our corollary will be proved if we show that the algebraic sum

$$A_G = \bigoplus \mathbb{L}^2(G, \boldsymbol{\pi}) \subseteq \mathcal{C}(G)$$

is *dense* in the Hilbert space $\mathbb{L}^2(G)$.

But A_G consists of coefficients of finite dimensional representations of G (we have proved that all finite dimensional representations are completely reducible), and the Peter-Weyl theorem 3.17 has shown that A_G is dense in $\mathcal{C}(G)$ for the uniform norm. A fortiori A_G will be dense in $\mathbb{L}^2(G)$ for the quadratic norm.

□

Corollary 4.7. *Any (continuous, topologically) irreducible representation of a compact group G in a Banach space is finite dimensional.*

Proof. Let $\boldsymbol{\sigma} : G \to Gl(E)$ be such a representation and let E' denote the (topological) dual of E, namely E' is the Banach space of continuous linear forms on E. By the Hahn-Banach theorem, for each $0 \neq x \in E$, there is a continuous linear form $x' \in E'$ with $x'(x) \neq 0$. For $u \in E'$ and $v \in E$ we can consider the corresponding coefficient of $\boldsymbol{\sigma}$

$$c_v^u \in \mathcal{C}(G) : g \to c_v^u(g) = \langle u, \boldsymbol{\sigma}(g)v \rangle$$

Letting v vary in E, we get a linear mapping

$$Q : E \to \mathcal{C}(G) \subseteq \mathbb{L}^2(G) \quad , \quad v \to c_v^u$$

Since G is compact and the mappings $g \to \boldsymbol{\sigma}(g)v$ are continuous (by the definition of *continuous* representations), the sets $\boldsymbol{\sigma}(G)v$ are compact, hence bounded in E (for each $v \in E$). By the uniform boundedness principle (Banach-Steinhaus theorem), the set $\boldsymbol{\sigma}(G)$ of operators in E is equicontinuous and bounded

$$\sup_{g\in G} ||\boldsymbol{\sigma}(g)|| = M < \infty$$

Hence

$$\begin{aligned} ||Qv|| &= \sup_{g\in G} |c_v^u(g)| = \sup_{g\in G} |\langle u, \boldsymbol{\sigma}(g)v\rangle| \\ &\leq ||u||_{E'} \sup_{g\in G} ||\boldsymbol{\sigma}(g)v||_E \leq M||u||_{E'}||v||_E \end{aligned}$$

This proves that Q is continuous from E into $\mathcal{C}(G)$ (equipped with the uniform norm); it's kernel is a proper and closed subspace $F \neq E$ if $u \neq 0$ (in this case $u(v) \neq 0$ for some $v \in E$ and thus $c_v^u(e) = \langle u, v \rangle = u(v) \neq 0$).

Take $v \in E$ with $Q(v) \neq 0$, and apply the orthogonal Hilbert sum decomposition of the preceding corollary to $Q(v)$.

$$\sum P_{\boldsymbol{\pi}}(Qv) = Qv \neq 0$$

with

$$P_{\boldsymbol{\pi}} = \text{ orthogonal projector from } \mathbb{L}^2(G) \text{ onto } \mathbb{L}^2(G, \boldsymbol{\pi})$$

This implies that there is a $\boldsymbol{\pi}$ (finite dimensional irreducible representation of G) with $P_{\boldsymbol{\pi}} Qv \neq 0$. For this $\boldsymbol{\pi}$, we consider the composite

$$E \xrightarrow{Q} \mathbb{L}^2(G) \xrightarrow{P_{\boldsymbol{\pi}}} \mathbb{L}^2(G, \boldsymbol{\pi})$$

and its kernel which is a proper closed subspace $F_{\boldsymbol{\pi}} \neq E$. But Q is a G-morphism (intertwining $\boldsymbol{\sigma}$ and the right regular representation)

$$c^u_{\boldsymbol{\sigma}(x)v}(g) = \langle u, \boldsymbol{\sigma}(g)\boldsymbol{\sigma}(x)v\rangle = c^u_v(gx)$$

which implies that $Q(\boldsymbol{\sigma}(x)v) = \rho(x)Q(v)$. Since $P_{\boldsymbol{\pi}}$ is also a G-morphism, the kernel $F_{\boldsymbol{\pi}}$ of the composite $P_{\boldsymbol{\pi}}Q$ must be an invariant subspace of E. However E is irreducible by assumption so that $F_{\boldsymbol{\pi}} = \{0\}$, and the composite $P_{\boldsymbol{\pi}}Q$ is one-to-one (into):

$$\text{dim } E \leq \text{dim } \mathbb{L}^2(G, \boldsymbol{\pi}) = (\text{dim } V)^2 \qquad \square$$

Definition 4.8. The dual $\widehat{G}$ of a compact group G is the **set** of equivalence classes of irreducible representations of G.

Let $\boldsymbol{\pi}$ be an irreducible representation of the compact group G and, $\varpi = [\boldsymbol{\pi}]$ its equivalence class ($\varpi \in \widehat{G}$). We say that $\boldsymbol{\pi}$ is a *model* of ϖ in this case i.e. when $\boldsymbol{\pi} \in \varpi$. The dimension of ϖ is defined as dim $(\boldsymbol{\pi}) =$ dim $(V_{\boldsymbol{\pi}})$ independently from the model chosen. Similarly the isotypical component of ϖ (in the right regular representation) is defined as $\mathbb{L}^2(G, \varpi) = \mathbb{L}^2(G, \boldsymbol{\pi})$ independently from the model $\boldsymbol{\pi}$ chosen for ϖ. By the finiteness of the dimension of the irreducible representations of the compact group G namely Corollaries 4.6 and 4.7

$$\mathbb{L}^2(G) = \widehat{\bigoplus}_{\varpi \in \hat{G}} \mathbb{L}^2(G, \varpi)$$

Instead of $\boldsymbol{\pi} \in \varpi \in \widehat{G}$ we shall write more simply $\boldsymbol{\pi} \in \widehat{G}$.

Proposition 4.9. *Let G be a compact group. Then the following properties are equivalent*
(i) The dual $\widehat{G}$ is countable.
(ii) $\mathbb{L}^2(G)$ is separable.
(iii) G is metrizable.

Proof. The equivalence between (i) and (ii) is obvious since $\mathbb{L}^2(G)$ is the Hilbert sum of the finite dimensional isotypical components $\mathbb{L}^2(G, \varpi)$ over the index set $\widehat{G}$. Moreover G can always be embedded in a product

$$\prod_{\widehat{G}} U(\boldsymbol{\pi}) \qquad \text{with } U(\boldsymbol{\pi}) \cong U_{\mathsf{dim}\ (\boldsymbol{\pi})}(\mathbb{C}) \qquad \text{(metrizable group)}$$

Since any countable product of metrizable topological spaces is metrizable, we see that (i)$\Rightarrow$(iii). Finally, the implication (iii)$\Rightarrow$(ii) is a classical application of the Stone-Weirstrass theorem.

4.1 Exercises

Exercise 4.10. Let V be a finite dimensional vector space and $\check{V}$ be its dual and for $u \in \check{V}$, $u \in V$ denote by $u \otimes v$ the operator $x \to u(x)v$ as defined in the beginning of this Section. Show
(a) $Tr(u \otimes v) = \langle u, v \rangle = u(v)$ (intrinsic definition of Tr),
(b) $(u \otimes v) \cdot (x \otimes y) = \langle u, y \rangle x \otimes v$,
(c) ${}^t Au \otimes Bv = B \cdot (u \otimes v) \cdot A$ for $A, B \in \mathrm{End}(V)$.
Moreover, identifying the dual of $\check{V} \otimes V$ with the dual of $V \otimes \check{V}$ in the obvious canonical way show
(d) ${}^t(u \otimes v) = v \otimes u$
Assume now that V is now a representation space for a group G. Using (a) and (c) prove
(e) $c_{u \otimes v} = c_v^u$ (i.e. $Tr(\boldsymbol{\pi}(x) \cdot u \otimes v) = \langle u, \boldsymbol{\pi}(x) v \rangle$).

Exercise 4.11. Let G be a compact group and $(\boldsymbol{\pi}, V)$ be a finite dimensional representation of G. We denote by V^G the subspace of invariants of V:

$$V^G = \{v \in V : \boldsymbol{\pi}(g)v = v \quad \forall g \in G\}$$

(a) Check that the operator $P = \int_G \boldsymbol{\pi}(g)dg$ is a projector from V onto V^G. If $\boldsymbol{\pi}$ is unitary, P is the orthogonal projector on this subspace.
(b) For two finite dimensional representations $\boldsymbol{\pi}$ and $\boldsymbol{\sigma}$ of G, consider $\mathrm{Hom}(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}})$ as a representation space of G via the action

$$g \cdot A = \boldsymbol{\pi}(g) \cdot A\boldsymbol{\sigma}(g)^{-1}$$

Observe that

$$\mathrm{Hom}(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}})^G = \mathrm{Hom}_G(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}})$$

and deduce a proof of the Lemma 4.3 from this observation.

(c) Using the G-isomorphism

$$\check{V}_{\boldsymbol{\pi}} \otimes V_{\boldsymbol{\sigma}} \longrightarrow \mathrm{Hom}(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}})$$

with $\check{V}_{\boldsymbol{\pi}} \otimes V_{\boldsymbol{\sigma}}$ being equipped with the representation $\check{\boldsymbol{\pi}} \otimes \boldsymbol{\sigma}$, conclude that the projector $\natural : \check{V}_{\boldsymbol{\pi}} \otimes V_{\boldsymbol{\sigma}} \to (\check{V}_{\boldsymbol{\pi}} \otimes V_{\boldsymbol{\sigma}})^G$ is given by $\int_G (\check{\boldsymbol{\pi}} \otimes \boldsymbol{\sigma})(g) dg$

5 Convolution, Plancherel formula and Fourier Inversion

Definition 5.1 (Convolution). On a compact group G, the convolution of two continuous functions f and g is defined as

$$f * g(x) = \int_G f(y)g(y^{-1}x)dy$$

Defining $f^*(x) = \overline{f(x^{-1})}$, we can also write

$$\begin{aligned} f * g(x) &= \int_G f(xy)g(y^{-1})dy = \int_G f(xy^{-1})g(y)dy \\ &= \int_G \overline{f^*(yx^{-1})}g(y)dy = \langle r(x^{-1})f^*, g\rangle \end{aligned}$$

The Cauchy-Schwartz inequality gives:

$$|f * g(x)| \leq ||f^*||_2||g||_2 = ||f||_2||g||_2$$

whence

$$||f * g||_\infty = \sup_{x \in G} |f * g(x)| \leq ||f||_2||g||_2$$

Consequently the convolution product can be extended by continuity from $\mathcal{C}(G)$ to $\mathbb{L}^2(G)$ and by definition

$$* : \mathbb{L}^2(G) \times \mathbb{L}^2(G) \to \mathcal{C}(G), \quad , \quad (f, g) \to f * g$$

is a continuous bilinear mapping still satisfying the above inequality. On the other hand the preceding formulas show:

$$\langle f, g\rangle = f^* * g(e) \quad , \quad ||f||_2^2 = f^* * f(e)$$

The convolution product for functions in $\mathbb{L}^1(G)$ can be defined by the same integral formula, but this will not converge for every $x \in G$ and the result will no longer be continuous in general. To see what happens, take $f, g \in \mathbb{L}^1(G)$. By Fubini's theorem,

$$\begin{aligned}\int |f * g(x)| dx &\leq \int dx \int dy |f(y) g(y^{-1}x)| \\ &= \int dy |f(y)| \int dx |g(y^{-1}x)| \\ &= ||g||_1 \int |f(y)| dy = ||f||_1 ||g||_1 < \infty\end{aligned}$$

These inequalities prove that $\int f(y)g(y^{-1}x)dy$ converges absolutely for almost all $x \in G$ and the result $f * g \in \mathbb{L}^1(G)$ satisfies:

$$||f * g||_1 \leq ||f||_1 ||g||_1$$

This convolution product is associative and $\mathbb{L}^1(G)$ is an *algebra for convolution*. This algebra has no unit element in general (more precisely, it has no unit when G is not discrete i.e. G is not finite). We shall see that this algebra $\mathbb{L}^1(G)$ is commutative exactly when G is commutative.

5.1 Integration of representations

Assume that $\boldsymbol{\pi}$ is a *unitary* representation of the compact group G in a Hilbert space H. We can "extend" $\boldsymbol{\pi}$ to a representation of $\mathbb{L}^1(G)$ by the formula

$$\boldsymbol{\pi}^1(f) = \int_G f(x)\boldsymbol{\pi}(x)dx \qquad\qquad (f \in \mathbb{L}^1(G))$$

These integrals converge absolutely in norm: $||\boldsymbol{\pi}(x)|| = 1$ implies

$$\int ||f(x)\boldsymbol{\pi}(x)||dx = \int |f(x)|dx = ||f||_1$$

Thus,

$$||\boldsymbol{\pi}^1(f)|| \leq ||f||_1 \qquad\qquad (f \in \mathbb{L}^1(G))$$

Although G is not really embedded in $\mathbb{L}^1(G)$ (when G is infinite), we consider $\boldsymbol{\pi}^1$ as an extension of $\boldsymbol{\pi}$. Later on we shall even drop the index 1, writing $\boldsymbol{\pi}$ instead of $\boldsymbol{\pi}^1$. The association

$$\boldsymbol{\pi} : G \to Gl(H) \quad \rightsquigarrow \quad \boldsymbol{\pi}^1 : \mathbb{L}^1(G) \to \mathsf{End}(H)$$

can even be made when $\boldsymbol{\pi}$ is a representation in a Banach space since $\boldsymbol{\pi}(G)$ is always a bounded set of End(H): being weakly compact, it is uniformly bounded.

5.2 Comparison of several norms

Let $A \in \mathrm{End}(V)$ be any endomorphism in V. Take any orthonormal basis (e_i) of V and assume that A is represented by the matrix (a^i_j) in the basis (e_i). Obviously

$$||A||^2_{HS} = \sum_{i,j} |a^i_j|^2$$

defines a norm on End(V) called the Hilbert-Schmidt norm (a priori this norm depends on the choice of the orthonormal basis (e_i)).

If B is another endomorphism, represented by the matrix (b^i_j) (in the same basis), a small computation shows that

$$Tr(A^*B) = \sum_{ij} \bar{a}^i_j b^i_j$$

This shows that the Hilbert-Schmidt norm is derived from the Hilbert-Schmidt inner product

$$\langle A, B \rangle_{HS} = \sum_{i,j} \bar{a}^i_j b^i_j = Tr(A^*B)$$

on End(V), and is in particular independent from the choice of the orthonormal basis (e_i) of V.

Return to a compact group G and a unitary irreducible representation $\boldsymbol{\pi} \in \widehat{G}$ in some finite dimensional Hilbert space $V = V_{\boldsymbol{\pi}}$. The spaces

$$\check{V} \otimes V \quad , \quad \text{End}(V) \quad , \quad \mathbb{L}^2(G, \boldsymbol{\pi})$$

are G-isomorphic. We shall give explicit isomorphisms between them keeping track of the various norms involved.

We have introduced the coefficients

$$c_v^u(x) = \langle u, \boldsymbol{\pi}(x) v \rangle \qquad\qquad (u \in \check{V}, v \in V)$$

and the more general coefficients (linear combinations of the preceding ones)

$$c_A(x) = Tr(A\boldsymbol{\pi}(x)) \qquad\qquad (A \in \mathsf{End}(V))$$

with the relation

$$c_A = c_v^u \qquad \text{for} \qquad A = u \otimes v$$

For fixed $u \in V$, the linear mapping

$$c^u : V \to \mathbb{L}^2(G, \boldsymbol{\pi}) \quad , \quad v \to c^u_v$$

is a G-morphism form $\boldsymbol{\pi}$ to r, the right regular representation.

Similarly if $v \in V$ is fixed,

$$\begin{aligned} l(s)c_v^u(x) &= c_v^u(s^{-1}x) = \langle u, \boldsymbol{\pi}(s^{-1})\boldsymbol{\pi}(x)v\rangle \\ &= \langle {}^t\boldsymbol{\pi}(s^{-1})u, \boldsymbol{\pi}(x)v\rangle = c_v^{\check{\boldsymbol{\pi}}(s)u}(x) \end{aligned}$$

shows that the linear mapping

$$c_v : \check{V} \to \mathbb{L}^2(G, \boldsymbol{\pi}) \quad , \quad u \to c_v^u$$

is a G-morphism from $\check{\boldsymbol{\pi}}$ to l (the left regular representation). Summing up,

$$c : \check{V} \otimes V \to \mathbb{L}^2(G, \boldsymbol{\pi}) \quad , \quad u \otimes v \to c_v^u$$

is a $\check{\boldsymbol{\pi}} \otimes \boldsymbol{\pi}$ to $l \times r$ (biregular representation) $G \times G$-morphism.

Note that

$$
\begin{aligned}
[l \times r(s,t)]c_A(x) &= c_A(s^{-1}xt) = Tr(A\boldsymbol{\pi}(s^{-1})\boldsymbol{\pi}(x)\boldsymbol{\pi}(t)) \\
&= Tr(\boldsymbol{\pi}(t)A\boldsymbol{\pi}(s)^{-1}\boldsymbol{\pi}(x)) = c_{\boldsymbol{\pi}(t)A\boldsymbol{\pi}(s)^{-1}}(x)
\end{aligned}
$$

So, the corresponding action of $G \times G$ on End(V) is defined by

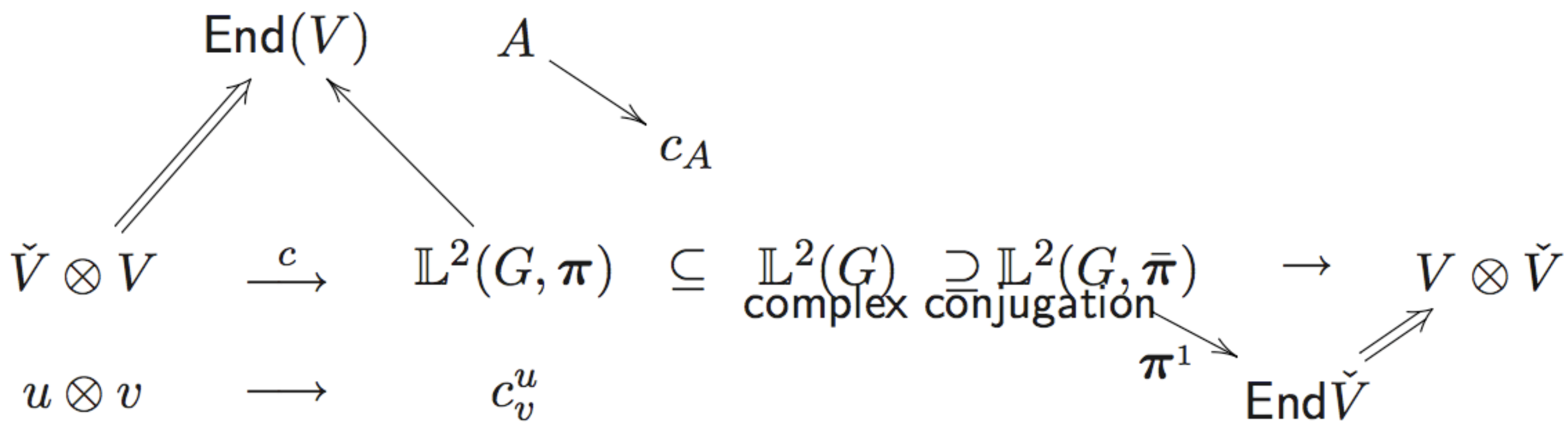
$$(s,t) \cdot A = \boldsymbol{\pi}(t)A\boldsymbol{\pi}(s)^{-1}$$

In the following diagram of G-morphisms, $\check{V} \otimes V$ is equipped with the inner product

$$\langle u \otimes v, x \otimes y \rangle = \overline{\langle u, x \rangle} \langle v, y \rangle$$

(We use the Riesz isomorphism between $\check{V}$ and $\bar{V}$). This inner product corresponds to the Hilbert-Schmidt norm

$$\langle u \otimes v, x \otimes y \rangle = Tr((u \otimes v)^* x \otimes y)$$


End(V)
A
c_A
V̌ ⊗ V
c
L²(G, π)
⊆
L²(G)
⊇ L²(G, π̄)
→
V ⊗ V̌
complex conjugation
π¹
End V̌
u ⊗ v
c_v^u

Schur's orthogonality relations (Theorem 4.5) say

$$\langle c_v^u, c_y^x \rangle = \frac{1}{\text{dim } \boldsymbol{\pi}} \overline{\langle u, x \rangle} \langle v, y \rangle \qquad (\text{dim } \boldsymbol{\pi} = \text{dim } V),$$

and hence

$$c : u \otimes v \to c_v^u \text{ is } \sqrt{\text{dim } \boldsymbol{\pi}}^{-1} \times \qquad \text{an isometry map}$$

The inverse of c is nearly the extension $\boldsymbol{\pi}^1$ of $\boldsymbol{\pi}$. Let us compute $\boldsymbol{\pi}^1(c_v^u)$. For this purpose, we apply this operator to a vector y and compute the inner product of the result with a vector x

$$\langle x, \boldsymbol{\pi}^1(c_v^u)y\rangle = \int c_v^u(s)\langle x, \boldsymbol{\pi}(s)y\rangle ds = \langle \bar{c}_v^u, c_y^x\rangle$$

This quantity will vanish when $\bar{\boldsymbol{\pi}}$ is not equivalent to $\boldsymbol{\pi}$.

However, Schur's relations give

$$\begin{aligned}\langle x, \boldsymbol{\pi}^1(\bar{c}_v^u)y\rangle &= \langle c_v^u, c_y^x\rangle = (\text{dim } \boldsymbol{\pi})^{-1}\overline{\langle u, x\rangle}\langle v, y\rangle \\ &= (\text{dim } \boldsymbol{\pi})^{-1}\langle x, \langle v, y\rangle u\rangle = \langle x, (\text{dim } \boldsymbol{\pi})^{-1} v \otimes u(y)\rangle\end{aligned}$$

This implies

$$\boldsymbol{\pi}^1(\bar{c}_v^u) = (\text{dim } \boldsymbol{\pi})^{-1} v \otimes u$$

and by linearity

$$\boldsymbol{\pi}^1(\bar{c}_A) = (\text{dim } \boldsymbol{\pi})^{-1} A^*$$

Since the $\bar{c}_A$ are the coefficients of $\bar{\boldsymbol{\pi}}$, we see that on $\mathbb{L}^2(G, \bar{\boldsymbol{\pi}})$, $f \to \boldsymbol{\pi}^1(f)$ is of the form

$$\boldsymbol{\pi}^1\Big|_{\mathbb{L}^2(G,\bar{\boldsymbol{\pi}})} = \sqrt{\text{dim } \boldsymbol{\pi}}^{-1} \times \text{ an isometry map}$$

The composite:

$$\begin{array}{ccccccc} \text{End}(V) & \to & \mathbb{L}^2(G, \boldsymbol{\pi}) & \xrightarrow{conj} & \mathbb{L}^2(G, \bar{\boldsymbol{\pi}}) & \to & \text{End}(\check{V}) \\ A & \to & c_A & & \bar{c}_A = f & \to & \boldsymbol{\pi}^1(f) \end{array}$$

can be identified with

$$(\text{dim } \boldsymbol{\pi})^{-1} \cdot (A \to A^*) = (\text{dim } \boldsymbol{\pi})^{-1} \times \text{ an isometry map}$$

NOTE: From now on, we write $\boldsymbol{\pi}^1(f)$ as simply $\boldsymbol{\pi}(f)$.

Theorem 5.2 (Plancherel theorem). *Let G be a compact group. For $\boldsymbol{\pi} \in \hat{G}$ denote by $P_{\boldsymbol{\pi}} : \mathbb{L}^2(G) \to \mathbb{L}^2(G, \boldsymbol{\pi})$, the orthogonal projector on the isotypical component of $\boldsymbol{\pi}$, and for $f \in \mathbb{L}^2(G)$, let $f_{\boldsymbol{\pi}} = P_{\boldsymbol{\pi}}(f)$ so that the series $\sum_{\hat{G}} f_{\boldsymbol{\pi}}$ converges to f in $\mathbb{L}^2(G)$. Then*

(a) $f_{\boldsymbol{\pi}}(x) =$ dim $\boldsymbol{\pi} \cdot Tr(\boldsymbol{\pi}(\check{f})\boldsymbol{\pi}(x))$. Here $\check{f}(x) = f(x^{-1})$

(b) $||f_{\boldsymbol{\pi}}||_2^2 =$ dim $\boldsymbol{\pi} \cdot ||\boldsymbol{\pi}(\check{f})||_{HS}^2$ (Hilbert-Schmidt norm on RHS).

(c) $||f||_2^2 = \sum_{\boldsymbol{\pi} \in \hat{G}}$ dim $\boldsymbol{\pi} \cdot ||\boldsymbol{\pi}(f)||_2^2$ (Parseval equality)

Proof. The orthogonal projection $f_{\boldsymbol{\pi}}$ of f in $\mathbb{L}^2(G, \boldsymbol{\pi})$, is the element c_A of this space having the same inner product with all elements c_B of $\mathbb{L}^2(G, \boldsymbol{\pi})$. Let us determine A as a function of f. We must have

$$\langle c_B, f \rangle = \langle c_B, f_{\boldsymbol{\pi}} \rangle = \langle c_B, c_A \rangle = (\text{dim } \boldsymbol{\pi})^{-1} Tr(B^* A)$$

But

$$\begin{aligned} \langle c_B, f \rangle &= \int \overline{c_B(x)} f(x) dx = \int \overline{Tr(B\boldsymbol{\pi}(x))} f(x) dx \\ &= \int Tr(\boldsymbol{\pi}(x^{-1}) B^*) f(x) dx = Tr(B^* \int \boldsymbol{\pi}(x^{-1}) f(x) dx) \\ &= Tr(B^* \boldsymbol{\pi}(\check{f})) \end{aligned}$$

Comparing the two results obtained for all B, we indeed find

$$A = (\text{dim } \boldsymbol{\pi}) \cdot \boldsymbol{\pi}(\check{f})$$

This gives

$$f_{\boldsymbol{\pi}}(x) = c_A(x) = Tr(A\boldsymbol{\pi}(x)) = (\text{dim } \boldsymbol{\pi}) \cdot Tr(\boldsymbol{\pi}(\check{f})\boldsymbol{\pi}(x))$$

as asserted in part (a). Moreover Schur's relations show that

$$||f_{\boldsymbol{\pi}}||_2^2 = ||c_A||_2^2 = (\text{dim } \boldsymbol{\pi})^{-1}||A||_{HS}^2 = (\text{dim } \boldsymbol{\pi}) \cdot ||\boldsymbol{\pi}(\check{f})||_{HS}^2$$

Thus (b) is proved and (c) follows from the observation that f and $\check{f}$ have the same $\mathbb{L}^2(G)$: we can interchange f and $\check{f}$. Also observe that the dimensions of $\boldsymbol{\pi}$ and $\check{\boldsymbol{\pi}}$ are the same.

□

5.3 Exercises

Exercise 5.3. Let G be a compact group, $f, g \in \mathbb{L}^1(G)$. For any representation $\boldsymbol{\sigma}$ of G (in a Banach space), prove

$$\boldsymbol{\sigma}(f * g) = \boldsymbol{\sigma}(f) \cdot \boldsymbol{\sigma}(g)$$

If $\boldsymbol{\sigma}$ is unitary, prove also $\boldsymbol{\sigma}(f) = \boldsymbol{\sigma}(f^*)$. (recall that $f^*(x) = \overline{f(x^{-1})}$)

Exercise 5.4. Show that the "extensions" of the regular representations of a compact group G are given by

$$l(f)(\varphi) = f * \varphi \quad , \quad r(g)(\varphi) = \varphi * \check{g}$$

where $f, g \in \mathbb{L}^1(G)$ and $\varphi \in \mathbb{L}^2(G)$ (recall that $\check{g}(x) = g(x^{-1})$).
Conclude from this that

$$||f * \varphi||_2 \leq ||f||_1 ||\varphi||_2$$

Moreover using exercise 5.3 deduce the associativity

$$(f * g) * \varphi = f * (g * \varphi)$$

Here $f, g \in \mathbb{L}^1(G)$ and $\varphi \in \mathbb{L}^2(G)$ or all three functions in $\mathcal{C}(G)$. Also check that for any representation $\boldsymbol{\sigma}$ of G

$$\boldsymbol{\sigma}(l(x)f) = \boldsymbol{\sigma}(x)\boldsymbol{\sigma}(f) \quad , \quad \boldsymbol{\sigma}(r(x)f) = \boldsymbol{\sigma}(f)\boldsymbol{\sigma}(x^{-1})$$

Exercise 5.5. Let G be a compact group and denote by $\mathbb{L}^1_{inv} = \mathbb{L}^1_{inv}(G)$ the closure of $\mathbb{L}^1(G)$ of the subspace of continuous functions f satisfying $f(xy) = f(yx)$ (or equivalently $f(y^{-1}xy) = f(x)$) for all $x, y \in G$. Show that $\mathbb{L}^1_{inv}$ is contained in the center of $\mathbb{L}^1(G)$ (as convolution algebra: prove that $f * g = g * f$ for $f, g \in \mathbb{L}^1_{inv}$). For any irreducible representation $\boldsymbol{\pi} : G \to Gl(V)$ prove that

$$\boldsymbol{\pi}(f) = (\text{dim } \boldsymbol{\pi})^{-1}\langle \chi, f\rangle 1_V \qquad (f \in \mathbb{L}^1_{inv})$$

where $\chi(g) = Tr(\boldsymbol{\pi}(g))$ and $\langle \chi, f\rangle = \int \chi(g)f(g)dg$.

Hint: Use Schur's lemma to prove that $\boldsymbol{\pi}(f)$ is a scalar operator and then take traces to determine the value of the constant in this multiple of the identity.

Exercise 5.6. Let $\boldsymbol{\sigma} : G \to Gl(V)$ be a unitary representation of a compact group G. Check that $\boldsymbol{\sigma}(1) = P$ (here 1 is the constant function $\mathbf{1}$ in $\mathbb{L}^1(G)$) is the orthogonal projector $V \to V^G$ on the subspace of G-invariants of V. (**Hint:** show that $1 * 1 = 1$ and $1^* = 1$ in $\mathbb{L}^1(G)$; more generally $1 * f = f * 1$ is the constant function $\int f(x)dx$.)

Exercise 5.7. Show that the "extended" left regular representation

$$l = l^1 : \mathbb{L}^1(G) \to \mathrm{End}(\mathbb{L}^2(G))$$

has trivial kernel $\{0\}$. (**Hint:** Let $0 \neq f \in \mathbb{L}^1(G)$ and construct a sequence $(g_n) \subseteq \mathcal{C}(G)$ with $g_n \geq 0, \int g_n(x)dx = 1$ and $l(f)(g_n) = f * g_n \to f \neq 0$). Conclude that if $0 \neq f \in \mathbb{L}^1(G)$, there exists a $\boldsymbol{\pi} \in \hat{G}$ such that $\boldsymbol{\pi}(f) \neq 0$. (continued on the next page)

Finally prove that

$$\mathbb{L}^1(G) \quad \text{commutative} \Longleftrightarrow G \quad \text{commutative}$$

Exercise 5.8. Let G be a compact group, $\boldsymbol{\pi} \in \hat{G}$ and consider the adjoint representation of G in $\mathrm{End}(V)$ ($V = V_{\boldsymbol{\pi}}$) defined by the following composition

$$\begin{array}{llllll} \text{Ad:} & G & \longrightarrow & G \times G & \longrightarrow & \mathrm{End}(V) \\ & s & \longrightarrow & (s,s) & \longrightarrow & (A \to \boldsymbol{\pi}(s)A\boldsymbol{\pi}(s)^{-1}) \\ & & & (s,t) & \longrightarrow & (A \to \boldsymbol{\pi}(t)A\boldsymbol{\pi}(s)^{-1}) \end{array}$$

Prove that the multiplicity of the identity representation in this adjoint representation is 1. (This identity representation acts on the subspace of scalar operators: Schur's lemma).

Exercise 5.9. The decomposition of the biregular representation of a compact group G in $\mathbb{L}^2(G)$ gives the decomposition of the left (resp. right) regular representation simply by the composition with

$$\begin{array}{rccc} i_1 : & G & \to & G \times G \\ & s & \to & (s,e) \end{array} \qquad \text{(resp. } \begin{array}{rccc} i_2 : & G & \to & G \times G \\ & s & \to & (e,s) \end{array} \text{)}$$

Conclude that

$$\begin{array}{ccccccc} l & \cong & \hat{\bigoplus} \check{\boldsymbol{\pi}} \otimes 1 & \cong & \hat{\bigoplus} \text{dim } \boldsymbol{\pi} \cdot \check{\boldsymbol{\pi}} & = & \text{dim } \boldsymbol{\pi} \cdot \boldsymbol{\pi} \\ r & \cong & \hat{\bigoplus} 1 \otimes \boldsymbol{\pi} & \cong & \hat{\bigoplus} \text{dim } \boldsymbol{\pi} \cdot \boldsymbol{\pi} & & \end{array}$$

6 Characters and Group algebras

Let $(\boldsymbol{\pi}, V)$ be a finite dimensional representation of a compact group G. The character $\chi = \chi_{\boldsymbol{\pi}}$ of $\boldsymbol{\pi}$ is the (complex valued) continuous function on G defined by

$$\chi(x) = Tr(\boldsymbol{\pi}(x))$$

Note that this is the function c_A for $A = \mathbf{id}_V \in \mathrm{End}(V)$.

When dim $(V) = 1$, χ and $\boldsymbol{\pi}$ can be identified. In this case χ is a homomorphism.

Quite generally, since the trace satisfies the identity $Tr(AB) = Tr(BA)$, we see that the characters of two *equivalent* representations are equal.

Moreover, characters satisfy

$$\chi(xy) = \chi(yx) \text{ as } \chi(y^{-1}xy) = \chi(x) \qquad (x, y \in G)$$

Thus characters are *invariant* functions

$$\chi \in C_{inv} = \{f \in \mathcal{C}(G) : f(y^{-1}xy) = f(x), \quad x \text{ and } y \in G\}$$

Observe that

$$\begin{aligned} \mathbb{L}^1_{inv} &= \text{closure of } \mathcal{C}(G) \text{ in } \mathbb{L}^1(G) \\ \mathbb{L}^2_{inv} &= \text{closure of } \mathcal{C}(G) \text{ in } \mathbb{L}^2(G) \end{aligned}$$

Invariant functions are also called *central* functions, they belong to the center $\mathbb{L}^1(G)$ with respect to convolution.

Still quite generally, the character of the contragredient $\check{\boldsymbol{\pi}}$ of $\boldsymbol{\pi}$ is given by

$$\chi_{\check{\boldsymbol{\pi}}}(x) = Tr(\check{\boldsymbol{\pi}}(x)) = Tr({}^t\boldsymbol{\pi}(x^{-1})) = Tr(\boldsymbol{\pi}(x^{-1})) = \chi(x^{-1})$$

hence $\chi_{\check{\boldsymbol{\pi}}} = \check{\chi}$

When $\boldsymbol{\pi}$ is unitary, $\boldsymbol{\pi}(x^{-1}) = \boldsymbol{\pi}(x)^*$ ($\check{\boldsymbol{\pi}}$ is equivalent to $\bar{\boldsymbol{\pi}}$) and $\check{\chi}$ is the complex conjugate of χ.

One can also check without difficulty that for two finite dimensional representations $\boldsymbol{\pi}$ and $\boldsymbol{\sigma}$ of G

$$\chi_{\boldsymbol{\pi}\oplus\boldsymbol{\sigma}} = \chi_{\boldsymbol{\pi}} + \chi_{\boldsymbol{\sigma}} \quad , \quad \chi_{\boldsymbol{\pi}\otimes\boldsymbol{\sigma}} = \chi_{\boldsymbol{\pi}} \cdot \chi_{\boldsymbol{\sigma}}$$

When $\boldsymbol{\pi}$ is irreducible, Schur's lemma shows that the elements z in the center Z of G are mapped to scalar operators by $\boldsymbol{\pi}$: $\boldsymbol{\pi}(z) = \lambda_z \mathbf{id}_V$ so that $\chi(z) = \lambda_z$dim (V). Thus the restriction of (dim $V)^{-1}\chi$ to the center Z is a homomorphism

$$\lambda : Z \longrightarrow \mathbb{C}^{\times}$$

This is the **central character** of $\boldsymbol{\pi}$: it is independent of the special model chosen in the equivalence class of $\boldsymbol{\pi}$.

In particular if $\lambda(z)$ is *not contained* in $\{\pm 1\}$, $\boldsymbol{\pi}$ and $\bar{\boldsymbol{\pi}}$ are not equivalent: their central characters are different.

Also observe that $\chi(e) =$ dim (V) (= dim $\boldsymbol{\pi}$)

Proposition 6.1. *Any continuous central function* $f \in \mathcal{C}_{inv}$ *on a compact group* G *is a uniform limit of linear combinations of characters of irreducible representations of* G.

Proof. Let $\epsilon > 0$. By the Peter-Weyl theorem 3.17, we know that there is a finite dimensional representation of $(\boldsymbol{\sigma}, V)$ and a $A \in \text{End}(V)$ with

$$|f(x) - Tr(A\boldsymbol{\sigma}(x))| < \epsilon \qquad\qquad (x \in G)$$

In this representation, replace x by one of its conjugates yxy^{-1}:

$$|f(x) - Tr(A\boldsymbol{\sigma}(yxy^{-1}))| = |f(x) - Tr(\boldsymbol{\sigma}(y^{-1})A\boldsymbol{\sigma}(y)\boldsymbol{\sigma}(x))| < \epsilon$$

Integrating over y, we have

$$|f(x) - Tr(B\boldsymbol{\sigma}(x))| < \epsilon \text{ where } B = \int \boldsymbol{\sigma}(y^{-1})A\boldsymbol{\sigma}(y)dy$$

By the invariance of the Haar measure, the operator B commutes with all the operators $\boldsymbol{\sigma}(x)$.

Hence, if we decompose $\boldsymbol{\sigma}$ into isotypical components

$$\boldsymbol{\sigma} \cong \bigoplus_{\boldsymbol{\pi}} n_{\boldsymbol{\pi}} \boldsymbol{\pi} \cong \bigoplus_{\boldsymbol{\pi}} \boldsymbol{\pi} \otimes 1_{n_{\boldsymbol{\pi}}}$$

the operator B will have the form

$$B = \bigoplus \mathbf{id}_{\text{dim } \boldsymbol{\pi}} \otimes B_{\boldsymbol{\pi}}$$

and

$$\begin{aligned} B\boldsymbol{\sigma}(x) &= \boldsymbol{\sigma}(x)B \cong \bigoplus \boldsymbol{\pi} \otimes B_{\boldsymbol{\pi}} \\ Tr(B\boldsymbol{\sigma}(x)) &= \sum a_{\boldsymbol{\pi}} \chi_{\boldsymbol{\pi}}(x) \qquad (a_{\boldsymbol{\pi}} = Tr(B_{\boldsymbol{\pi}})) \end{aligned}$$

This shows that

$$|f(x) - \sum_{\text{finite}} a_{\boldsymbol{\pi}} \chi_{\boldsymbol{\pi}}(x)| < \epsilon$$

$\square$

Theorem 6.2. *Let $\boldsymbol{\pi}$ and $\boldsymbol{\sigma}$ be two finite dimensional representations of a compact group G with respective characters $\chi_{\boldsymbol{\pi}}$ and $\chi_{\boldsymbol{\sigma}}$. Then*

$$\langle \chi_{\boldsymbol{\pi}}, \chi_{\boldsymbol{\sigma}} \rangle = \textit{dim Hom}_G(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}})$$

Proof. By Lemma 4.3, we know that the integral

$$\int \check{\boldsymbol{\pi}}(x) \otimes \boldsymbol{\sigma}(x) dx$$

is an expression for the projector

$$\begin{array}{rcl} \natural : \check{V}_{\boldsymbol{\pi}} \otimes V_{\boldsymbol{\sigma}} & \longrightarrow & (\check{V}_{\boldsymbol{\pi}} \otimes V_{\boldsymbol{\sigma}})^G \\ \mathsf{Hom}(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}}) & \longrightarrow & \mathsf{Hom}_G(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}}) \end{array}$$

The dimension of the image space is the trace of this projector. Thus

$$\langle \chi_{\boldsymbol{\pi}}, \chi_{\boldsymbol{\sigma}} \rangle = \int \overline{\chi_{\boldsymbol{\pi}}(x)} \chi_{\boldsymbol{\sigma}}(x) dx = \dim \mathsf{Hom}_G(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}})$$

□

Corollary 6.3. *Let* $\boldsymbol{\pi}$ *be a finite dimensional representation of* G*. Then*

$$\boldsymbol{\pi} \text{ is irreducible } \iff ||\chi_{\boldsymbol{\pi}}||_2 = \sqrt{\langle \chi_{\boldsymbol{\pi}}, \chi_{\boldsymbol{\pi}} \rangle} = 1$$

Corollary 6.4. *Let* $\boldsymbol{\pi}, \boldsymbol{\sigma} \in \hat{G}$*. Then*

$$\langle \chi_{\boldsymbol{\pi}}, \chi_{\boldsymbol{\sigma}} \rangle = \delta_{\boldsymbol{\pi}\boldsymbol{\sigma}} \quad (= 1 \text{ if } \boldsymbol{\pi} \text{ is } \textbf{equivalent} \text{ to } \boldsymbol{\sigma}, \text{ 0 otherwise})$$

Corollary 6.5. *Let* $\boldsymbol{\sigma}$ *be a finite dimensional representation of* G *and* $\boldsymbol{\sigma} = \bigoplus_I n_{\boldsymbol{\pi}} \boldsymbol{\pi}$ *(summation over a finite subset* $I \subseteq \hat{G}$*) be a decomposition into irreducible components. Then*
(a) $n_{\boldsymbol{\pi}} = \langle \chi_{\boldsymbol{\pi}}, \chi_{\boldsymbol{\sigma}} \rangle$ *is well determined*
(b) $||\chi_{\boldsymbol{\sigma}}||^2 = \sum_I n_{\boldsymbol{\pi}}^2$

Proof. Let $V_{\boldsymbol{\sigma}} = \bigoplus (V_{\boldsymbol{\tau}} \otimes \mathbb{C}^{n_{\boldsymbol{\tau}}})$. Since each G-morphism $V_{\boldsymbol{\sigma}} \to V_{\boldsymbol{\pi}}$ must vanish on all isotypical components $V_\tau \otimes \mathbb{C}^{n_{\boldsymbol{\tau}}}$ where $\boldsymbol{\tau}$ is not equivalent to $\boldsymbol{\pi}$, we have

$$\begin{aligned} \mathrm{Hom}_G(V_{\boldsymbol{\sigma}}, V_{\boldsymbol{\pi}}) &= \mathrm{Hom}_G(V_{\boldsymbol{\pi}} \otimes \mathbb{C}^{n_{\boldsymbol{\pi}}}, V_{\boldsymbol{\pi}}) \\ &= \mathbb{C}^{n_{\boldsymbol{\pi}}} \otimes \mathrm{Hom}_G(V_{\boldsymbol{\pi}}, V_{\boldsymbol{\pi}}) = \mathbb{C}^{n_{\boldsymbol{\pi}}} \end{aligned}$$

This proves assertion (a). Moreover (b) follows from Pythagoras theorem and Corollary 6.3

Corollary 6.6. *The set of characters* $(\chi_{\boldsymbol{\pi}})_{\boldsymbol{\pi} \in \hat{G}}$ *is an orthonormal basis of the Hilbert space* $\mathbb{L}^2(G)_{inv}$. *Every function* $f \in \mathbb{L}^2(G)_{inv}$ *can be expanded in the series*

$$f = \sum_{\hat{G}} \langle \chi_{\boldsymbol{\pi}}, f \rangle \chi_{\boldsymbol{\pi}} \qquad \text{(convergence in } \mathbb{L}^2(G)\text{)}$$

Theorem 6.7. *Let G be a compact group. For $\boldsymbol{\pi} \in \hat{G}$, let $P_{\boldsymbol{\pi}}$ denote the projector $\mathbb{L}^2(G) \to \mathbb{L}^2(G, \boldsymbol{\pi})$ onto the isotypical component of $\boldsymbol{\pi}$ (in the right regular representation). Then $P_{\boldsymbol{\pi}}$ is given by the convolution with the normalized character $\vartheta_{\boldsymbol{\pi}} = dim\ \boldsymbol{\pi} \cdot \chi_{\boldsymbol{\pi}}$*

$$P_{\boldsymbol{\pi}} : f \to f_{\boldsymbol{\pi}} = P_{\boldsymbol{\pi}} f = f * \vartheta_{\boldsymbol{\pi}}$$

Proof. We have already seen that

$$f_{\boldsymbol{\pi}}(x) = \text{dim } \boldsymbol{\pi} \cdot Tr(\boldsymbol{\pi}(\check{f})\boldsymbol{\pi}(x))$$

Thus

$$\begin{aligned} f_{\boldsymbol{\pi}}(x) &= \text{dim } \boldsymbol{\pi} \cdot Tr\left(\int f(y^{-1})\boldsymbol{\pi}(y)\boldsymbol{\pi}(x)dy\right) \\ &= \text{dim } \boldsymbol{\pi} \cdot \int f(y)Tr(\boldsymbol{\pi}(y^{-1}x))dy = f * \vartheta_{\boldsymbol{\pi}}(x) \end{aligned}$$

□

Exercise 6.8. Let H_1 and H_2 be two Hilbert spaces. Prove that for any operator A in $H_1 \otimes H_2$ which commutes to all operators $T \otimes 1$ $T \in \mathrm{End}(H_1)$ can be written in the form $1 \otimes B$ for some $B \in \mathrm{End}\ (H_2)$. (**Hint:** Introduce an orthonormal basis (e_i) of H_1 and write A as an matrix of blocks with respect to this basis

$$A(e_j \otimes x) = \sum_i e_i \otimes A^i_j x \qquad\qquad (A^i_j \in \mathrm{End}(H_2)).$$

Using the commutations $(P_j \otimes 1)A = A(P_j \otimes 1)$ where P_j is the orthogonal projector on $\mathbb{C}e_j$, conclude that $A^i_j = 0$ for $i \neq j$. Finally, using the commutation relations of A with the operators $U_{ji} \otimes 1$

$$U_{ji}(e_i) = e_j \quad , \quad U_{ji}(e_k) = 0 \text{ for } k \neq i,$$

conclude that $A^i_i = B \in \mathrm{End}(H_2)$ is independent of i.)

Exercise 6.9. Check that the formula

$$f = \sum \langle \chi_{\boldsymbol{\pi}}, f \rangle \chi_{\boldsymbol{\pi}}$$

coincides with the Fourier inversion formula.

7 Induced representations and Frobenius-Weil reciprocity

Suppose that K is a closed (hence compact) subgroup of G.

Recall that $K\backslash G$ is the space of right cosets Kg, $g \in G$.

Suppose for the moment that $K\backslash G$ is finite, i.e. $K\backslash G = \{Kg_1, \ldots, Kg_n\}$ for some n so that G is the disjoint union of $Kg_1, \ldots, Kg_n$.

For each $s \in G$ we have an associated permutation $\boldsymbol{\pi}(s)$ of $\{1, \ldots, n\}$ that sends i to the unique j with $Kg_i s^{-1} = Kg_j$.

We can define an representation ρ of G on $\mathbb{C}^n$ by

$$\rho(s)(a_1, \ldots, a_n) := (a_{\boldsymbol{\pi}^{-1}(1)}, \ldots, a_{\boldsymbol{\pi}^{-1}(n)})$$

Equivalently, if we think of $\mathbb{C}^n$ as the space of functions from $K\backslash G$ into $\mathbb{C}$, then, for $s \in G$ and a coset $L \in K\backslash G$,

$$\rho(s)f(L) = f(Ls)$$

The space of functions from $K\backslash G$ into $\mathbb{C}$ can be identified with the space of functions from G to $\mathbb{C}$ that are constant on right cosets of K, that is, with the space of functions $f : G \to \mathbb{C}$ such that

$$f(kx) = f(x), \quad k \in K,\ x \in G$$

Note that ρ is just the right regular representation for G restricted to this subspace.

Can we build other representations of G by similar constructions?

We return to the case where K is an arbitrary closed subgroup of G (so that $K\backslash G$ is not necessarily finite).

The canonical projection $G \xrightarrow{p} K\backslash G$ pushes the Haar measure ds on G forward to a measure $d\dot{x}$ on $K\backslash G$ characterized by

$$\int_{K\backslash G} f(\dot{x})d\dot{x} = \int_G f(p(s))ds \qquad\qquad (f \in \mathcal{C}(K\backslash G))$$

Lemma 7.1. *Negligible sets in $K\backslash G$ (relative to the measure $d\dot{x}$) are those sets N for which $p^{-1}(N)$ is negligible in G (relative to the Haar measure ds of G). Moreover, for any $f \in \mathcal{C}(G)$ (or by extension for any $f \in \mathbb{L}^1(G)$)*

$$\int_{K\backslash G} \left[\int_K f(kx)\, dk\right] d\dot{x} = \int_G f(x)dx$$

In particular, the measure $d\dot{x}$ is invariant under left translations from G in $K\backslash G$.

Proof. Exercise.

Now let $(\boldsymbol{\sigma}, V_{\boldsymbol{\sigma}})$ be a unitary representation of K. We define the Hilbert space $\mathbb{L}^2(G, V_{\boldsymbol{\sigma}})$ as the completion of the space $\mathcal{C}(G, V_{\boldsymbol{\sigma}})$ of continuous functions $G \to V_{\boldsymbol{\sigma}}$ with the norm

$$||f||^2 = \int_G ||f(x)||^2 dx$$

The norm under the integral sign is computed in $V_{\boldsymbol{\sigma}}$.

Note that $\mathbb{L}^2(G, V_{\boldsymbol{\sigma}})$ is a Hilbert space with inner product

$$\langle f, g \rangle = \int_G \langle f(x), g(x) \rangle dx$$

The inner product under the integral sign is computed in $V_{\boldsymbol{\sigma}}$.

The elements

$$f \in \mathbb{L}^2(G, V_{\boldsymbol{\sigma}}) \text{ such that } f(kx) = \boldsymbol{\sigma}(k) f(x) \qquad \text{for all } k \in K$$

constitute a subspace $H \subseteq \mathbb{L}^2(G, V_{\boldsymbol{\sigma}})$.

Since $||f(x)||$ only depends on the coset Kx of x for $f \in H$ ($\boldsymbol{\sigma}$ is assumed to be unitary), we can take the norm and inner product on H to be defined by

$$\begin{aligned} ||f||_H^2 &= \int_{K \backslash G} ||f(\dot{x})||^2 \, d\dot{x} \\ \langle f, g \rangle_H &= \int_{K \backslash G} \langle f(\dot{x}), g(\dot{x}) \rangle \, d\dot{x} \end{aligned}$$

As before, the norm and the inner product under the integral sign are computed in $V_{\boldsymbol{\sigma}}$.

Note that if $f \in H$, r is the analogue of the right regular representation of G on $\mathbb{L}^2(G, V_{\boldsymbol{\sigma}})$ (that is, $(r(s)h)(x) = h(xs)$ for $h \in \mathbb{L}^2(G, V_{\boldsymbol{\sigma}})$ and $s, x \in G$) and $k \in K$, then

$$(r(s)f)(kx) = f(kxs) = \boldsymbol{\sigma}(k)f(xs) = \boldsymbol{\sigma}(k)(r(s)f)(x),$$

so that $r(s)f \in H$ also. That is, H is r-invariant.

The *induced representation* $\rho = \mathrm{Ind}_K^G(\boldsymbol{\sigma})$ is by definition the "right regular representation" r of G restricted to $H \subseteq \mathbb{L}^2(G, V_{\boldsymbol{\sigma}})$.

The induced representation is unitary.

For example, if $\boldsymbol{\sigma}$ is the identity representation of K (in dimension 1), $V_{\boldsymbol{\sigma}} = \mathbb{C}$, $\mathbb{L}^2(G, V_{\boldsymbol{\sigma}}) = \mathbb{L}^2(G)$, then H is simply $\mathbb{L}^2(K \setminus G)$ and we get the construction considered at the start of this section.

Write H^G for the G-fixed elements in H (with respect to the action given by ρ).

As before, write $V_{\boldsymbol{\sigma}}^K$ for the K-fixed elements of $V_{\boldsymbol{\sigma}}$ (with respect to the action given by $\boldsymbol{\sigma}$).

Proposition 7.2. *(1) The linear map $H^G \to V_{\boldsymbol{\sigma}}^K$ given by $f \to f(e)$ is an isomorphism of vector spaces.*
(2) Let $(\boldsymbol{\pi}, H_{\boldsymbol{\pi}})$ be a unitary representation of G. Then there is an equivalence

$$\boldsymbol{\pi} \otimes \mathit{Ind}_K^G(\boldsymbol{\sigma}) \xrightarrow{\sim} \mathit{Ind}_K^G(\boldsymbol{\pi}|_K \otimes \boldsymbol{\sigma}) \quad : \quad H_{\boldsymbol{\pi}} \hat{\otimes} H \to \widetilde{H}$$

given by $v \otimes f \to \varphi$ with $\varphi(x) = [\boldsymbol{\pi}(x)v] \otimes f(x)$

Proof. The elements of H^G are certainly functions $f : G \to V_{\boldsymbol{\sigma}}$ which are (equal nearly everywhere to a) constant

$$f(x) = f(ex) = r(x)f(e) = f(e)$$

In particular,

$$f(k) = f(e), \quad k \in K$$

By definition of H, $f \in H$ implies

$$f(k) = f(ke) = \boldsymbol{\sigma}(k)f(e), \quad k \in K$$

Thus, $f(e) = \boldsymbol{\sigma}(k)f(e)$ for all $k \in K$ and so $f(e) \in V_{\boldsymbol{\sigma}}^K$, giving part (1) of the proposition.

To check part (2), let us first show that the functions φ (as defined in the proposition) belong to the space of the induced representation $\mathrm{Ind}_K^G(\boldsymbol{\pi}|_K \otimes \boldsymbol{\sigma})$

$$\begin{aligned} \varphi(kx) &= [\boldsymbol{\pi}(kx)v] \otimes f(kx) = [\boldsymbol{\pi}(k)\boldsymbol{\pi}(x)v] \otimes [\boldsymbol{\sigma}(k)f(x)] \\ &= [[\boldsymbol{\pi}|_K \otimes \boldsymbol{\sigma}](k)](\varphi(x)) \end{aligned}$$

Next we show that $v \otimes f \to \varphi$ is a G-morphism (intertwining $\boldsymbol{\pi} \otimes \rho$) and $\tilde{\rho} = \mathrm{Ind}_K^G(\boldsymbol{\pi}|_K \otimes \boldsymbol{\sigma})$, which we recall is the restriction of the the right regular representation to $\tilde{H}$. Note that

$$[\boldsymbol{\pi} \otimes \rho](s)(v \otimes f) = [\boldsymbol{\pi}(s)v] \otimes [\rho(s)f]$$

is mapped to the function on G given by

$$\begin{aligned} x \quad &\mapsto \quad [\boldsymbol{\pi}(x)\boldsymbol{\pi}(s)v] \otimes [(\rho(s)f)(x)] = [\boldsymbol{\pi}(xs)v] \otimes f(xs) \\ &= \quad (\tilde{\rho}(s)\varphi)(x) \end{aligned}$$

as desired.

Now we check that $v \otimes f \to \varphi$ is isometric and hence injective. If (e_i) is an orthonormal basis of $H_{\boldsymbol{\pi}}$, every element of $H_{\boldsymbol{\pi}} \hat{\otimes} H$ can be written uniquely as $\sum e_i \otimes f_i$ with $\sum ||f_i||^2 < \infty$, and such an element has its image the function $\varphi = \sum \varphi_i$ given by

$$x \mapsto \sum \boldsymbol{\pi}(x) e_i \otimes f_i(x)$$

The norm of φ is

$$\begin{aligned} ||\varphi||^2_{\tilde{H}} &= \int_{K\backslash G} ||\varphi(x)||^2 d\dot{x} = \int_{K\backslash G} ||\sum \boldsymbol{\pi}(x) e_i \otimes f_i(x)||^2 d\dot{x} \\ &= \int_{K\backslash G} \sum ||f_i(x)||^2 d\dot{x} = \sum \int_{K\backslash G} ||f_i(x)||^2 d\dot{x} \\ &= \sum ||f_i||^2_H = ||\sum e_i \otimes f_i||^2 \end{aligned}$$

The third equality is justified by the fact that $\boldsymbol{\pi}(x) e_i$ is also an orthonormal basis of $H_{\boldsymbol{\pi}}$, since $\boldsymbol{\pi}$ is unitary.

Finally to see that $v \otimes f \to \varphi$ is onto, it is enough to see that all continuous functions $\Phi \in \tilde{H}$ belong to the image. The function Φ has a unique expansion in the orthonormal basis $(\boldsymbol{\pi}(x)e_i)$ of $H_{\boldsymbol{\pi}}$:

$$\begin{aligned} \Phi(x) &= \sum \boldsymbol{\pi}(x)e_i \otimes f_i(x) \qquad\qquad (f_i(x) \in V) \\ ||\Phi(x)||^2 &= \sum ||f_i(x)||^2 \end{aligned}$$

Therefore

$$\begin{aligned} \sum \boldsymbol{\pi}(kx)e_i \otimes f_i(kx) &= \Phi(kx) = [\boldsymbol{\pi}(k) \otimes \boldsymbol{\sigma}(k)]\Phi(x) \\ &= \sum [\boldsymbol{\pi}(k)\boldsymbol{\pi}(x)e_i] \otimes [\boldsymbol{\sigma}(k)f_i(x)] \end{aligned}$$

The uniqueness of the decomposition gives $f_i(kx) = \boldsymbol{\sigma}(k)f_i(x)$. Thus, $f_i \in H$, as required. □

Note that if we consider $\mathbb{C}$ as a trivial G- or K- space, then

$$\mathrm{Hom}_G(\mathbb{C}, H) \cong H^G$$

and

$$\mathrm{Hom}_K(\mathbb{C}, V_{\boldsymbol{\sigma}}) \cong V_{\boldsymbol{\sigma}}^K$$

(We identify $A \in \mathrm{Hom}_G(\mathbb{C}, H)$ with the image $h \in H$ of $1 \in \mathbb{C}$ and observe that the assumption that A intertwines with the induced representation H is equivalent to the assumption that $\rho(g)h = h$ for all $g \in G$. A similar comment holds for $\mathrm{Hom}_K(\mathbb{C}, V_{\boldsymbol{\sigma}})$.)

Therefore, the isomorphism of part (1) of the preceding proposition can be written as

$$\mathrm{Hom}_G(\mathbb{C}, H) \xrightarrow{\sim} \mathrm{Hom}_K(\mathbb{C}, V_{\boldsymbol{\sigma}})$$

This form admits the following generalization.

Theorem 7.3 (Frobenius-Weil). *Let $(\boldsymbol{\pi}, H_{\boldsymbol{\pi}})$ be a unitary representation of a compact group group G and $(\boldsymbol{\sigma}, V_{\boldsymbol{\sigma}})$ a unitary representation of one of its closed subgroups K. Put $\rho = \mathit{Ind}_K^G(\boldsymbol{\sigma})$ and $H = H_\rho$. Then there is a canonical isomorphism*

$$\mathit{Mor}_G(H_{\boldsymbol{\pi}}, H_\rho) \xrightarrow{\sim} \mathit{Mor}_K(H_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}})$$

where we take for morphisms between two representations spaces, the Hilbert-Schmidt morphisms. (If our spaces are finite dimensional, then Mor_G is what we have written before as Hom_G and, similarly, Mor_K is just Hom_K.)

Proof. We have seen in finite dimensions that in the identification

$$\check{H}_{\boldsymbol{\pi}} \otimes H_\rho \xrightarrow{\sim} \mathsf{Hom}(H_{\boldsymbol{\pi}}, H_\rho)$$

the representation $\check{\boldsymbol{\pi}} \otimes \rho$ is transformed into the representation

$$A \to \rho(s) A \boldsymbol{\pi}(s)^{-1} \qquad (s \in G, A \in \mathsf{Hom}(H_{\boldsymbol{\pi}}, H_\rho))$$

Much the same thing happens for infinite dimensional unitary representations , we complete the algebraic tensor product of Hilbert spaces and get an isomorphism with the space of Hilbert-Schmidt operators.

$$\check{H}_{\boldsymbol{\pi}} \hat{\otimes} H_\rho \xrightarrow{\sim} \mathsf{Mor}_{HS}(H_{\boldsymbol{\pi}}, H_\rho)$$

Thus G-morphisms $H_{\boldsymbol{\pi}} \to H_\rho$ correspond to G-invariants in $\check{H}_{\boldsymbol{\pi}} \hat{\otimes} H_\rho$. In other words,

$$\mathrm{Mor}_G(H_{\boldsymbol{\pi}}, H_\rho) \cong (\check{H}_{\boldsymbol{\pi}} \hat{\otimes} H_\rho)^G$$

Since $\check{H}_{\boldsymbol{\pi}} \hat{\otimes} H_\rho = \tilde{H}$ can be identified with the space of the representation of G induced from the representation $\boldsymbol{\pi}|_K \otimes \boldsymbol{\sigma}$ of K (part (2) of Proposition 7.2), we infer from part (1) of Proposition 7.2 that

$$(\check{H}_{\boldsymbol{\pi}} \hat{\otimes} H_\rho)^G \cong \tilde{H}^G \cong (\check{H}_{\boldsymbol{\pi}} \hat{\otimes} V_{\boldsymbol{\sigma}})^K$$

The conclusion follows from the identity

$$(\check{H}_{\boldsymbol{\pi}} \hat{\otimes} V_{\boldsymbol{\sigma}})^K \cong \mathrm{Mor}_K(H_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}})$$

□

Corollary 7.4. *Consider* $(\boldsymbol{\pi}, H_{\boldsymbol{\pi}}) \in \hat{G}$ *and* $(\boldsymbol{\sigma}, V_{\boldsymbol{\sigma}}) \in \hat{K}$*. Then the multiplicity of* $\boldsymbol{\pi}$ *in* $\mathit{Ind}_K^G(\boldsymbol{\sigma})$ *is the same as the multiplicity of* $\boldsymbol{\sigma}$ *in* $\boldsymbol{\pi}|_K$

Proof. Denote the (infinite dimensional in general) space of $\rho = \mathrm{Ind}_K^G(\boldsymbol{\sigma})$ by H_ρ.

Any G-morphism $H_{\boldsymbol{\pi}} \to H_\rho$ must send $H_{\boldsymbol{\pi}}$ into the isotypical component $\boldsymbol{\pi}$ in H_ρ: this isotypical component is isomorphic to a $\oplus_I H_{\boldsymbol{\pi}}$ whence

$$\mathrm{Mor}_G(H_{\boldsymbol{\pi}}, \oplus_I H_{\boldsymbol{\pi}}) = \bigoplus_I \mathrm{Mor}_G(H_{\boldsymbol{\pi}}, H_{\boldsymbol{\pi}}) = \bigoplus_I \mathbb{C}$$

by Schur's lemma.

On the other hand, every K-morphism $H_{\boldsymbol{\pi}} \to V_{\boldsymbol{\sigma}}$ must vanish off of those components of $H_{\boldsymbol{\pi}}$ into a direct sum of irreducibles (for K) which are not equivalent to $(\boldsymbol{\sigma}, V_{\boldsymbol{\sigma}})$.

Let us write the direct sum of the copies of $(\boldsymbol{\sigma}, V_{\boldsymbol{\sigma}})$ in $H_{\boldsymbol{\pi}}$ as $V_{\boldsymbol{\sigma}} \otimes \mathbb{C}^m$. Then

$$\mathsf{Mor}_K(H_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}}) = \mathsf{Mor}_K(V_{\boldsymbol{\sigma}} \otimes \mathbb{C}^m, V_{\boldsymbol{\sigma}}) \xrightarrow{\sim} \mathsf{Mor}(\mathbb{C}^m, \mathsf{Mor}_K(V_{\boldsymbol{\sigma}}, V_{\boldsymbol{\sigma}}))$$

To see the isomorphism

$$\mathrm{Mor}_K(V_{\boldsymbol{\sigma}} \otimes \mathbb{C}^m, V_{\boldsymbol{\sigma}}) \xrightarrow{\sim} \mathrm{Mor}(\mathbb{C}^m, \mathrm{Mor}_K(V_{\boldsymbol{\sigma}}, V_{\boldsymbol{\sigma}})),$$

note that if (e_i) is a basis for $\mathbb{C}^m$ we can write any element of $V_{\boldsymbol{\sigma}} \otimes \mathbb{C}^m$ as $\sum_i v_i \otimes e_i$. Any map $A \in \mathrm{Mor}_K(V_{\boldsymbol{\sigma}} \otimes \mathbb{C}^m, V_{\boldsymbol{\sigma}})$, is specified by the m maps $v \mapsto A(v \otimes e_i) = A_i$ belonging to $\mathrm{Mor}_K(V_{\boldsymbol{\sigma}}, V_{\boldsymbol{\sigma}})$, and we can think of these m maps as a single map from $\mathbb{C}^m$ to $\mathrm{Mor}_K(V_{\boldsymbol{\sigma}}, V_{\boldsymbol{\sigma}})$ defined by $e_i \mapsto A_i$.

Since $\mathrm{Mor}_K(V_{\boldsymbol{\sigma}}, V_{\boldsymbol{\sigma}}) = \mathbb{C}\mathbf{id}_V$ (Schur's lemma), we have

$$\mathrm{Mor}_K(H_{\boldsymbol{\pi}}, V_{\boldsymbol{\sigma}}) \cong \text{ dual of } \mathbb{C}^m$$

The isomorphism of the theorem implies equality of the dimension of the spaces. They are respectively

$$\begin{aligned} \mathrm{Card}(I) &= \text{multiplicity of } \boldsymbol{\pi} \text{ in } \rho = \mathrm{Ind}_K^G(\boldsymbol{\sigma}) \\ m &= \text{multiplicity of } \boldsymbol{\sigma} \text{ in } \boldsymbol{\pi}\big|_K \end{aligned}$$

□

8 Representations of the symmetric group

8.1 Young subgroups, tableaux and tabloids

If $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ is a partition of n, then write $\lambda \vdash n$. We also use the notation $|\lambda| = \sum_i \lambda_i$, so that a partition of n satisfies $|\lambda| = n$

Definition 8.1. Suppose $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l) \vdash n$. The *Ferrers diagram* or *shape*, of λ is an array of n dots having l left-justified rows with row i containing λ_i dots for $1 \leq i \leq l$.

Definition 8.2. For any set T, let $\mathcal{S}_T$ be the set of permutations of T. Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \vdash n$. Then the corresponding *Young subgroup* of $\mathcal{S}_n$ is

$$\mathcal{S}_\lambda = \mathcal{S}_{\{1,2,\ldots\lambda_1\}} \times \mathcal{S}_{\{\lambda_1+1,\lambda_1+2,\ldots,\lambda_1+\lambda_2\}} \times \mathcal{S}_{\{n-\lambda_l+1,n-\lambda_l+2,\ldots,n\}}$$

Now consider the representation $(1 \uparrow_{\mathcal{S}_\lambda}^{\mathcal{S}_n})$, by which we mean the representation of $\mathcal{S}_n$ induced by the trivial representation of the subgroup $\mathcal{S}_\lambda$. If $\pi_1, \pi_2, \ldots, \pi_k$ is a transversal for $\mathcal{S}_\lambda$, then the vector space

$$V^\lambda = \mathbb{C}\{\pi_1\mathcal{S}_\lambda, \pi_2\mathcal{S}_\lambda, \ldots, \pi_k\mathcal{S}_\lambda\}$$

is a module for our induced representation .

Definition 8.3. Suppose $\lambda \vdash n$. A *Young tableau of shape* λ, is an array t obtained by replacing the dots of the Ferrers diagram with the numbers $1, 2, \ldots, n$ bijectively.

Definition 8.4. Two λ-tableaux t_1 and t_2 are *row equivalent*, $t_1 \sim t_2$, if the corresponding rows of the two tableaux contain the same elements. A *tabloid of shape* λ or λ-tabloid is then

$$\{t\} = \{t_1 | t_1 \sim t\}$$

where shape$(t) = \lambda$

Now $\pi \in \mathcal{S}_n$ acts on a tableau $t = (t_{i,j})$ of shape $\lambda \vdash n$ as follows:

$$\pi t = (\pi(t_{i,j}))$$

This induces an action on tabloids by letting

$$\pi\{t\} = \{\pi t\}$$

Exercise: Check that this is well defined, namely independent of the choice of t.

Definition 8.5. Suppose $\lambda \vdash n$. Let

$$M^\lambda = \mathbb{C}\{\{\mathbf{t}_1\}, \ldots, \{\mathbf{t}_k\}\}$$

where $\{t_1\}, \ldots, \{t_k\}$, is a complete list of λ-tabloids. Then M^λ is called the *permutation module corresponding to* λ.

Definition 8.6. Any G-module M is *cyclic* if there is a $\mathbf{v} \in M$ such that

$$M = \mathbb{C}G\mathbf{v}$$

where $G\mathbf{v} = \{g\mathbf{v} | g \in G\}$. In this case we say that M is *generated by* **v**.

Proposition 8.7. *If $\lambda \vdash n$, then M^λ is cyclic, generated by any given λ-tabloid. In addition, dim $M^\lambda = n!/\lambda!$, the number of λ-tabloids.*

Theorem 8.8. *Consider $\lambda \vdash n$ with the Young subgroup $\mathcal{S}_\lambda$ and tabloid $\{t^\lambda\}$, as before. Then $V^\lambda = \mathbb{C}\mathcal{S}_n\mathcal{S}_\lambda$ and $M^\lambda = \mathbb{C}\mathcal{S}_n\{\mathbf{t}^\lambda\}$ are isomorphic as $\mathcal{S}_n$-modules.*

Proof. Let $\pi_1, \pi_2, \ldots, \pi_k$ be a transversal for $\mathcal{S}_\lambda$. Define a map:

$$\theta : V^\lambda \to M^\lambda$$

by $\theta(\pi_i\mathcal{S}_\lambda) = \{\pi_i\mathbf{t}^\lambda\}$ for $i = 1, 2, \ldots, k$ and linear extension. It is not hard to verify that θ is the desired $\mathcal{S}_n$-isomorphism of modules. □

8.2 Dominance and Lexicographic ordering

Definition 8.9. Suppose $\lambda = (\lambda_1, \ldots, \lambda_l)$ and $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ are partitions of n. Then λ dominates μ, written as $\lambda \trianglerighteq \mu$ if

$$\lambda_1 + \lambda_2 + \ldots + \lambda_i \geq \mu_1 + \mu_2 + \ldots \mu_i$$

for all $i \geq 1$. If $i > l$ (respectively, $i > m$), then we take λ_i (respectively, μ_i) to be zero.

Lemma 8.10 (Dominance lemma for partitions). *Let t^λ and s^μ be tableaux of shape λ and μ respectively. If for each index i, the elements of row i of s^μ are all in different columns in t^λ, then $\lambda \trianglerighteq \mu$.*

Proof. By hypothesis, we can sort the entries in each column of t^λ so that the elements of rows $1, 2, \ldots, i$ of s^μ all occur in the first i rows of t^λ.

Now note that

$$\begin{aligned}
\lambda_1 + \lambda_2 + \cdots + \lambda_i &= \text{number of elements in the first } i \text{ rows of } t^\lambda \\
&\geq \text{number of elements of } s^\mu \text{ in the first } i \text{ rows of } t^\lambda \\
&= \mu_1 + \mu_2 + \cdots + \mu_i
\end{aligned}$$

□

Definition 8.11. Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ and $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ be partitions of n. Then $\lambda < \mu$ in *lexicographic order* if for some index i,

$$\lambda_j = \mu_j \text{ for } j < i \text{ and } \lambda_i < \mu_i$$

Proposition 8.12. *If $\lambda, \mu \vdash n$ with $\lambda \trianglerighteq \mu$ then $\lambda \geq \mu$*

Proof. If $\lambda \neq \mu$, then find the first index i where they differ. Thus, $\sum_{j=1}^{i-1} \lambda_j = \sum_{j=1}^{i-1} \mu_j$ and $\sum_{j=1}^{i} \lambda_j > \sum_{j=1}^{i} \mu_j$ (since $\lambda \triangleright \mu$). So $\lambda_i > \mu_i$. □

8.3 Specht Modules

Definition 8.13. Suppose now that the tableau t has rows $R_1, R_2, \ldots R_l$ and columns $C_1, C_2, \ldots, C_k$. Then

$$R_t = \mathcal{S}_{R_1} \times \mathcal{S}_{R_2} \times \ldots \times \mathcal{S}_{R_l}$$

and

$$C_t = \mathcal{S}_{C_1} \times \mathcal{S}_{C_2} \times \ldots \times \mathcal{S}_{C_k}$$

are the *row-stabilizer* and *column-stabilizer* of t respectively.

Note that our equivalence classes can be expressed as $\{t\} = R_t t$.

In general, given a subset $H \subseteq \mathcal{S}_n$, we can form the group algebra elements

$$H^+ = \sum_{\pi \in H} \pi$$

and

$$H^- = \sum_{\pi \in H} \mathsf{sgn}(\pi)\pi$$

For a tableau t, the element R_t^+ is already implicit in the corresponding tabloid by the remark at the end of the previous paragraph. However we will also need to make use of

$$\kappa_t \stackrel{\texttt{def}}{=} C_t^- = \sum_{\pi \in C_t} \mathsf{sgn}(\pi)\pi$$

Note that if t has columns $C_1, C_2, \ldots, C_k$, then κ_t factors as

$$\kappa_t = \kappa_{C_1}\kappa_{C_2}\ldots\kappa_{C_k}$$

Definition 8.14. If t is a tableau, then the associated polytabloid is

$$\mathbf{e}_t = \kappa_t\{\mathbf{t}\}$$

Lemma 8.15. *Let t be a tableau and π be a permutation. Then*

1. $R_{\pi t} = \pi R_t \pi^{-1}$

2. $C_{\pi t} = \pi C_t \pi^{-1}$

3. $\kappa_{\pi t} = \pi \kappa_t \pi^{-1}$

4. $\mathbf{e}_{\pi t} = \pi \mathbf{e}_t$

Proof. 1. We have the following equivalent statements:

$$
\begin{aligned}
\sigma \in R_{\pi t} &\longleftrightarrow \sigma\{\pi t\} = \{\pi t\} \\
&\longleftrightarrow \pi^{-1}\sigma\pi\{t\} = \{t\} \\
&\longleftrightarrow \pi^{-1}\sigma\pi \in R_t \\
&\longleftrightarrow \sigma \in \pi R_t \pi^{-1}
\end{aligned}
$$

The proofs of 2 and 3 are similar to that of part 1.

4. We have

$$\mathbf{e}_{\pi t} = \kappa_{\pi t}\{\boldsymbol{\pi}\mathbf{t}\} = \pi\kappa_t\pi^{-1}\{\boldsymbol{\pi}\mathbf{t}\} = \pi\kappa_t\{\mathbf{t}\} = \pi\mathbf{e}_t$$

□

Definition 8.16. For any partition λ, the corresponding *Specht module*, S^λ is the submodule of M^λ spanned by the polytabloids $\mathbf{e}_t$, where t is of shape λ

Proposition 8.17. *The S^λ are cyclic modules generated by any given polytabloid.*

Proof. This follows from part 4 of Lemma 8.15

8.4 The Submodule theorem

Recall that $H^- = \sum_{\pi \in H} (\operatorname{sgn} \pi)\pi$ for any subset $H \subseteq \mathcal{S}_n$. If $H = \{\pi\}$, then we write π^- for H^-. We need the unique inner product on M^λ for which

$$\langle \{\mathbf{t}\}, \{\mathbf{s}\} \rangle = \delta_{\{t\},\{s\}} \tag{8.1}$$

Lemma 8.18 (Sign Lemma). *Let $H \leq \mathcal{S}_n$ be a subgroup. Then*
1. If $\pi \in H$, then

$$\pi H^- = H^- \pi = (\mathit{sgn}\,\pi) H^-$$

Equivalently, $\pi^- H^- = H^-$.
2. For any $\mathbf{u}, \mathbf{v} \in M^\lambda$,

$$\langle H^- \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, H^-, \mathbf{v} \rangle$$

3. If the transposition $(b, c) \in H$, then we can factor

$$H^- = k(\epsilon - (b, c))$$

where $k \in \mathbb{C}[\mathcal{S}_n]$
4. If t is a tableau with b, c in the same row of t and $(b, c) \in H$, then

$$H^- \{\mathbf{t}\} = \mathbf{0}$$

Proof. Exercise ∎

Corollary 8.19. *Let $t = t^\lambda$ be a λ-tableau and $s = s^\mu$ be a μ-tableau, where $\lambda, \mu \vdash n$. If $\kappa_t\{\mathbf{s}\} \neq 0$, then $\lambda \trianglerighteq \mu$. And if $\lambda = \mu$, then $\kappa_t\{\mathbf{s}\} = \pm \mathbf{e}_t$*

Proof. Suppose b and c are two elements in the same row of s^μ. Then they cannot be in the same column of t^λ, for if so then $\kappa_t = k(\epsilon - (b, c))$ and $\kappa_t\{\mathbf{s}\} = 0$ by parts 3 and 4 in the preceding lemma. Thus the dominance lemma 8.10 yields $\lambda \trianglerighteq \mu$.

If $\lambda = \mu$, then we must have $\{s\} = \pi\{t\}$ for some $\pi \in C_t$ by the same argument that established the dominance lemma. Using part 1 of the preceeding lemma yields

$$\kappa_t\{\mathbf{s}\} = \kappa_t \pi\{\mathbf{t}\} = (\mathsf{sgn}\, \pi)\kappa_t\{\mathbf{t}\} = \pm \mathbf{e}_t$$

□

Corollary 8.20. *If $\mathbf{u} \in M^\mu$ and shape $t = \mu$, then $\kappa_t \mathbf{u}$ is multiple of $\mathbf{e}_t$.*

Proof. We can write $\mathbf{u} = \sum_i c_i \{s_i\}$, where the s_i are μ-tableaux. By the previous corollary, $\kappa_t \mathbf{u} = \sum_i \pm c_i \mathbf{e}_t$. □

Theorem 8.21 (Submodule Theorem). *Let U be a submodule of M^μ. Then*

$$U \supseteq S^\mu \quad \text{or} \quad U \subseteq S^{\mu\perp}$$

In particular the S^μ are irreducible.

Proof. Consider $\mathbf{u} \in U$ and a μ-tableau t. By the preceding corollary, we know that $\kappa_t \mathbf{u} = f\mathbf{e}_t$ for some field element f. There are two cases, depending on which multiples can arise.

Suppose that there exists a $\mathbf{u}$ and a t with $f \neq 0$. Then since $\mathbf{u}$ is in the submodule U, we have $fe_t = \kappa_t \mathbf{u} \in U$. Thus $\mathbf{e}_t \in U$ (since f is nonzero) and $S^\mu \subseteq U$ (since S^μ is cyclic).

On the other hand, suppose we always have $\kappa_t \mathbf{u} = 0$. We claim that this forces $U \subseteq S^{\mu\perp}$. Consider any $\mathbf{u} \in U$. Given an arbitrary μ- tableau t, we can apply part 2 of the sign lemma to obtain

$$\langle \mathbf{u}, \mathbf{e}_t \rangle = \langle \mathbf{u}, \kappa_t \{\mathbf{t}\} \rangle = \langle \kappa_t \mathbf{u}, \{\mathbf{t}\} \rangle = \langle 0, \{\mathbf{t}\} \rangle = 0$$

Since the $\mathbf{e}_t$ span S^μ, we have $\mathbf{u} \in S^{\mu\perp}$, as claimed.

□

Proposition 8.22. *Suppose $\theta \in Hom_{\mathcal{S}_n}(S^\lambda, M^\mu)$ is nonzero. Then $\lambda \trianglerighteq \mu$. Moreover, if $\lambda = \mu$, then θ is multiplication by a scalar.*

Proof. Since $\theta \neq 0$, there is some basis vector $\mathbf{e}_t$ such that $\theta(\mathbf{e}_t) \neq 0$. Because $\langle \cdot, \cdot \rangle$ is an inner product with complex scalars, $M^\lambda = S^\lambda \oplus S^{\lambda\perp}$. Thus we can extend θ to an element of $\text{Hom}_{\mathcal{S}_n}(M^\lambda, M^\mu)$ by setting $\theta(S^{\lambda\perp}) = 0$. So

$$0 \neq \theta(\mathbf{e}_t) = \theta(\kappa_t\{\mathbf{t}\}) = \kappa_t\theta(\{\mathbf{t}\}) = \kappa_t(\sum_i c_i\{s_i\})$$

where the s_i are μ-tableaux. By Corollary 8.19, we have $\lambda \trianglerighteq \mu$.

In the case $\lambda = \mu$, Corollary 8.20 yields $\theta(\mathbf{e}_t) = c\mathbf{e}_t$ for some constant c. So for any permutation π,

$$\theta(\mathbf{e}_{\pi t}) = \theta(\pi\mathbf{e}_t) = \pi\theta(\mathbf{e}_t) = \pi(c\mathbf{e}_t) = c\mathbf{e}_{\pi t}$$

Thus θ is multiplication by c. $\square$

Theorem 8.23. *The S^λ for $\lambda \vdash n$ form a complete list of irreducible $\mathcal{S}_n$-modules.*

Proof. The S^λ are irreducible by the submodule theorem.

Since we have the right number of modules for a full set, it suffices to show that they are pairwise inequivalent. But if $S^\lambda \cong S^\mu$, then there is a nonzero homomorphism $\theta \in \text{Hom}_{\mathcal{S}_n}(S^\lambda, M^\mu)$, since $S^\mu \subseteq M^\mu$. Thus $\lambda \trianglerighteq \mu$ (Proposition 8.22). Similarly $\mu \trianglerighteq \lambda$, so $\lambda = \mu$. $\square$

8.5 Standard Tableaux and a Basis for S^λ

Definition 8.24. A tableau t is *standard* if the rows and columns of t are increasing sequences. In this case we also say that the corresponding tabloid and polytabloid are standard.

Theorem 8.25. *The set*

$$\{e_t : t \text{ is a standard } \lambda\text{-tableau}\}$$

is a basis for S^λ.

Proof. Omitted

8.6 The Branching Rule

Definition 8.26. If λ is a diagram, then an *inner corner of* λ is a node $(i, j) \in \lambda$ whose removal leaves the Ferrers diagram of a partition. Any partition obtained by such a removal is denoted by λ^-. An *outer corner* of λ is a node $(i, j) \notin \lambda$ whose addition produces the Ferrers diagram of a partition. Any partition obtained by such an addition is denoted by λ^+

Lemma 8.27. *We have*

$$f^{\lambda} = \sum_{\lambda^-} f^{\lambda^-}$$

Proof. Every standard tableau of shape $\lambda \vdash n$ consists of n in some inner corner together with a standard tableau of shape $\lambda^- \vdash n-1$. The result follows.

$\square$

Theorem 8.28 (Branching Rule). *If $\lambda \vdash n$, then*
1. $S^\lambda \downarrow_{\mathcal{S}_{n-1}} \cong \bigoplus_{\lambda^-} S^{\lambda^-}$*, and*
2. $S^\lambda \uparrow^{\mathcal{S}_{n+1}} \cong \bigoplus_{\lambda^+} S^{\lambda^+}$

Proof. 1. Let the inner corners of λ appear in rows $r_1 < r_2 < \cdots < r_k$. For each i, let λ^i denote the partition of λ^- obtained by removing the corner cell in row r_i. In addition, if n is at the end of row r_i of tableau t (respectively, in row r_i of tabloid $\{t_i\}$), then t^i (respectively, $\{t^i\}$) will be the array obtained by removing the n.

Now given any group G with module V and submodule W, it is easy to see that

$$V \cong W \oplus (V/W),$$

where V/W is the quotient space. Thus it suffices to find a chain of subspaces

$$\{\mathbf{0}\} = V^{(0)} \subset V^{(1)} \subset V^{(2)} \subset \cdots \subset V^{(k)} = S^\lambda$$

such that $V^{(i)}/V^{(i-1)} \cong S^{\lambda^i}$ as $\mathcal{S}_{n-1}$ modules for $1 \leq i \leq k$. Let $V^{(i)}$ be the vector space spanned by the standard polytabloids $\mathbf{e}_t$ where n appears in t at the end of one of rows r_1 through r_i. We show that the $V^{(i)}$ are our desired modules as follows.

Define maps $\theta_i : M^\lambda \to M^{\lambda^i}$ by linearly extending

$$\{\mathbf{t}\} \stackrel{\theta_i}{\to} \begin{cases} \{\mathbf{t_i}\} & \text{if } n \text{ in row } r_i \text{ of } \{t\}, \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

Verify that θ_i is an $\mathcal{S}_{n-1}$-homomorphism. Furthermore, for standard t we have

$$\mathbf{e}_t \stackrel{\theta_i}{\to} \begin{cases} \mathbf{e}_{t^i} & \text{if } n \text{ is in row } r_i \text{ of } t, \\ \mathbf{0} & \text{if } n \text{ is in row } r_j \text{ of } t, \text{ where } j < i. \end{cases}$$

This is because any tabloid appearing in $\mathbf{e}_t$, t standard, has n in the same row or higher than in t.

Since the standard polytabloids form a basis for the corresponding Specht module,

$$\theta_i V^{(i)} = S^{\lambda^i}$$

and

$$V^{(i-1)} \subseteq \mathrm{ker}\theta_i.$$

We can construct the chain

$$\{0\} = V^{(0)} \subseteq V^{(1)} \cap \mathrm{ker}\theta_1 \subseteq V^{(1)} \subseteq V^{(2)} \cap \mathrm{ker}\theta_2 \subseteq V^{(2)} \subseteq \cdots \subseteq V^{(k)} = S^{\lambda}$$

But

$$\dim \frac{V^{(i)}}{V^{(i)} \cap \mathrm{ker}\theta_i} = \dim \theta_i V^{(i)} = f^{\lambda^i}$$

By the preceding lemma, the dimensions of these quotients add up to dim S^λ. Since this leaves no space to insert extra modules, the chain must have equality for the first, third, etc. containments. Furthermore,

$$\frac{V^{(i)}}{V^{(i-1)}} \cong \frac{V^{(i)}}{V^{(i)} \cap \text{ker}\theta_i} \cong S^{\lambda^i}$$

as desired.

2. We will show that this part follows from the first by Frobenius reciprocity. In fact, parts 1 and 2 can be shown to be equivalent by the same method.

Let χ^λ be the character of S^λ. If $S^\lambda \uparrow^{\mathcal{S}_{n+1}} \cong \oplus_{\mu \vdash n+1} m_\mu S^\mu$, then by taking characters, $\chi^\lambda \uparrow^{\mathcal{S}_{n+1}} \cong \sum_{\mu \vdash n+1} m_\mu \chi^\mu$.

The multiplicities are given by

$$
\begin{aligned}
m_\mu &= \langle \chi^\lambda \uparrow^{\mathcal{S}_{n+1}}, \chi^\mu \rangle \\
&= \langle \chi^\lambda, \chi^\mu \downarrow_{\mathcal{S}_n} \rangle \\
&= \langle \chi^\lambda, \sum_{\mu^-} \chi^{\mu^-} \rangle \\
&= \begin{cases} 1 & \text{if } \lambda = \mu^- \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} 1 & \text{if } \mu = \lambda^+ \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

9 Symmetric Functions

9.1 Symmetric functions in general

Let $x = (x_1, x_2, \ldots)$ be a set of indeterminates. A *homogeneous symmetric function of degree* n over $\mathbb{Q}$ is a formal power series

$$f(x) = \sum_{\alpha} c_\alpha x^\alpha$$

(a) $\alpha = (\alpha_1, \alpha_2, \ldots)$ ranges over all sequence of non-negative integers that sum to n (i.e. weak compositions of n)

(b) $c_\alpha \in \mathbb{Q}$

(c) x^α stands for the monomial $x_1^{\alpha_1} x_2^{\alpha_2} \ldots$

(d) $f(x_{w(1)}, x_{w(2)}, \ldots) = f(x_1, x_2, \ldots)$ for every permutation w of the positive integers.

The set of all homogeneous symmetric functions of degree n over $\mathbb{Q}$ is denoted as Λ^n.

If $f \in \Lambda^m$ and $g \in \Lambda^n$, then it is clear that $fg \in \Lambda^{m+n}$ (where fg is a product of the formal power series). Hence if we define

$$\Lambda = \Lambda^0 \oplus \Lambda^1 \oplus \cdots \qquad \text{(vector space direct sum)}$$

Then Λ has the structure of a $\mathbb{Q}$-algebra

9.2 Monomial Symmetric Functions

Given $\lambda = (\lambda_1, \lambda_2, \ldots) \vdash n$, define a symmetric function $m_\lambda(x) \in \Lambda^n$ by

$$m_\lambda = \sum_\alpha x^\alpha$$

where the sum ranges over all distinct permutations $\alpha = (\alpha_1, \alpha_2, \ldots)$ of the entries of the vector $\lambda = (\lambda_1, \lambda_2, \ldots)$.

We call m_λ a *monomial symmetric function*. Clearly if $f = \sum_\alpha c_\alpha x^\alpha \in \Lambda^n$ then $f = \sum_{\lambda \vdash n} c_\lambda m_\lambda$. The set $\{m_\lambda : \lambda \vdash n\}$ is a (vector space) basis for Λ^n and hence that

$$\dim \Lambda^n = p(n)$$

the number of partitions of n. Moreover the set $\{m_\lambda : \lambda \in \text{Par}\}$ is a basis for Λ.

9.3 Elementary Symmetric Functions

We define the *elementary symmetric functions* e_λ for $\lambda \in \mathrm{Par}$ by the formula

$$\begin{aligned} e_n &= m_{1^n} = \sum_{i_1 < \ldots < i_n} x_{i_1} \cdots x_{i_n}, \qquad n \geq 1 \quad (\text{with } e_0 = m_\emptyset = 1) \\ e_\lambda &= e_{\lambda_1} e_{\lambda_2} \cdots, \qquad \text{if } \lambda = (\lambda_1, \lambda_2, \ldots) \end{aligned}$$

Suppose that $A = (a_{ij})_{i,j\geq 1}$ is an integer matrix with finitely many nonzero entries with row and column sums

$$\begin{aligned} r_i &= \sum_j a_{ij} \\ c_j &= \sum_i a_{ij} \end{aligned}$$

Define the *row-sum vector* row(A) and *column-sum vector* col(A) by

$$\begin{aligned} \text{row}(A) &= (r_1, r_2, \ldots) \\ \text{column}(A) &= (c_1, c_2, \ldots) \end{aligned}$$

A $(0,1)$-*matrix* is a matrix all of whose entries are either 0 or 1.

Proposition 9.1. *Let $\lambda \vdash n$, and let $\alpha = (\alpha_1, \alpha_2, \ldots)$ be a weak composition of n. Then the coefficient $M_{\lambda\alpha}$ of x^α in e_λ is equal to the number of $(0,1)$-matrices $A = (a_{ij})_{i,j\geq 1}$ satisfying row$(A) = \lambda$ and col$(A) = \alpha$. That is,*

$$e_\lambda = \sum_{\mu \vdash n} M_{\lambda\mu} m_\mu \tag{9.1}$$

Corollary 9.2. *Let* $M_{\lambda\mu}$ *be given by Equation (9.1). Then* $M_{\lambda\mu} = M_{\mu\lambda}$*. That is, the transition matrix between the bases* $\{m_\lambda : \lambda \vdash n\}$ *and* $\{e_\lambda : \lambda \vdash n\}$ *is a symmetric matrix.*

Proposition 9.3. *We have*

$$\prod_{i,j}(1+x_iy_j) = \sum_{\lambda,\mu} M_{\lambda\mu} m_\lambda(x) m_\mu(y) \tag{9.2}$$

$$= \sum_{\lambda} m_\lambda(x) e_\lambda(y) \tag{9.3}$$

Here λ and μ range over Par. (It suffices to take $|\lambda| = |\mu|$, since otherwise $M_{\lambda\mu} = 0$.)

Proof. A monomial $x_1^{\alpha_1} x_2^{\alpha_2} \cdots y_1^{\beta_1} y_2^{\beta_2} \cdots = x^\alpha y^\beta$ appearing in the expansion of $\prod(1 + x_i y_j)$ is obtained by choosing a $(0, 1)$-matrix $A = (a_{ij})$ with finitely many 1's, satisfying

$$\prod_{i,j} (x_i y_j)^{a_{ij}} = x^\alpha y^\beta$$

But

$$\prod_{i,j} (x_i y_j)^{a_{ij}} = x^{\text{row}(A)} y^{\text{col}(A)}$$

so the coefficient of $x^\alpha y^\beta$ in the product $\prod(1 + x_i y_j)$ is the number of $(0, 1)$-matrices satisfying row$(A) = \alpha$ and col$(A) = \beta$. Hence equation (9.2) follows. Equation (9.3) is then a consequence of (9.1)

□

Theorem 9.4. *Let $\lambda, \mu \vdash n$. Then $M_{\lambda\mu} = 0$ unless $\lambda' \trianglerighteq \mu$, while $M_{\lambda\lambda'} = 1$. Hence the set $\{e_\lambda : \lambda \vdash n\}$ is a basis for Λ^n (so $\{e_\lambda : \lambda \in \mathit{Par}\}$ is a basis for Λ). Equivalently, $e_1, e_2, \ldots$ are algebraically independent and generate Λ as a $\mathbb{Q}$-algebra, which we write as*

$$\Lambda = \mathbb{Q}[e_1, e_2, \ldots]$$

Proof. Suppose $M_{\lambda\mu} \neq 0$ so by Proposition 9.1 there is a $(0,1)$-matrix A with row$(A) = \lambda$ and col$(A) = \mu$. Let A' be the matrix with row$(A') = \lambda$ and with its $1's$ left justified, i.e. $A'_{ij} = 1$ precisely for $1 \leq j \leq \lambda_i$. For any i the number of $1's$ in the first i columns of A' clearly is not less than the number of $1's$ in the first i columns of A, so by definition of dominance order we have col$(A') \trianglerighteq$ col$(A) = \mu$.

But col$(A') = \lambda'$, so $\lambda' \trianglerighteq \mu$ as desired. Moreover it is easy to see that A' is the *only* $(0,1)$-matrix with row$(A') = \lambda$ and col$(A') = \lambda'$, so $M_{\lambda,\lambda'} = 1$.

The previous argument shows the following. Let $\lambda^1, \lambda^2, \ldots, \lambda^{p(n)}$ be an ordering of Par(n) that is compatible with the dominance order and such that the "reverse conjugate" order $(\lambda^{p(n)})', \ldots, (\lambda^2)', (\lambda^1)'$ is also compatible with the dominance order. (Exercise: give an example of such an order). Then the matrix $(M_{\lambda\mu})$, with the row order $\lambda^1, \lambda^2, \ldots$ and column order $(\lambda^1)', (\lambda^2)', \ldots$ is upper triangular with $1's$ on the main diagonal. Hence it is invertible, so $\{e_\lambda : \lambda \vdash n\}$ is a basis for Λ^n. (In fact it is a basis for $\Lambda^n_{\mathbb{Z}}$ since the diagonal entries are actually $1's$, and not merely nonzero.)

The set $\{e_\lambda : \lambda \in \mathsf{Par}\}$ consists of all monomials $e_1^{a_1} e_2^{a_2} \ldots$ (where $a_i \in \mathbb{N}$, $\sum a_i < \infty$). Hence the linear independence of $\{e_\lambda : \lambda \in \mathsf{Par}\}$ is equivalent to the algebraic independence of $e_1, e_2, \ldots$ as desired

$\square$

9.4 Complete Homogeneous Symmetric functions

Define the *complete homogeneous symmetric functions* (or just *complete symmetric functions*) h_λ for $\lambda \in$ Par by the formulas

$$h_n = \sum_{\lambda \vdash n} m_\lambda = \sum_{i_1 \leq \ldots \leq i_n} x_{i_1} \cdots x_{i_n} \quad \text{(with } h_0 = m_\emptyset = 1\text{)}$$

$$h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots \quad \text{if } \lambda = (\lambda_1, \lambda_2, \ldots) \tag{9.4}$$

Proposition 9.5. *Let $\lambda \vdash n$, and let $\alpha = (\alpha_1, \alpha_2, \ldots)$ be a weak composition of n. Then the coefficient $N_{\lambda\alpha}$ of x^α in h_λ is equal to the number of $\mathbb{N}$-matrices $A = (a_{ij})$ satisfying row$(A) = \lambda$ and col$(A) = \alpha$. That is,*

$$h_\lambda = \sum_{\mu \vdash n} N_{\lambda\mu} m_\mu, \tag{9.5}$$

Corollary 9.6. *Let* $N_{\lambda\mu}$ *be given by Equation (9.5). Then* $N_{\lambda\mu} = N_{\mu\lambda}$ *, i.e the transition matrix between the bases* $\{m_\lambda : \lambda \vdash n\}$ *and* $\{h_\lambda : \lambda \vdash n\}$ *is a symmetric matrix.*

Note that by a Corollary in the next section (Corollary 9.9), it follows that the set $\{h_\lambda : \lambda \vdash n\}$ is indeed a basis.

Proposition 9.7. *We have*

$$\prod_{i,j}(1 - x_i y_j)^{-1} = \sum_{\lambda,\mu} N_{\lambda\mu} m_\lambda(x) m_\mu(y) \tag{9.6}$$

$$= \sum_{\lambda} m_\lambda(x) h_\lambda(y) \tag{9.7}$$

where λ and μ range over Par (and where it suffices to take $|\lambda| = |\mu|$).

9.5 An Involution

Since $\Lambda = \mathbb{Q}[e_1, e_2, \ldots]$, an algebra endomorphism $f : \Lambda \to \Lambda$ is determined uniquely by its values $f(e_n), n \geq 1$; and conversely any choice of $(f(e_n)) \in \Lambda$ determines an endomorphism f. Define an endomorphism $\omega : \Lambda \to \Lambda$ by $\omega(e_n) = h_n$, $n \geq 1$. Thus (since ω preserves multiplication), $\omega(e_\lambda) = h_\lambda$ for all partitions λ.

Theorem 9.8. *The endomorphism ω is an involution i.e. $\omega^2 = 1$ (the identity automorphism), or equivalently $\omega(h_n) = e_n$. (Thus, $\omega(h_\lambda) = e_\lambda$ for all partitions λ.)*

Proof. Consider the formal power series

$$H(t) := \sum_{n \geq 0} h_n t^n \quad \in \Lambda\,[[t]]$$

$$E(t) := \sum_{n \geq 0} e_n t^n \quad \in \Lambda\,[[t]]$$

Check the identities

$$H(t) = \prod_n (1 - x_n t)^{-1} \tag{9.8}$$

$$E(t) = \prod_n (1 + x_n t) \tag{9.9}$$

Hence $H(t)E(-t) = 1$. Equating the coefficients of t^n on both sides yields

$$0 = \sum_{i=0}^{n} (-1)^i e_i h_{n-i}, \qquad n \geq 1 \tag{9.10}$$

Conversely, if $\sum_{i=0}^{n}(-1)^i u_i h_{n-i} = 0$ for all $n \geq 1$, for certain $u_i \in \Lambda$ with $u_0 = 1$, then $u_i = e_i$. Now apply ω to Equation (9.10) to obtain

$$0 = \sum_{i=0}^{n}(-1)^i h_i \omega(h_{n-i}) = (-1)^n \sum_{i=0}^{n}(-1)^i \omega(h_i) h_{n-i},$$

whence $\omega(h_i) = e_i$ as desired.

$\square$

Corollary 9.9. *The set* $\{h_\lambda : \lambda \vdash n\}$ *is a basis for* Λ^n *(so* $\{h_\lambda : \lambda \in \mathit{Par}\}$ *is a basis for* Λ*). Equivalently,* $h_1, h_2, \ldots$ *are algebraically independent and generate* Λ *as a* $\mathbb{Q}$*-algebra, which we write as*

$$\Lambda = \mathbb{Q}[h_1, h_2, \ldots]$$

Proof. Theorem 9.8 shows that the endomorphism $\omega : \Lambda \to \Lambda$ defined by $\omega(e_n) = h_n$ is invertible (since $\omega = \omega^{-1}$), and hence is an automorphism of Λ. The proof now follows from Theorem 9.4.

□

9.6 Power Sum Symmetric Functions

We define a fourth set p_λ of symmetric functions indexed by $\lambda \in \mathrm{Par}$ and called the *power sum symmetric functions* as follows:

$$p_n = m_n = \sum_i x_i^n, \;\; n \geq 1 \quad (\text{with } p_0 = m_\emptyset = 1)$$

$$p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots \quad \text{if } \lambda = (\lambda_1, \lambda_2, \ldots)$$

Proposition 9.10. *Let* $\lambda = (\lambda_1, \ldots, \lambda_l) \vdash n$*, where* $l = l(\lambda)$*, and set*

$$p_\lambda = \sum_{\mu \vdash n} R_{\lambda\mu} m_\mu \tag{9.11}$$

Let $k = l(\mu)$*. Then* $R_{\lambda\mu}$ *is equal to the number of ordered partitions* $\pi = (B_1, B_2, \ldots B_k)$ *of* $\{1, \ldots, l\}$ *such that*

$$\mu_j = \sum_{i \in B_j} \lambda_i, \quad 1 \le j \le k \tag{9.12}$$

Proof. $R_{\lambda\mu}$ is the coefficient of $x^\mu = x_1^{\mu_1} x_2^{\mu_2} \ldots$ in

$$p_\lambda = (\sum x_i^{\lambda_1})(\sum x_i^{\lambda_2}) \cdots$$

To obtain the monomial x^μ in the expansion of the product, we choose the term $x_{i_j}^{\lambda_j}$ from each factor $\sum x_i^{\lambda_j}$ so that $\prod_j x_{i_j}^{\lambda_j} = x^\mu$. Define $B_r = \{j : i_j = r\}$. Then $(B_1, \ldots B_k)$ will be an ordered partition of $\{1, \ldots, l\}$ satisfying Equation (9.12), and conversely every such ordered partition gives rise to a term x^μ.

□

Corollary 9.11. *Let* $R_{\lambda\mu}$ *be as in Equation (9.11). Then* $R_{\lambda\mu} = 0$ *unless* $\mu \trianglerighteq \lambda$*, while*

$$R_{\lambda\lambda} = \prod_i m_i(\lambda)! \tag{9.13}$$

where $\lambda = \langle 1^{m_1(\lambda)} 2^{m_2(\lambda)} \cdots \rangle$*. Hence* $\{p_\lambda : \lambda \vdash n\}$ *is a basis for* Λ^n *(so* $\{p_\lambda : \lambda \in \mathit{Par}\}$ *is a basis for* Λ*). Equivalently,* $p_1, p_2, \ldots$ *are algebraically independent and generate* Λ *as a* $\mathbb{Q}$*-algebra, i.e.*

$$\Lambda = \mathbb{Q}[p_1, p_2, \ldots]$$

We now consider the effect of the involution ω on p_λ. For any partition $\lambda = \langle 1^{m_1(\lambda)} 2^{m_2(\lambda)} \cdots \rangle$ define

$$z_\lambda = 1^{m_1} m_1! 2^{m_2} m_2! \cdots \tag{9.14}$$

If $w \in \mathcal{S}_n$, then the *cycle type* $\rho(w)$ of w is the partition $\rho(w) = (\rho_1, \rho_2, \ldots) \vdash n$ such that the cycle lengths of w (in its factorization into disjoint cycles) are $\rho_1, \rho_2, \ldots$. The number of permutations $w \in \mathcal{S}_n$ of a fixed cycle type $\rho = \langle 1^{m_1} 2^{m_2} \cdots \rangle$ is given by

$$\begin{aligned} \#\{w \in \mathcal{S}_n : \rho(w) = \rho\} &= \frac{n!}{1^{m_1} m_1! 2^{m_2} m_2! \cdots} \\ &= n! z_\rho^{-1} \end{aligned} \tag{9.15}$$

The set $\{v \in \mathcal{S}_n : \rho(v) = \rho\}$ is just the conjugacy class in $\mathcal{S}_n$ containing w. For any finite group G, the order $\# K_w$ of the conjugacy class K_w is equal to the index $[G : C(w)]$ of the centralizer of w.

Hence:

Proposition 9.12. *Let $\lambda \vdash n$. Then z_λ is equal to the number of permutations $v \in \mathcal{S}_n$ that commute with a fixed w_λ of cycle type λ.*

For a partition $\lambda = \langle 1^{m_1} 2^{m_2} \cdots \rangle$, define

$$\varepsilon_\lambda = (-1)^{m_2+m_4+\cdots} = (-1)^{n-l(\lambda)} \tag{9.16}$$

Thus for any $w \in \mathcal{S}_n$, $\varepsilon_{p(w)}$ is $+1$ if w is an even permutation and -1 otherwise, so the map $\mathcal{S}_n \to \{\pm 1\}$ defined by $w \mapsto \varepsilon_{p(w)}$ is the usual "sign homomorphism"

Proposition 9.13. *We have*

$$\prod_{i,j}(1-x_iy_j)^{-1} = \exp\sum_{n\geq 1}\frac{1}{n}p_n(x)p_n(y)$$

$$= \sum_{\lambda} z_\lambda^{-1}p_\lambda(x)p_\lambda(y) \tag{9.17}$$

$$\prod_{i,j}(1+x_iy_j) = \exp\sum_{n\geq 1}\frac{1}{n}(-1)^{n-1}p_n(x)p_n(y)$$

$$= \sum_{\lambda} z_\lambda^{-1}\varepsilon_\lambda p_\lambda(x)p_\lambda(y) \tag{9.18}$$

Proof. We have

$$\begin{aligned}\log \prod_{i,j} (1 - x_i y_j)^{-1} &= \sum_{i,j} \log (1 - x_i y_j)^{-1} \\ &= \sum_{i,j} \sum_{n \geq 1} \frac{1}{n} x_i^n y_j^n \\ &= \sum_{n \geq 1} \frac{1}{n} \left(\sum_i x_i^n \right) \left(\sum_j y_j^n \right) \\ &= \sum_{n \geq 1} \frac{1}{n} p_n(x) p_n(y)\end{aligned}$$

□

Proposition 9.14. *Let $\lambda \vdash n$. Then*

$$\omega p_\lambda = \varepsilon_\lambda p_\lambda$$

In other words, p_λ is an eigenvector for ω corresponding to the eigenvalue ε_λ.

Proof. Regard ω as acting on symmetric functions in the variables $y = (y_1, y_2, \ldots)$. Those in the variables x are regarded as scalars.

Apply ω to Equation (9.17). We obtain

$$\begin{aligned}
\omega \sum_{\lambda} z_{\lambda}^{-1} p_{\lambda}(x) p_{\lambda}(y) &= \omega \prod_{i,j} (1 - x_i y_j)^{-1} \\
&= \sum_{v} m_v(x) \omega h_v(y) \quad \text{(by (9.6))} \\
&= \sum_{v} m_v(x) e_v(y) \quad \text{(by Theorem 9.8)} \\
&= \prod_{i,j} (1 + x_i y_j) \quad \text{(by (9.2))} \\
&= \sum_{\lambda} z_{\lambda}^{-1} \varepsilon_{\lambda} p_{\lambda}(x) p_{\lambda}(y) \quad \text{(by 9.18)}
\end{aligned}$$

Since the $p_\lambda(x)$'s are linearly independent, their coefficients in the first and last sums of the above chain of equalities must be the same. In other words, $\omega p_\lambda(y) = \varepsilon_\lambda p_\lambda(y)$, as desired. □

Proposition 9.15. *We have*

$$h_n = \sum_{\lambda \vdash n} z_\lambda^{-1} p_\lambda \tag{9.19}$$

$$e_n = \sum_{\lambda \vdash n} \varepsilon_\lambda z_\lambda^{-1} p_\lambda \tag{9.20}$$

$$\tag{9.21}$$

Proof. Substituting $y = (t, 0, 0, \ldots)$ in Equation (9.17) immediately yields Equation (9.19). Equation (9.20) is similiarly obtained from (9.18), or by applying ω to (9.19). □

9.7 A Scalar product

Define a scalar product on Λ by requiring that $\{m_\lambda\}$ and $\{h_\mu\}$ be dual bases i.e.

$$\langle m_\lambda, h_\mu \rangle = \delta_{\lambda\mu} \tag{9.22}$$

for all $\lambda, \mu \in$ Par. Notice that $\langle \cdot, \cdot \rangle$ respects the grading of Λ, in the sense that if f and g are homogeneous then $\langle f, g \rangle = 0$ unless deg $f =$ deg g.

Proposition 9.16. *The scalar product* $\langle \cdot, \cdot \rangle$ *is symmetric, i.e.* $\langle f, g \rangle = \langle g, f \rangle$ *for all* $f, g \in \Lambda$*.*

Proof. The result is equivalent to Corollary 9.6. More specifically, it suffices by linearity to prove $\langle f, g \rangle = \langle g, f \rangle$ for some bases $\{f\}$ and $\{g\}$ of Λ. Take $\{f\} = \{g\} = \{h_\lambda\}$. Then

$$\langle h_\lambda, h_\mu \rangle = \left\langle \sum_\nu N_{\lambda\nu} m_\nu, h_\mu \right\rangle = N_{\lambda\mu} \tag{9.23}$$

Since $N_{\lambda\mu} = N_{\mu\lambda}$ by Corollary 9.6, we have $\langle h_\lambda, h_\mu \rangle = \langle h_\mu, h_\lambda \rangle$ as desired. □

Lemma 9.17. *Let* $\{u_\lambda\}$ *and* $\{v_\lambda\}$ *be bases of* Λ *such that for all* $\lambda \vdash n$ *we have* $u_\lambda, v_\lambda \in \Lambda^n$. *Then* $\{u_\lambda\}$ *and* $\{v_\lambda\}$ *are dual bases if and only if*

$$\sum_\lambda u_\lambda(x) v_\lambda(y) = \prod_{i,j} (1 - x_i y_j)^{-1}$$

Proof. Write $m_\lambda = \sum_\rho \zeta_{\lambda\rho} u_\rho$ and $h_\mu = \sum_\nu \eta_{\mu\nu} v_\nu$. Thus

$$\delta_{\lambda\mu} = \langle m_\lambda, h_\mu \rangle = \sum_{\rho,\nu} \zeta_{\lambda\rho}\eta_{\mu\nu}\langle u_\rho, v_\nu\rangle \tag{9.24}$$

For each fixed $n \geq 0$, regard ζ and η as matrices indexed by Par(n), and let A be the matrix defined by $A_{\rho\nu} = \langle u_\rho, v_\nu \rangle$. Then (9.24) is equivalent to $I = \zeta A \eta^t$, where ^{t} denotes the transpose and I is the identity matrix. Therefore

$$\begin{aligned} \{u_\lambda\} \text{ and } \{v_\lambda\} \text{ are dual bases } &\iff A = I \\ &\iff I = \zeta\eta^t \\ &\iff I = \zeta^t\eta \\ &\iff \delta_{\rho\nu} = \sum_\lambda \zeta_{\lambda\rho}\eta_{\lambda\nu} \qquad (9.25) \end{aligned}$$

Now by Proposition 9.7 we have

$$\begin{aligned}
\prod_{i,j}(1-x_iy_j)^{-1} &= \sum_{\lambda} m_\lambda(x)h_\lambda(y) \\
&= \sum_{\lambda}\left(\sum_{\rho}\zeta_{\lambda\rho}u_\rho(x)\right)\left(\sum_{\nu}\eta_{\lambda\nu}v_\nu(y)\right) \\
&= \sum_{\rho,\nu}\left(\sum_{\lambda}\zeta_{\lambda\rho}\eta_{\lambda\nu}\right)u_\rho(x)v_\nu(y)
\end{aligned}$$

Since the power series $u_\rho(x)v_\nu(y)$ are linearly independent over $\mathbb{Q}$, the proof follows from (9.25). $\square$

Proposition 9.18. *We have*

$$\langle p_\lambda, p_\mu \rangle = z_\lambda \delta_{\lambda\mu} \tag{9.26}$$

Hence the p_λ's form an orthogonal basis of Λ. (They don't form an orthonormal basis, since $\langle p_\lambda, p_\lambda \rangle \neq 1$)

Proof. By Proposition 9.13 and Lemma 9.17 we see that $\{p_\lambda\}$ and $\{p_\mu / z_\mu\}$ are dual bases, which is equivalent to (9.26). □

Corollary 9.19. *The scalar product* $\langle \cdot, \cdot \rangle$ *is positive definite i.e.* $\langle f, f \rangle \geq 0$ *for all* $f \in \Lambda$*, with equality if and only if* $f = 0$*.*

Proof. Write (uniquely) $f = \sum_\lambda c_\lambda p_\lambda$. Then

$$\langle f, f \rangle = \sum c_\lambda^2 z_\lambda$$

The proof follows since each $z_\lambda > 0$. □

Proposition 9.20. *The involution* ω *is an isometry, i.e.* $\langle \omega f, \omega g \rangle = \langle f, g \rangle$ *for all* $f, g \in \Lambda$*.*

Proof. By the bilinearity of the scalar product, it suffices to take $f = p_\lambda$ and $g = p_\mu$. The result then follows from Propositions 9.14 and 9.18. □

9.8 The Combinatorial Definition of Schur Functions

The fundamental combinatorial objects associated with Schur functions are semistandard tableaux. Let λ be a partition. A *semistandard (Young) tableaux* (SSYT) of *shape* λ is an array $T = (T_{ij})$ of positive integers of shape λ (i.e., $1 \le i \le l(\lambda)$, $1 \le j \le \lambda_i$) that is weakly increasing in every row and strictly increasing in every column. The *size* of an SSYT is its number of entries.

If T is an SSYT of shape λ then we write $\lambda = \text{sh}(T)$. Hence the size of T is just $|\text{sh}(T)|$. We may also think of an SSYT of shape λ as the Ferrers diagram of λ whose boxes have been filled with positive integers (satisfying certain conditions).

We say that T has *type* $\alpha = (\alpha_1, \alpha_2, \ldots)$, denoted $\alpha = \text{type}(T)$, if T has $\alpha_i = \alpha_i(T)$ parts equal to i. For any SSYT T of type α (or indeed for any multiset on $\mathbb{P}$ with possible additional structure), write

$$x^T = x_1^{\alpha_1(T)} x_2^{\alpha_2(T)} \cdots$$

There is a generalization of SSYTs of shape λ that fits naturally into the theory of symmetric functions. If λ and μ are partitions with $\mu \subseteq \lambda$ (i.e. $\mu_i \leq \lambda_i$ for all i), then define a *semistandard tableau* of *(skew) shape* λ/μ to be an array $T = (T_{ij})$ of positive integers of shape λ/μ (i.e. $1 \leq i \leq l(\lambda)$, $\mu_i < j \leq \lambda_i$) that is weakly increasing in every row and strictly increasing in every column.

We can similarly extend the definition of a Ferrers diagram of shape λ to one of shape λ/μ. Thus an SSYT of shape λ/μ may be regarded as a Ferrers diagram of shape λ/μ whose boxes have been filled with positive integers (satisfying certain conditions), just for “ordinary shapes” λ.

The definitions of type(T) and x^T carry over directly from SSYTs T of ordinary shape to those of skew shape.

Definition 9.21. Let λ/μ be a skew shape. The *skew Schur function* $s_{\lambda/\mu} = s_{\lambda/\mu}(x)$ of *shape* λ/μ in the variables $x = (x_1, x_2, \ldots)$ is the formal power series

$$s_{\lambda/\mu}(x) = \sum_T x^T$$

summer over all SSYTs T of shape λ/μ. If $\mu = \emptyset$, then we call $s_\lambda(x)$ the *Schur function* of shape λ.

Theorem 9.22. *For any skew shape λ/μ, the skew Schur function $s_{\lambda/\mu}$ is a symmetric function.*

Proof. It suffices to show that $s_{\lambda/\mu}$ is invariant under interchanging x_i and x_{i+1}. Suppose that $|\lambda/\mu| = n$ and that $\alpha = (\alpha_1, \alpha_2, \ldots)$ is a weak composition of n. Let

$$\tilde{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \alpha_{i+1}, \alpha_i, \alpha_{i+2}, \ldots).$$

If $\mathcal{T}_{\lambda/\mu,\alpha}$ denotes the set of all SSYTs of shape λ/μ and type α, then we seek the bijection $\varphi : \mathcal{T}_{\lambda/\mu,\alpha} \to \mathcal{T}_{\lambda/\mu,\tilde{\alpha}}$.

Let $T \in \mathcal{T}_{\lambda/\mu,\alpha}$. Consider the parts of T equal to i or $i+1$. Some columns of T will contain no such parts, while some others will contain two such parts, viz., one i and $i+1$. These columns we ignore. The remaining parts equal to i or $i+1$ occur once in each column, and consist of rows with a certain number r of i's followed by a certain number s of $i+1$'s. (Of course r and s depend on the row in question.) For example, a portion of T could loook as follows:

$$\begin{array}{ccccc}
 & & & & i \\
i & i & \underbrace{i \quad i}_{r=2} & \underbrace{i+1 \quad i+1 \quad i+1 \quad i+1}_{s=4} & i+1 \\
i+1 & i+1 & & &
\end{array}$$

In each such row convert the r i's and s $i+1$'s to s i's and r $i+1$'s

$$\begin{array}{ccccccc} & & & & & & i \\ i & & i & \underbrace{i \quad i \quad i \quad i}_{s=4} & & \underbrace{i+1 \quad i+1}_{r=2} & i+1 \\ i+1 & & i+1 & & & & \end{array}$$

It's easy to see that the resulting array $\varphi(T)$ belongs to $\mathcal{T}_{\lambda/\mu,\tilde{\alpha}}$, and that φ establishes the desired bijection.

$\square$

If $\lambda \vdash n$ and α is a weak composition of n, then let $K_{\lambda\alpha}$ denote the number of SSYTs of shape λ and type α. $K_{\lambda\alpha}$ is called a *Kostka number*. By Definition 9.21 we have

$$s_\lambda = \sum_\alpha K_{\lambda\alpha} x^\alpha$$

summed over all weak compositions α of n, so by Theorem 9.22 we have

$$s_\lambda = \sum_{\mu \vdash n} K_{\lambda\mu} m_\mu \tag{9.27}$$

More generally, we can define the *skew Kostka number* $K_{\lambda/\nu,\alpha}$ as the number of SSYTs of shape λ/ν and type α, so that if $|\lambda/\nu| = n$ then

$$s_{\lambda/\nu} = \sum_{\mu \vdash n} K_{\lambda/\nu,\mu} m_\mu \tag{9.28}$$

Consider the number $K_{\lambda,1^n}$, also denoted by f^λ. By definition, f^λ is the number of ways to insert the numbers $1, 2, \ldots, n$ into the shape $\lambda \vdash n$, each number appearing once, so that every row and column is increasing. Such an array is called a *standard Young tableau*(SYT) (or just *standard tableau*) of shape λ. The number f^λ has several alternative combinatorial interpretations as given by the following proposition.

Proposition 9.23. *Let $\lambda \in$ Par. Then the number f^λ counts the objects in items (a)-(e) below. We illustrate these objects with the case $\lambda = (3,2)$.*

(a) Chains of partitions. Saturated chains in the interval $[\emptyset, \lambda]$ of Young's lattice Y, or equivalently, sequences $\emptyset = \lambda^0, \lambda^1, \ldots, \lambda^n = \lambda$ of partitions (which we identify with their diagrams) such that λ^i is obtained from λ^{i-1} by adding a single square.

$\emptyset$	$\subset$	*1*	$\subset$	*2*	$\subset$	*3*	$\subset$	*31*	$\subset$	*32*
$\emptyset$	$\subset$	*1*	$\subset$	*2*	$\subset$	*21*	$\subset$	*31*	$\subset$	*32*
$\emptyset$	$\subset$	*1*	$\subset$	*2*	$\subset$	*21*	$\subset$	*22*	$\subset$	*32*
$\emptyset$	$\subset$	*1*	$\subset$	*11*	$\subset$	*21*	$\subset$	*31*	$\subset$	*32*
$\emptyset$	$\subset$	*1*	$\subset$	*11*	$\subset$	*21*	$\subset$	*22*	$\subset$	*32*

(b) Linear extensions. Let P_λ be the poset whose elements are the squares of the diagram of λ, with t covering s if t lies directly to the

right or directly below s(with no squares in between). Such posets are just the finite order ideals of $\mathbb{N} \times \mathbb{N}$. Then $f^{\lambda} = e(P_{\lambda})$, the number of linear extensions of P_{λ}

(c) Ballot sequences. Ways in which n voters can vote sequentially in an election for candidates $A_1, A_2, \ldots,$ so that for all i, A_i receives λ_i votes, and so that A_i never trails A_{i+1} in the voting. (We denote such a voting sequence as $a_1 a_2 \cdots a_n$, where the k-th voter votes for A_{a_k}.)

$$11122 \quad 11212 \quad 11221 \quad 12112 \quad 12121$$

(d) Lattice permutations. Sequences $a_1 a_2 \cdots a_n$ in which i occurs λ_i times, and such that in any left factor $a_1 a_2 \cdots a_j$, the number of i's is at least as great as the number of $i+1$'s (for all i). Such a sequence is called a lattice permutation(or Yamanouchi word or ballot sequence) of type λ.

$$11122 \quad 11212 \quad 11221 \quad 12112 \quad 12121$$

(e) Lattice paths. Lattice $0 = v_0, v_1, \ldots, v_n$ *in* $\mathbb{R}^l$ *(where* $l = l(\lambda)$*) from the orign* v_0 *to* $v_n = (\lambda_1, \lambda_2, \ldots, \lambda_l)$*, with each step a unit coordinate vector, and staying within the region (or cone)* $x_1 \geq x_2 \geq \cdots \geq x_l \geq 0$.

Define a *reverse* SSYT or *column-strict plane partition* of (skew) shape λ/μ to be an array of positive integers of shape λ/μ that is weakly decreasing in rows and strictly decreasing in columns. Define the type α of a reverse SSYT exactly as for ordinary SSYT.

Define $\hat{K}_{\lambda/\mu,\alpha}$ to be the number of reverse SSYTs of shape λ/μ and type α.

Proposition 9.24. *Let λ/μ be a skew partition of n, and let α be a weak composition of n. Then $\hat{K}_{\lambda/\mu,\alpha} = K_{\lambda/\mu,\alpha}$.*

Proof. Suppose that T is a reverse SSYT of shape λ and type $\alpha = (\alpha_1, \alpha_2, \ldots)$. Let k denote the largest part of T. The transformation $T_{ij} \mapsto k + 1 - T_{ij}$ shows that $\widehat{K}_{\lambda\alpha} = K_{\lambda\bar{\alpha}}$, where $\bar{\alpha} = (\alpha_k, \alpha_{k-1}, \ldots, \alpha_1, 0, 0, \ldots)$. But by Theorem 9.22 we have $K_{\lambda\bar{\alpha}} = K_{\lambda\alpha}$, and the proof is complete. □

Proposition 9.25. *Suppose that μ and λ are partitions with $|\mu| = |\lambda|$ and $K_{\lambda\mu} \neq 0$. Then $\lambda \trianglerighteq \mu$. Moreover $K_{\lambda\lambda} = 1$.*

Proof. Suppose $K_{\lambda\mu} \neq 0$. By definition, there exists an SSYT T of shape λ and type μ. Suppose that a part $T_{ij} = k$ appears below the k-th row (i.e. $i > k$). Then we have $1 \leq T_{1k} < T_{2k} < \cdots < T_{ik} = k$ for $i > k$, which is impossible. Hence the parts $1, 2, \ldots, k$ all apppear in the first k rows, so that $\mu_1 + \mu_2 + \cdots + \mu_k \leq \lambda_1 + \lambda_2 + \cdots + \lambda_k$, as desired. Moreover, if $\mu = \lambda$ then we must have $T_{ij} = i$ for all (i, j), so $K_{\lambda\lambda} = 1$. □

Corollary 9.26. *The Schur functions s_λ with $\lambda \in$ Par(n) form a basis for Λ^n, so $\{s_\lambda : \lambda \in \text{Par}\}$ is a basis for Λ. In fact, the transition matrix $K_{\lambda\mu}$ which expresses the s_λ's in terms of the m_μ's, with respect to any linear ordering of Par(n) that extends dominance order, is lower triangular with 1's on the main diagonal.*

Proof. Proposition 9.25 is equivalent to the assertion about $K_{\lambda\mu}$. Since a lower triangular matrix with 1's on the main diagonal is invertible, it follows that $\{s_\lambda : \lambda \in \text{Par}(n)\}$ is a $\mathbb{Q}$-basis for Λ^n. □

Corollary 9.27. *The Schur functions s_λ with $\lambda \in$ Par(n) form a basis for Λ^n, so $\{s_\lambda : \lambda \in \text{Par}\}$ is a basis for Λ. In fact, the transition matrix $K_{\lambda\mu}$ which expresses the s_λ's in terms of the m_μ's, with respect to any linear ordering of Par(n) that extends dominance order, is lower triangular with 1's on the main diagonal.*

Proof. Proposition 9.25 is equivalent to the assertion about $K_{\lambda\mu}$. Since a lower triangular matrix with 1's on the main diagonal is invertible, it follows that $\{s_\lambda : \lambda \in \text{Par}(n)\}$ is a $\mathbb{Q}$-basis for Λ^n. □

9.9 The RSK Algorithm

The basic operation of the RSK algorithm consists of the *row insertion* $P \leftarrow k$ of a positive integer k into a nonskew SSYT $P = (P_{ij})$. The operation $P \leftarrow k$ is defined as follows: Let r be the largest integer such that $P_{1,r-1} \leq k$. (If $P_{11} > k$ then let $r = 1$.) If P_{1r} doesn't exist (i.e. P has $r - 1$ columns), then simply place k at the end of the first row. The insertion process stops, and the resulting SSYT is $P \leftarrow k$. If on the other hand, P has at least r columns, so that P_{1r} exists, then replace P_{1r} by k. The element then "bumps" $P_{ir} := k'$ into the second row, i.e. insert k' into the second row of P by the insertion rule just described. Continue until an element is inserted at the end of a row (possibly as the first element of the next row). The resulting array is $P \leftarrow k$.

Lemma 9.28. *(a) When we insert k into an SSYT P, then the insertion path moves to the left. More precisely, if $(r,s),(r+1,t) \in I(P \leftarrow k)$ then $t \leq s$.*

(b) Let P be an SSYT, and let $j \leq k$. Then $I(P \leftarrow j)$ lies strictly to the left of $I((P \leftarrow j) \leftarrow k)$. More precisely, if $(r,s) \in I(P \leftarrow j)$ and $(r,t) \in I((P \leftarrow j) \leftarrow k)$, then $s < t$. Moreover, $I((P \leftarrow j) \leftarrow k)$ does not extend below the bottom of $I(P \leftarrow j)$. Equivalently

$$\#I((P \leftarrow j) \leftarrow k) \leq \#I(P \leftarrow j)$$

Proof. (a) Suppose that $(r, s) \in I(P \leftarrow k)$. Now either $P_{r+1,s} > P_{rs}$ (since P is strictly increasing in columns) or else there is no $(r+1, s)$ entry of P. In the first case, P_{rs} cannot get bumped to the right of column s without violating the fact that the rows of $P \leftarrow k$ are weakly increasing, since P_{rs} would be to the right of $P_{r+1,s}$ on the same row. The second case is clearly impossible, since we would otherwise have a gap in row $r+1$. Hence (a) is proved.

(b) Since a number can only bump a strictly larger number, it follows that k is inserted in the first row of $P \leftarrow j$ strictly to the right of j. Since the first row of P is weakly increasing, j bumps an element no larger than the element k bumps. Hence by induction $I(P \leftarrow j)$ lies strictly to the left of $I((P \leftarrow j) \leftarrow k)$.

The bottom element b of $I(P \leftarrow j)$ was inserted at the end of its row. By what was just proved, if $I((P \leftarrow j) \leftarrow k)$ has an element c in this row, then it lies to the right of b. Hence c was inserted at the end of the row, so the insertion procedure terminates. It follows that $I((P \leftarrow j) \leftarrow k)$ can never go below the bottom of $I(P \leftarrow j)$. $\square$

Corollary 9.29. *If P is an SSYT and $k \geq 1$, then $P \leftarrow k$ is also an SSYT.*

Proof. It is clear that the rows of $P \leftarrow k$ are weakly increasing. Now a number a can only bump a larger number b. By Lemma 9.28(a), b does not move to the right when when it is bumped. Hence b is inserted below a number that is strictly smaller than b, so $P \leftarrow k$ remains an SSYT. □

Now let $A = (a_{ij})$ be a $\mathbb{N}$-matrix with finitely many nonzero entries. We will say that A is an $\mathbb{N}$-matrix of *finite support*. We can think of A as either an infinite matrix or as an $m \times n$ matrix when $a_{ij} = 0$ for $i > m$ and $j > n$.

Associate with A a *generalized permutation* of *two-line array* w_A defined by

$$w_A = \begin{pmatrix} i_1 & i_2 & i_3 & \dots & i_m \\ j_1 & j_2 & j_3 & \dots & j_m \end{pmatrix} \tag{9.29}$$

where (a) $i_1 \leq i_2 \leq \cdots \leq i_m$

(b) if $i_r = i_s$, and $r \leq s$, then $j_r \leq j_s$,

(c) for each pair (i, j), there are exactly a_{ij} values of r for which $(i_r, j_r) = (i, j)$

It is easily seen that A determines a unique two line array w_A satisfying $(a) - (c)$, and conversely any such array corresponds to a unique A.

We now associate with A (or w_A) a pair (P, Q) of SSYTs of the same shape, as follows. Let w_A be given by (9.29). Begin with $(P(0), Q(0)) = (\emptyset, \emptyset)$ (where $\emptyset$ denotes the empty SSYT). If $t < m$ and $(P(t), Q(t))$ are defined, then let

(a) $P(t+1) = P(t) \leftarrow j_{t+1}$;

(b) $Q(t+1)$ be obtained from $Q(t)$ by inserting i_{t+1} (leaving all parts of $Q(t)$ unchanged) so that $P(t+1)$ and $Q(t+1)$ have the same shape.

The process ends at $(P(m), Q(m))$, and we define $(P, Q) = (P(m), Q(m))$. We denote this correspondence by $A \xrightarrow{A} (P, Q)$ and call it the *RSK algorithm*. We call P the *insertion tableau* and Q the *recording tableau* of A or of w_A

Theorem 9.30. *The RSK algorithm is a bijection between* $\mathbb{N}$*-matrices* $A = (a_{ij})_{i,j \geq 1}$ *of finite support and ordered pairs* (P, Q) *of SSYT of the same shape. In this correspondence,*

$$j \text{ occurs in } P \text{ exactly } \sum_i a_{ij} \text{ times} \tag{9.30}$$

$$i \text{ occurs in } Q \text{ exactly } \sum_j a_{ij} \text{ times} \tag{9.31}$$

(These last two conditions are equivalent to type$(P) =$ *col*(A)*, type*$(Q) =$ *row*(A)*).*

Proof. By Corollary 9.29, P is an SSYT. Clearly, by definition of the RSK algorithm P and Q have the same shape, and also (9.30) and (9.31) hold. Thus we must show the following: (a) Q is an SSYT , and (b) the RSK algorithm is a bijection, i.e., given (P, Q) , one can uniquely recover A.

To prove (a), first note that since the elements of Q are inserted in weakly increasing order, it follows that the rows and columns of Q are weakly increasing. Thus we must show that the columns of Q are strictly increasing, i.e. no two equal elements of the top row of w_A can end up in the same column of Q. But if $i_k = i_{k+1}$ in the top row, then we must $j_k \leq j_{k+1}$. Hence by Lemma 9.28(b), the insertion path of j_{k+1} will always lie strictly to the right of the path for j_k, and will never extend below the bottom of j_k's insertion path. It follows that the bottom elements of the two insertion paths lie in different columns, so the columns of the Q are strictly increasing as desired.

The above argument establishes an important property of the RSK algorithm: *Equal elements of Q are inserted strictly left to right.*

It remains to show that the RSK algorithm is a bijection. Thus given $(P, Q) = (P(m), Q(m))$, let Q_{rs} be the rightmost occurrence of the largest entry of Q (where Q_{rs} is the element of Q in row r and column s). Since equal elements of Q are inserted left to right, it follows that $Q_{rs} = i_m$, $Q(m-1) = Q(m) \setminus Q_{rs}$ (i.e., $Q(m)$ with the element Q_{rs} deleted), and that P_{rs} was the last element of P to be bumped into place after inserting j_m into $P(m-1)$. But it is then easy to reverse the insertion procedure $P(m-1) \leftarrow j_m$.

P_{rs} must have been bumped by the rightmost element $P_{r-1,t}$ of row $r-1$ of P that is smaller than P_{rs}. Hence remove P_{rs} from P, replace $P_{r-1,t}$ with P_{rs}, and continue by replacing the rightmost element of row $r-2$ of P that is smaller than $P_{r-1,t}$ with $P_{r-1,t}$,etc. Eventually some element j_m is removed from the first row of P. We have thus uniquely recovered (i_m, j_m) and $(P(m-1), Q(m-1))$. By iterating this procedure we recover the entire two-line array w_A. Hence the RSK algorithm is injective.

To show surjectivity, we need to show that applying the procedure of the previous paragraph to an arbitrary pair (P, Q) of SSYTs of the same shape always yields a valid two-line array

$$w_A = \begin{pmatrix} i_1 & i_2 & i_3 & \dots & i_m \\ j_1 & j_2 & j_3 & \dots & j_m \end{pmatrix} \qquad (9.32)$$

Clearly, $i_1 \le i_2 \le \cdots \le i_m$, so we need to show that if $i_k = i_{k+1}$ then $j_k \le j_{k+1}$. Let $i_k = Q_{rs}$ and $i_{k+1} = Q_{uv}$, so $r \ge u$ and $s < v$. When we begin to apply inverse bumping to P_{uv}, it occupies the end of its row (row u).

Hence when we apply inverse bumping to P_{rs}, its "inverse insertion path" intersects row u strictly to the left of the column v. Thus at row u the inverse insertion path of P_{rs} lies strictly to the left of that of P_{uv}. By a simple induction argument (essentially the "inverse" of Lemma 9.28(b)), the entire inverse insertion path of P_{rs} lies strictly to the left of that of P_{uv}. In particular, before removing i_{k+1} the two elements j_k and j_{k+1} appear in the first row with j_k to the left j_{k+1}. Hence $j_k \leq j_{k+1}$ as desired, completing the proof. □

When the RSK algorithm is applied to a permutation matrix A (or a permutation $w \in \mathcal{S}_n$), the resulting tableaux P, Q are just standard Young tableaux (of the same shape). Conversely, if P and Q are SYTs of the same shape, then the matrix A satisfying $A \overset{RSK}{\rightarrow} (P, Q)$ is a permutation matrix. Hence the RSK algorithm sets up a bijection between the symmetric group $\mathcal{S}_n$ and pairs (P, Q) of SYTs of the same shape $\lambda \vdash n$. In particular, if f^λ denotes the number of SYTs of shape λ, then we have the fundamental identity

$$\sum_{\lambda \vdash n} (f^\lambda)^2 = n! \tag{9.33}$$

Although permutation matrices are very special cases $\mathbb{N}$-matrices of finite support, in fact the RSK algorithm for arbitrary $\mathbb{N}$-matrices A can be reduced to the case of permutation matrices. Namely, given the two line array w_A, say of length n, replace the first row by $1, 2, \ldots, n$. Suppose the second row of w_A has c_i i's. Then replace the 1's in the second row from left-to-right with $1, 2, \ldots, c_1$, next the 2's from left to-right with $c_1 + 1, c_2 + 1, \ldots, c_1 + c_2$ etc. until the second row becomes a permutation of $1, 2, \ldots, n$. Denote the resulting two-line array by $\tilde{w}_A$.

Lemma 9.31. *Let*

$$w_A = \begin{pmatrix} i_1 & i_2 & i_3 & \dots & i_n \\ j_1 & j_2 & j_3 & \dots & j_n \end{pmatrix}$$

be a two-line array, and let

$$\tilde{w}_A = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \tilde{j}_1 & \tilde{j}_2 & \tilde{j}_3 & \dots & \tilde{j}_n \end{pmatrix}$$

Suppose that $\tilde{w}_A \stackrel{RSK}{\rightarrow} (\tilde{P}, \tilde{Q})$. Let (P, Q) be the tableaux obtained from $\tilde{P}$ and $\tilde{Q}$ by replacing k in $\tilde{P}$ by i_k, and $\tilde{j}_k$ in $\tilde{Q}$ by j_k. Then $w_A \stackrel{RSK}{\rightarrow} (P, Q)$. In other words, the operation $w_A \mapsto \tilde{w}_A$ "commutes" with the RSK algorithm.

Proof. Suppose that when the number j is inserted into a row at some stage of the RSK algorithm, it occupies the k-th position in the row. If this number j were replaced by a larger number $j + \epsilon$, smaller than any element of the row which is greater than j, then $j + \epsilon$ would also be inserted in at the k-th position. From this we see that the insertion procedure for elements $j_1, j_2, \ldots, j_n$ exactly mimics that for $\tilde{j}_1, \tilde{j}_2, \ldots \tilde{j}_n$, and the proof follows. □

The process of replacing w_A with $\tilde{w}_A$, P with $\tilde{P}$, etc is called *standardization*

9.10 Some consequences of the RSK algorithm

Theorem 9.32 (Cauchy identity). *We have*

$$\prod_{i,j}(1 - x_i y_j)^{-1} = \sum_{\lambda} s_\lambda(x) s_\lambda(y) \tag{9.34}$$

Proof. Write

$$\prod_{i,j}(1 - x_i y_j)^{-1} = \prod_{i,j}\left[\sum_{a_{ij}\geq 0}(x_i y_j)^{a_{ij}}\right] \tag{9.35}$$

A term $x^\alpha y^\beta$ in this expansion is obtained by choosing an $\mathbb{N}$-matrix $A^t = (a_{ij})^t$ (the transpose of A) of finite support with row$(A) = \alpha$ and col$(A) = \beta$.

Hence the coefficient of $x^\alpha y^\beta$ in (9.35) is the number of $N_{\alpha\beta}$ of $\mathbb{N}$-matrices A with row$(A) = \alpha$ and col$(A) = \beta$. This statement is also equivalent to (9.6). On the other hand the coefficient of $x^\alpha y^\beta$ in $\sum_\lambda s_\lambda(x)s_\lambda(y)$ is the number of pairs (P, Q) of SSYT of the shape λ such that type$(P) = \alpha$ and type$(Q) = \beta$. The RSK algorithm sets up a bijection between the matrices A and the tableau pairs (P, Q), so the proof follows. $\square$

Corollary 9.33. *The Schur functions form an orthonormal basis for* Λ*, i.e.* $\langle s_\lambda, s_\mu \rangle = \delta_{\lambda\mu}$

Proof. Combine Corollary 9.27 and Lemma 9.17. □

Corollary 9.34. *Fix partitions $\mu, \nu \vdash n$. Then*

$$\sum_{\lambda \vdash n} K_{\lambda\mu} K_{\lambda\nu} = N_{\mu\nu} = \langle h_\mu, h_\nu \rangle$$

where $K_{\lambda\mu}$ and $K_{\lambda\nu}$ denote Kostka numbers, and $N_{\mu\nu}$ is the number $\mathbb{N}$-matrices A with row$(A) = \mu$ and col$(A) = \nu$.

Proof. Take the coefficient of $x^\mu y^\nu$ on both sides of (9.34). □

Corollary 9.35. *We have*

$$h_\mu = \sum_\lambda K_{\lambda\mu} s_\lambda \tag{9.36}$$

In other words, if $M(u,v)$ denotes the transition matrix from the basis $\{v_\lambda\}$ to the basis $\{u_\lambda\}$ of Λ (so that $u_\lambda = \sum_\mu M(u,v)_{\lambda\mu} v_\mu$), then

$$M(h,s) = M(s,m)^t$$

We give three proofs of this corollary, all essentially equivalent

First proof. Let $h_\mu = \sum_\lambda a_{\lambda\mu} s_\lambda$. By Corollary 9.33, we have $a_{\lambda\mu} = \langle h_\mu, s_\lambda \rangle$. Since $\langle h_\mu, m_\nu \rangle = \delta_{\mu\nu}$ by definition (9.22) of the scalar product $\langle , \rangle$, we have from Equation (9.27) that $\langle h_\mu, hs_\lambda \rangle = K_{\lambda\mu}$. □

Second Proof. Fix μ. Then

$$\begin{aligned} h_\mu &= \sum_A x^{\mathsf{col}(A)} \\ &= \sum_{(P,Q)} x^Q && \text{by the RSK algorithm} \\ &= \sum_\lambda K_{\lambda\mu} \sum_Q x^Q \\ &= \sum_\lambda K_{\lambda\mu} s_\lambda \end{aligned}$$

where (i) A ranges over all $\mathbb{N}$-matrices with row$(A) = \mu$

(ii) (P, Q) ranges over all pairs of SSYT of the same shape with type$(P) = \mu$ and

(iii) Q ranges over all SSYT of shape λ. □

Third proof Take the coefficient of $m_\mu(x)$ on both sides of the identity

$$\sum_\lambda m_\lambda(x) h_\lambda(y) = \sum_\lambda s_\lambda(x) s_\lambda(y)$$

The two sides are equal by (9.7) and (9.34) □

Corollary 9.36. *We have*

$$h_1^n = \sum_{\lambda \vdash n} f^\lambda s_\lambda \tag{9.37}$$

Proof. Take the coefficients of $x_1 x_2 \ldots x_n$ on both sides of (9.34). To obtain a bijective proof, consider the RSK algorithm $A \stackrel{RSK}{\longrightarrow} (P, Q)$ when col$(A) = \langle 1^n \rangle$.

9.11 Symmetry of the RSK algorithm

Theorem 9.37. *Let A be an $\mathbb{N}$-matrix of finite support, and suppose that $A \stackrel{RSK}{\rightarrow} (P, Q)$. Then $A^t \stackrel{RSK}{\rightarrow} (Q, P)$, where ^{t} denotes the transpose.*

Proof. Let $w_A = \binom{u}{v}$ be the two-line array associated to A. Hence $w'_A = \binom{v}{u}_{\text{sorted}}$ i.e., sort the columns of $\binom{v}{u}$ so that the columns are weakly increasing in lexicographic order. It follows from Lemma 9.31 that we may assume u and v have no repeated elements.

Consider

$$w_A = \left(\begin{array}{ccc} u_1 & \dots & u_n \\ v_1 & \dots & v_n \end{array} \right) = \left(\begin{array}{c} u \\ v \end{array} \right)$$

Suppose the u_i's and v_j's are distinct, define the *inversion poset* $I = I(A) = I(\binom{u}{v})$ as follows. The vertices of I are the columns of $\binom{u}{v}$. For notational convenience, we denote a column $\begin{smallmatrix} a \\ b \end{smallmatrix}$ as ab. Define $ab < cd$ in I if $a < c$ and $b < d$.

Lemma 9.38. *The map* $\varphi : I(A) \longrightarrow I(A^t)$ *defined by* $\varphi(ab) = \varphi(ba)$ *is an isomorphism of posets.*

Now given the inversion poset $I = I(A)$, define I_1 to be the set of minimal elements of I, then I_2 to be the set of minimal elements of $I - I_1$, then I_3 to be the set of minimal elements of $I - I_1 - I_2$ etc. Note that since I_i is an antichain of I, its elements can be labeled

$$(u_{i1}, v_{i1}), (u_{i2}, v_{i2}), \ldots, (u_{in_i}, v_{in_i}) \tag{9.38}$$

where $n_i = \#I_i$ such that

$$\begin{array}{cccc} u_{i1} < & u_{i2} < & \ldots & < u_{in_i} \\ v_{i1} > & v_{i2} > & \ldots & > v_{in_i} \end{array} \tag{9.39}$$

Lemma 9.39. *Let $I_1, \ldots, I_d$ be the (nonempty) antichains defined above, labeled as in (9.39). Let $A \stackrel{RSK}{\rightarrow} (P, Q)$. Then the first row of P is $v_{1n_1} v_{2n_2} \cdots v_{dn_d}$, while the first row of Q is $u_{11} u_{21} \cdots u_{d1}$. Moreover, if $(u_k, v_k) \in I_i$, then v_k is inserted into the i-th column of the first row of the tableau $P(k-1)$ in the RSK algorithm.*

Proof. Induction on n, the case $n = 1$ being trivial. Assume the assertion for $n - 1$, and let

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \\ v_1 & v_2 & \cdots & v_n \end{pmatrix}, \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & \cdots & u_{n-1} \\ v_1 & v_2 & \cdots & v_{n-1} \end{pmatrix}$$

Let $P(n-1), Q(n-1)$ be the tableaux obtained after inserting $v_1, \ldots, v_{n-1}$, and let the antichains $I'_i := I_i(\binom{\tilde{u}}{\tilde{v}})$, $1 \leq i \leq e$ (where $e = d-1$ or $e = d$), be given by $(\tilde{u}_{i1}, \tilde{v}_{i1}), \ldots, (\tilde{u}_{im_i}, \tilde{v}_{im_i})$ where $\tilde{u}_{i1} < \cdots < \tilde{u}_{im_i}$ and $\tilde{v}_{i1} > \cdots \tilde{v}_{im_i}$. By the induction hypothesis, the first row of $P(n-1)$ is $\tilde{v}_{1m_1}\tilde{v}_{2m_2}\cdots\tilde{v}_{em_e}$, while the first row of Q is $\tilde{u}_{11}\tilde{u}_{21}\cdots\tilde{u}_{e1}$. Now we insert v_n into $P(n-1)$. If $\tilde{v}_{im_i} > v_n$, then $I'_i \cup (u_n, v_n)$ is an antichain of $I(\binom{u}{v})$. Hence $(u_n, v_n) \in I_i(\binom{u}{v})$ if i is the *least* index for which $\tilde{v}_{im_i} > v_n$. If there is no such i, then (u_n, v_n) is the unique element of the antichain $I_d(\binom{u}{v})$ of $I(\binom{u}{v})$. These conditions mean that v_n is inserted into the i-th column of $P(n-1)$, as claimed. We start a new i-th column exactly when $v_n = v_{d1}$, in which case $u_n = u_{d1}$, so u_n is inserted into the i-th column of the first row of $Q(n-1)$, as desired.

Proof. **of Theorem 9.37** If the antichain $I_i(\binom{u}{v})$ is given by 9.38 such that (264) is satisfied, then by Lemma 9.38 the antichain $I_i(\binom{v}{u})$ is just

$$(v_{im_i}, u_{im_i}), \ldots (v_{i2}, u_{i2}), (v_{i1}, u_{i1},)$$

where

$$\begin{aligned} v_{im_i} < \quad \ldots \quad < v_{i2} \quad < v_{i1} \\ u_{im_i} > \quad \ldots \quad u_{i2} > \quad > u_{i1} \end{aligned}$$

Hence by Lemma 9.39, if $A^t \stackrel{RSK}{\rightarrow} (P', Q')$, then the first row of P' is $u_{11}u_{21}\cdots u_{d1}$, and the first row of Q' is $v_{im_1}v_{2m_2}\cdots v_{dm_d}$. Thus by Lemma 9.39, the first rows of P' and Q' agree with the first rows of P and Q, respectively.

When the RSK algorithm is applied to $\binom{u}{v}$, the element $v_{ij}, 1 \leq j < m_i$, gets bumped into the second row of P *before* the element $1 \leq s < m_r$, if and only if $u_{i,j+1} < u_{r,s+1}$. Let $\bar{P}$ and $\bar{Q}$ denote P and Q with their first rows removed. It follows that

$$\begin{pmatrix} a \\ b \end{pmatrix} := \begin{pmatrix} u_{12}\cdots & u_{1m_1} & u_{22}\cdots & u_{2m_2} & \cdots u_{d2} & \cdots u_{dm_d} \\ v_{11}\cdots & v_{1m_1-1} & v_{21}\cdots & v_{2m_2-1} & \cdots v_{d1} & \cdots v_{dm_d-1} \end{pmatrix}_{\text{sorted}}$$

$$\stackrel{RSK}{\to} (\bar{P}, \bar{Q})$$

Similarly let $(\bar{P}', \bar{Q}')$ denote P' and Q' with their first rows removed. Applying the same argument to $\binom{v}{u}$ rather than $\binom{u}{v}$ yields

$$\begin{pmatrix} a' \\ b' \end{pmatrix} := \begin{pmatrix} v_{1m_1-1} \cdots & v_{11} & v_{2m_2-1} \cdots & v_{21} & \cdots v_{dm_d-1} & \cdots v_{d1} \\ u_{1m_1} \cdots & u_{12} & u_{2m_2} \cdots & u_{22} & \cdots u_{dm_d} & \cdots u_{d2} \end{pmatrix}_{\text{sorted}}$$

$$\stackrel{RSK}{\longrightarrow} (\bar{P}', \bar{Q}')$$

But $\binom{a}{b} = \binom{b'}{a'}_{\text{sorted}}$, so by induction on n (or on the number of rows) we have $(\bar{P}', \bar{Q}') = (\bar{Q}, \bar{P})$ and the proof follows. $\square$

Corollary 9.40. *Let A be a $\mathbb{N}$-matrix of finite support, and let $A \stackrel{RSK}{\longrightarrow} (P,Q)$. Then $A \stackrel{RSK}{\longrightarrow} (P,Q)$. Then A is symmetric i.e. $(A = A^t)$ if and only if $P = Q$.*

Proof. Immediete from the fact that $A^t \stackrel{RSK}{\longrightarrow} (Q,P)$. □

Corollary 9.41. *Let* $A = A^t$ *and* $A \overset{RSK}{\rightarrow} (P, P)$*, and let* $\alpha = (\alpha_1, \alpha_2, \dots)$ *where* $\alpha_i \in \mathbb{N}$ *and* $\sum \alpha_i < \infty$*. Then the map* $A \mapsto P$ *establishes a bijection between symmetric* $\mathbb{N}$*-matrices with row*$(A) = \alpha$ *and SSYTs of type* α*.*

Proof. Follows from Corollary 9.40 and Theorem 9.30.

Corollary 9.42. *We have*

$$\frac{1}{\prod_i(1-x_i)\cdot\prod_{i<j}(1-x_ix_j)} = \sum_{\lambda} s_\lambda(x) \qquad (9.40)$$

summed over all $\lambda \in$ Par.

Proof. The coefficient of x^α on the left side is the number of symmetric $\mathbb{N}$-matrices A with row$(A) = \alpha$ while the coefficient of x^α on the right hand side is the number of SSYTs of type α. Now apply Corollary 9.41. □

Corollary 9.43. *We have*

$$\sum_{\lambda \vdash n} f^{\lambda} = \#\{w \in \mathcal{S}_n : w^2 = 1\}$$

the number of involutions in $\mathcal{S}_n$*.*

Proof. Let $w \in \mathcal{S}_n$ and $w \overset{RSK}{\to} (P, Q)$ where P and Q are SYT of the same shape $\lambda \vdash n$. The permutation matrix corresponding to w is symmetric if and only if $w^2 = 1$. By Theorem 9.37 this is the case if and only if $P = Q$, and the proof follows.

Alternatively, take the coefficient of $x_1, \cdots x_n$ on both sides of (9.40) □

9.12 The dual RSK Algorithm

There is a variation of the RSK algorithm that is related to the product $\prod(1+x_iy_j)$ in the same way that the RSK algorithm itself is related to $\prod(1-x_iy_j)^{-1}$. We call this variation the *dual RSK algorithm* and denote it by $A \stackrel{RSK^*}{\rightarrow} (P,Q)$. The matrix A will now be a $(0,1)$ matrix of finite support. Form the two line array w_A just as before. The RSK^* algorithm proceeds exactly like the RSK algorithm, except that an element i bumps the leftmost element $\geq i$, rather than the leftmost element $> i$. (In particular, RSK and RSK* agree for permutation matrices.) It follows that each row of P is *strictly* increasing.

Theorem 9.44. *The RSK* algorithm is a bijection between $(0,1)$-matrices A of finite support and pairs (P,Q) such that P^t (the transpose of P) and Q are SSYTs with sh$(P) =$ sh(Q). Moreover, col$(A) =$ type(P) and row$(A) =$ type(Q).*

Theorem 9.45. *We have*

$$\prod_{i,j}(1 + x_i y_j) = \sum_{\lambda} s_{\lambda'}(x) s_{\lambda}(y)$$

Lemma 9.46. *Let ω_y denote ω acting on the y variables only (so we regard the x_i's as constants commuting with ω). Then*

$$\omega_y \prod (1 - x_i y_j)^{-1} = \prod (1 + x_i y_j)$$

Proof. We have

$$\begin{aligned}
\omega_y \prod (1 - x_i y_j)^{-1} &= \omega_y \sum_\lambda m_\lambda(x) h_\lambda(y) \quad (\text{ by Proposition 9.7}) \\
&= \sum_\lambda m_\lambda(x) e_\lambda(y) \quad (\text{ by Theorem 9.8 }) \\
&= \prod (1 + x_i y_j) \quad (\text{ by Proposition 9.3 })
\end{aligned}$$

Theorem 9.47. *For every $\lambda \in$ Par we have*

$$\omega s_\lambda = s_{\lambda'}$$

Proof. We have

$$\begin{aligned}
\sum_\lambda s_\lambda(x) s_{\lambda'}(y) &= \prod (1 + x_i y_j) \quad (\text{ by Theorem } 9.45) \\
&= \omega_y \prod (1 - x_i y_j)^{-1} \quad (\text{ by Lemma } 9.46) \\
&= \omega_y \sum_\lambda s_\lambda(x) s_\lambda(y) \quad (\text{ by Theorem } 9.32) \\
&= \sum_\lambda s_\lambda(x) \omega_y(s_\lambda(y))
\end{aligned}$$

Take the coefficient of $s_\lambda(x)$ on both sides. Since the $s_\lambda(x)$'s are linearly independent, we obtain $s_{\lambda'}(y) = \omega_y(s_\lambda(y))$, or just $s_{\lambda'} = \omega s_\lambda$. □

9.13 The Classical definition of the Schur functions

Let $\alpha = (\alpha_1, \alpha_2, \ldots \alpha_n) \in \mathbb{N}^n$ and $w \in \mathcal{S}_n$. As usual write $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and define

$$w(x^\alpha) = x_1^{\alpha_{w(1)}} \cdots x_n^{\alpha_{w(n)}}$$

Now define

$$a_\alpha = a_\alpha(x_1, \ldots, x_n) = \sum_{w \in \mathcal{S}_n} \varepsilon_w w(x^\alpha) \tag{9.41}$$

where

$$\varepsilon_w = \begin{cases} 1 & \text{if } w \text{ is an even permutation} \\ -1 & \text{if } w \text{ is an odd permutation} \end{cases}$$

Note that the right-hand side of equation (9.41) is just the expansion of a determinant, namely

$$a_\alpha = \det(x_i^{\alpha_j})_{i,j=1}^n$$

Note also that a_α is skew-symmetric, i.e. $w(a_\alpha) = \varepsilon_w a_\alpha$, so $a_\alpha = 0$ unless all the α_i's are distinct. Hence assume that $\alpha_1 > \alpha_2 > \cdots > \alpha_n \geq 0$, so $\alpha = \lambda + \delta$, where $\lambda \in \text{Par}$, $l(\lambda) \leq n$, and $\delta = \delta_n = (n-1, n-2, \ldots, 0)$. Since $\alpha_j = \lambda_j + n - j$, we get

$$a_\alpha = a_{\lambda+\delta} = \det\left(x_i^{\lambda_j+n-j}\right)_{i,j=1}^n \tag{9.42}$$

For instance,

$$a_{421} = a_{211+210} = \begin{vmatrix} x_1^4 & x_1^2 & x_1^1 \\ x_2^4 & x_2^2 & x_2^1 \\ x_3^4 & x_3^2 & x_3^1 \end{vmatrix}$$

Note in particular that

$$a_\delta = \mathsf{det}(x_i^{n-j}) = \prod_{1 \leq i < j \leq n} (x_i - x_j) \qquad (9.43)$$

the *Vandermonde determinant*. If for some $i \neq j$ we put $x_i = x_j$ in a_α, then because a_α is skew-symmetric (or because the i-th row and j-th row of the determinant (9.42) become equal), we obtain 0.

Hence a_α is divisible by $x_i - x_j$ and thus by a_δ (in the ring $\mathbb{Z}[x_1, \ldots x_n]$). Thus $a_\alpha / a_\delta \in \mathbb{Z}[x_1, \ldots, x_n]$. Moreover, since a_α and a_δ are skew-symmetric, the quotient is symmetric, and is clearly homogeneous of degree $|\alpha| - |\delta| = |\lambda|$. In other words, $a_\alpha / a_\delta \in \Lambda_n^{|\lambda|}$.

Theorem 9.48. *We have*

$$a_{\lambda+\delta}/a_{\delta} = s_{\lambda}(x_1, \ldots, x_n)$$

Proof. There are many proofs of this result. We give one that can be extended to give an important result on skew Schur functions(Theorem 9.51).

Applying ω to 9.36 and replacing λ by λ' yields

$$e_\mu = \sum_\lambda K_{\lambda'\mu} s_\lambda$$

Since the matrix $(K_{\lambda'\mu})$ is invertible, it suffices to show that

$$e_\mu(x_1, \ldots, x_n) = \sum K_{\lambda'\mu} \frac{a_{\lambda+\delta}}{a_\delta}$$

or equivalently (always working with n variables),

$$a_\delta e_\mu = \sum_\lambda K_{\lambda'\mu} a_{\lambda+\delta} \tag{9.44}$$

Since both sides of (9.44) are skew-symmetric, it is enough to show that the coefficient of $x^{\lambda+\delta}$ in $a_\delta e_\mu$ is $K_{\lambda'\mu}$. We multiply a_δ by e_μ by successively multiplying $e_{\mu_1}, e_{\mu_2}, \ldots$. Each partial product $a_\delta e_{\mu_1} \cdots e_{\mu_k}$ is skew-symmetric, so any term $x_1^{i_1} \cdots x_n^{i_n}$ appearing in $a_\delta e_{\mu_1} \cdots e_{\mu_k}$ has all exponents i_j *distinct*. When we multiply such a term $x_1^{i_1} \cdots x_n^{i_n}$ by a term $x_{m_1} \cdots x_{m_j}$ from $e_{\mu_{k+1}}$ (so $j = \mu_{k+1}$), either two exponents become equal, or the exponents maintain their relative order. If two exponents become equal then that term disappears from $a_\delta e_{\mu_1} \cdots e_{\mu_{k+1}}$. Hence to get the term $x^{\lambda+\delta}$, we must start with the term x^δ in a_δ and successively multiply by a term x^{α^1} of e_{μ_1}, then x^{α^2} of e_{μ_2} etc., keeping the exponents strictly decreasing. The number of ways to do this is the coefficient of $x^{\lambda+\delta}$ in $a_\delta e_\mu$.

Given the terms $x^{\alpha^1}, x^{\alpha^2}, \ldots$ as above, define an SSYT $T = T(\alpha^1, \alpha^2, \ldots)$ as follows: Column j of T contains an i if the variable x_j occurs in x^{α^i} (i.e. the j-th coordinate of α^i is equal to 1). For example, suppose $n = 4,\ \lambda = 5332,\ \lambda' = 44311,\ \lambda + \delta = 8542,\ \mu = 3222211,\ x^{\alpha^1} = x_1 x_2 x_3,\ x^{\alpha^2} = x_1 x_2,\ x^{\alpha^3} = x_3 x_4,\ x^{\alpha^4} = x_1 x_2, x^{\alpha^5} = x_1 x_4,\ x^{\alpha^6} = x_1,\ x^{\alpha^7} = x_3$. Then T is given by

$$
\begin{array}{l}
1113 \\
2235 \\
447 \\
5 \\
6
\end{array}
$$

It is easy to see that the map $(\alpha^1, \alpha^2, \ldots) \mapsto T(\alpha^1, \alpha^2, \ldots)$ gives a bijection between ways of building up the term $x^{\lambda+\delta}$ from x^{δ} (according to the rules above) and SSYT of shape λ' and type μ, so the proof follows. $\square$

From the combinatorial definition of Schur functions it is clear that $s_\lambda(x_1, \ldots, x_n) = 0$ if $l(\lambda) > n$. It is not hard to check thatt dim $(\Lambda_n) = \#\{\lambda \in \text{Par} : l(\lambda) \leq n\}$. It follows that the set $\{s_\lambda(x_1, \ldots, x_n) : l(\lambda) \leq n\}$ is a basis for Λ_n. (This also follows frm a simple extension of the proof of Corollary 9.27). We define on Λ_n a scalar product $\langle , \rangle_n$ by requiring that $\{s_\lambda(x_1, \ldots, x_n)\}$ is an orthonormal basis. If $f, g \in \Lambda$, then we write $\langle f, g\rangle_n$ as short for $\langle f(x_1, \ldots .x_n), g(x_1, \ldots, x_n)\rangle_n$. Thus

$$\langle f, g\rangle = \langle f, g\rangle_n$$

provided that every monomial appearing in f involves at most n distinct variables e.g., if deg $f \leq n$

Corollary 9.49. *If $f \in \Lambda_n$, $l(\lambda) \leq n$, and $\delta = (n-1, n-2, \ldots, 1, 0)$, then*

$$\langle f, s_\lambda \rangle_n = [x^{\lambda+\delta}] a_\delta f$$

the coefficient of $x^{\lambda+\delta}$ in $a_\delta f$.

Proof. All functions will be in the variables $x_1, \ldots, x_n$. Let $f = \sum_{l(\lambda) \leq n} c_\lambda s_\lambda$. Then by Theorem 9.48 we have

$$a_\delta f = \sum_{l(\lambda) \leq n} c_\lambda a_{\lambda+\delta},$$

so

$$\langle f, s_\lambda \rangle_n = c_\lambda = [x^{\lambda+\delta}] a_\delta f.$$

□

Let us now consider a “skew generalization” of Theorem 9.48. We continue to work in n variables $x_1, \ldots, x_n$. For any $\lambda\nu \in$ Par, $l(\lambda) \leq n, l(\nu) \leq n$, consider the expansion

$$s_\nu e_\mu = \sum_\lambda L^{\lambda'}_{\nu'\mu} s_\lambda,$$

or equivalently (multiplying by a_δ),

$$a_{\nu+\delta} e_\mu = \sum_\lambda L^{\lambda'}_{\nu'\mu} a_{\lambda+\delta} \tag{9.45}$$

Arguing as in the proof of Theorem 9.48 shows that $L^{\lambda'}_{\nu'\mu}$ is equal to the number of ways to write

$$\lambda + \delta = \nu + \delta + \alpha^1 + \alpha^2 + \cdots + \alpha^k,$$

Here $l(\mu) = k$, each α^i is a $(0,1)$-vector with μ_i 1's, and each a partial sum $\nu + \delta + \alpha^1 + \cdots + \alpha^i$ has strictly decreasing coordinates. Define a skew SSYT $T = T_{\lambda'/\nu'}(\alpha^1, \ldots, \alpha^k)$ of shape λ'/ν' and type μ by the condition that i appears in column j of T if the j-th coordinate of α^i is a 1. This establishes a bijection which shows that $L^{\lambda'}_{\nu'\mu}$ is equal to the skew Kostka number $K_{\lambda'/\nu',\mu}$, the number of skew SSYTs of shape λ'/ν' and type μ (see Equation (9.28)). (If $\nu' \not\subseteq \lambda'$ then this number is 0.)

Corollary 9.50. *We have*

$$s_\nu e_\mu = \sum_\lambda K_{\lambda'/\nu',\mu} s_\lambda. \tag{9.46}$$

Proof. Divide (9.45) by a_δ and let $n \to \infty$. □

Theorem 9.51. *For any $f \in \Lambda$, we have*

$$\langle f s_\nu, s_\lambda \rangle = \langle f, s_{\lambda/\nu} \rangle.$$

In other words, the two linear transformations $M_\nu : \Lambda \to \Lambda$ and $D_\nu : \Lambda \to \Lambda$ defined by $M_\nu f = s_\nu f$ and $D_\nu s_\lambda = s_{\lambda/\nu}$ are adjoint with respect to the scalar product $\langle \cdot, \cdot \rangle$. In particular

$$\langle s_\mu s_\nu, s_\lambda \rangle = \langle s_\mu, s_{\lambda/\nu} \rangle. \tag{9.47}$$

Proof. Apply ω to (9.46) and replace ν by ν' and λ by λ'. We obtain

$$s_\nu h_\mu = \sum_\lambda K_{\lambda/\nu,\mu} s_\lambda$$

Hence

$$\langle s_\nu h_\mu, s_\lambda \rangle = K_{\lambda/\nu,\mu} = \langle h_\mu, s_{\lambda/\nu} \rangle, \tag{9.48}$$

by (9.28) and the fact that $\langle h_\mu, m_\rho \rangle = \delta_{\mu\rho}$ by definition of $\langle \cdot, \cdot \rangle$. But equation (9.48) is linear in h_μ, so since $\{h_\mu\}$ is a basis for Λ, the proof follows. $\square$

Theorem 9.52. *For any $\lambda, \nu \in$ Par we have $\omega s_{\lambda/\nu} = s_{\lambda'/\nu'}$.*

Proof. By equation (9.47) and the fact that ω is an isometry, we have

$$\langle \omega(s_\mu s_\nu), \omega s_\lambda \rangle = \langle \omega s_\mu, \omega s_{\lambda/\nu} \rangle.$$

Hence by Theorem 9.47 we get

$$\langle s_{\mu'} s_{\nu'}, s_{\lambda'} \rangle = \langle s_{\mu'}, \omega s_{\lambda/\nu} \rangle \tag{9.49}$$

On the other hand, substituting λ', μ', ν' for λ, μ, ν respectively in (9.47) yields

$$\langle s_{\mu'} s_{\nu'}, s_{\lambda'} \rangle = \langle s_{\mu'}, s_{\lambda'/\nu'} \rangle \tag{9.50}$$

From Equations (9.49) and (9.50) there follows $\omega s_{\lambda/\nu} = s_{\lambda'/\nu'}$ □

9.14 The Jacobi-Trudi Identity

Theorem 9.53. *Let* $\lambda = (\lambda_1, \ldots, \lambda_n)$ *and* $\mu = (\mu_1, \ldots, \mu_n) \subseteq \lambda$. *Then*

$$s_{\lambda/\mu} = \textit{det}(h_{\lambda_i - \mu_j - i + j})_{i,j=1}^{n} \tag{9.51}$$

where we set $h_0 = 1$ *and* $h_k = 0$ *for* $k < 0$.

Proof. Let $c^\lambda_{\mu\nu} = \langle s_\lambda, s_\mu s_\nu \rangle$, so

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

Then

$$\begin{aligned}
\sum_\lambda s_{\lambda/\mu}(x) s_\lambda(y) &= \sum_{\lambda,\nu} c^\lambda_{\mu\nu} s_\nu(x) s_\lambda(y) \\
&= \sum_\nu s_\nu(x) s_\mu(y) s_\nu(y) \\
&= s_\mu(y) \sum_\nu h_\nu(x) m_\nu(y)
\end{aligned}$$

Let $y = (y_1, \ldots, y_n)$. Multiplying by $a_\delta(y)$ gives

$$\begin{aligned}\sum_\lambda s_{\lambda/\mu}(x) a_{\lambda+\delta}(y) &= \left(\sum_\nu h_\nu(x) m_\nu(y)\right) a_{\mu+\delta}(y) \\ &= \left(\sum_{\alpha \in \mathbb{N}^n} h_\alpha(x) y^\alpha\right) \left(\sum_{w \in \mathcal{S}_n} \varepsilon_w y^{w(\mu+\delta)}\right) \\ &= \sum_{w \in \mathcal{S}_n} \sum_\alpha \varepsilon_w h_\alpha(x) y^{\alpha + w(\mu+\delta)}\end{aligned}$$

Now take the coefficient of $y^{\lambda+\delta}$ on both sides (so we are looking for terms where $\lambda + \delta = \alpha + w(\mu + \delta)$). We get

$$\begin{aligned}s_{\lambda/\mu}(x) &= \sum_{w \in \mathcal{S}_n} \varepsilon_w h_{\lambda+\delta-w(\mu+\delta)}(x) \qquad (9.52) \\ &= \det(h_{\lambda_i - \mu_j - i + j}(x))_{i,j=1}^n\end{aligned}$$

□

Corollary 9.54 (Dual Jacobi-Trudi identity). *Let* $\mu \subseteq \lambda$ *with* $\lambda_1 \leq n$. *Then*

$$s_{\lambda/\mu} = \mathit{det}(e_{\lambda' - \mu'_j - i + j})_{i,j=1}^{n} \tag{9.53}$$

Let $f^{\lambda/\mu}$ be the number of SYT of shape λ/μ.

Corollary 9.55. *Let* $|\lambda/\mu| = N$ *and* $l(\lambda) \leq n$. *Then*

$$f^{\lambda/\mu} = N! \mathit{det} \left(\frac{1}{(\lambda_i - \mu_j - i + j)!} \right)_{i,j=1}^{n} \tag{9.54}$$

9.15 The Murnaghan-Nakayama Rule

A skew shape λ/μ is *connected* if the interior of the diagram of λ/μ, regarded as a union of *solid* squares, is a connected (open) set. A *border strip* (or *rim hook or ribbon*) is a connected skew shape with no 2×2 square.

Given positive integers $a_1, \ldots, a_k$, there is a unique border strip λ/μ (up to translation) with a_i squares in row i (i.e. $a_i = \lambda_i - \mu_i$). It follows that the number of border strips of size n (up to translation) is 2^{n-1}. Define the height $\text{ht}(B)$ of a border strip B to be one less than its number of rows.

Theorem 9.56. *For any* $\mu \in$ *Par and* $r \in \mathbb{N}$ *we have*

$$s_\mu p_r = \sum_\lambda (-1)^{ht(\lambda/\mu)} s_\lambda, \tag{9.55}$$

summed over all partitions $\lambda \supseteq \mu$ *for which* λ/μ *is a border strip of size* r.

Proof. Let $\delta = (n-1, n-2, \ldots, 0)$, and let all functions be in the variables $x_1, \ldots, x_n$. In equation 9.41 let $\alpha = \mu + \delta$ and multiply by p_r. We get

$$a_{\mu+\delta} p_r = \sum_{j=1}^{n} a_{\mu+\delta+r\epsilon_j}, \tag{9.56}$$

where ϵ_j is the sequence with a 1 in the j-th place and 0 elsewhere. Arrange the sequence $\mu + \delta + r\epsilon_j$ in descending order. If it has two terms equal, then it will contribute nothing to (9.56). Otherwise there is some $p \leq q$ for which

$$\mu_{p-1} + n - p + 1 > \mu_q + n - q + r > \mu_p + n - p,$$

in which case $a_{\mu+\delta+r\epsilon_j} = (-1)^{q-p} a_{\lambda+\delta}$, where λ is the partition

$$\lambda = (\mu_1, \ldots, \mu_{p-1}, \mu_q + p - q + r, \mu_p + 1, \ldots, \mu_{q-1} + 1, \mu_{q+1}, \ldots, \mu_n)$$

Such partitions are precisely those for which λ/μ is a border strip B of size r, and $q-p$ is just $\mathsf{ht}(B)$. Hence

$$a_{\mu+\delta}p_r = \sum_{\lambda}(-1)^{\mathsf{ht}(\lambda/\mu)}a_{\lambda+\delta}$$

Divide by a_δ and let $n \to \infty$ to obtain 9.55 □

Let $\alpha = (\alpha_1, \alpha_2, \ldots)$ be a weak composition of n. Define a *border-strip tableaux* (or *rim-hook tableaux*) of shape λ/μ (where $|\lambda/\mu| = n$) and type α to be an assignment of positive integers to the squares of λ/μ such that

(a) every row and column is weakly increasing

(b) the integer i appears α_i times, and

(c) the set of squares occupied by i forms a border strip.

Equivalently, one may think of a border-strip tableau as a sequence $\mu = \lambda^0 \subseteq \lambda^1 \subseteq \cdots \subseteq \lambda^r \subseteq \lambda$ of partitions such that each skew shape λ^i/λ^{i+1} is a border strip of size α_i (including the empty border-strip $\emptyset$ when $\alpha_i = 0$).

Define the height $\mathrm{ht}(T)$ of a border-strip tableau T to be

$$\mathrm{ht}(T) = \mathrm{ht}(B_1) + \mathrm{ht}(B_2) \cdots + \mathrm{ht}(B_k)$$

where $B_1, \ldots, B_k$ are the (nonempty) border strips appearing in T.

Theorem 9.57. *We have*

$$s_\mu p_\alpha = \sum_\lambda \chi^{\lambda/\mu}(\alpha) s_\lambda, \tag{9.57}$$

where

$$\chi^{\lambda/\mu}(\alpha) = \sum_T (-1)^{ht(T)} \tag{9.58}$$

summed over all border-strip tableaux of shape λ/μ and type α

Corollary 9.58. *We have*

$$p_\alpha = \sum_\lambda \chi^\lambda(\alpha) s_\lambda \tag{9.59}$$

where $\chi^\lambda(\alpha)$ *is given by (9.58) with* $\mu = \emptyset$.

Corollary 9.59. *We have*

$$s_{\lambda/\mu} = \sum_{\nu} z_\nu^{-1} \chi^{\lambda/\mu}(\nu) p_\nu, \tag{9.60}$$

where $\chi^{\lambda/\mu}(\nu)$ is given by (9.58).

Proof. We have from (9.57) that

$$\begin{aligned} \chi^{\lambda/\mu}(\nu) &= \langle s_\mu p_\nu, s_\lambda \rangle \\ &= \langle p_\nu, s_{\lambda/\mu} \rangle, \end{aligned}$$

and the proof follows from Proposition 9.18. □

The orthogonality properties of the bases $\{s_\lambda\}$ and $\{p_\lambda\}$ translate into orthogonality relations satisfied by the coefficients $\chi^\lambda(\mu)$.

Proposition 9.60. *(a) Fix μ, ν. Then*

$$\sum_\lambda \chi^\lambda(\mu)\chi^\lambda(\nu) = z_\mu \delta_{\mu\nu}$$

(b) Fix λ, μ. Then

$$\sum_\nu z_\nu^{-1}\chi^\lambda(\nu)\chi^\mu(\nu) = \delta_{\lambda\mu}$$

Proof. (a) Expand p_μ and p_ν by (9.59) and take $\langle p_\mu, p_\nu\rangle$.

(b) Expand s_λ and s_μ by (9.60) and take $\langle s_\lambda, s_\mu\rangle$. □

10 Characters of the Symmetric and Unitary Groups

10.1 Characters of the Symmetric Group

Let $\mathbf{CF}^n$ denote the set of all class functions (i.e. functions constant on conjugacy classes) $f : \mathcal{S}_n \rightarrow \mathbb{Q}$. Recall that $\mathbf{CF}^n$ has a natural scalar product defined by

$$\langle f, g \rangle = \frac{1}{n!} \sum_{w \in \mathcal{S}_n} f(w) g(w)$$

Sometimes by abuse of notation we write $\langle \phi, \gamma \rangle$ instead of $\langle f, g \rangle$ when ϕ and γ are representations of $\mathcal{S}_n$ with characters f and g.

If $\alpha = (\alpha_1, \ldots, \alpha_l)$ is a vector of positive integers and $|\alpha| := \alpha_1 + \cdots + \alpha_l = n$, then recall the Young subgroup $\mathcal{S}_\alpha \subseteq \mathcal{S}_n$ given by

$$\mathcal{S}_\alpha = \mathcal{S}_{\alpha_1} \times \mathcal{S}_{\alpha_2} \times \cdots \times \mathcal{S}_{\alpha_l}$$

where $\mathcal{S}_{\alpha_1}$ permutes $1, 2, \ldots, \alpha_1$; $\mathcal{S}_{\alpha_2}$ permutes $\alpha_1 + 1, \alpha_1 + 2, \ldots \alpha_1 + \alpha_2$ etc.

Consider the following linear transformations $\text{ch} : \mathbf{CF}^n \to \Lambda^n$ called the *Frobenius characteristic maps*. If $f \in \mathbf{CF}^n$, then

$$\begin{aligned} \text{ch } f &= \frac{1}{n!} \sum_{w \in \mathcal{S}_n} f(w) p_{\rho(w)} \\ &= \sum_{\mu} z_\mu^{-1} f(\mu) p_\mu \end{aligned}$$

where $f(\mu)$ denotes $f(w)$ for any type $\rho(w) = \mu$. Equivalently, extending the ground field $\mathbb{Q}$ to the algebra Λ and defining $\Psi(w) = p_{\rho(w)}$, we have

$$\text{ch } f = \langle f, \Psi \rangle \qquad (10.1)$$

Note that if f_μ is the class function defined by

$$f_\mu(w) = \begin{cases} 1, & \text{if } \rho(w) = \mu, \\ 0, & \text{otherwise.} \end{cases}$$

then $\mathrm{ch} f_\mu = z_\mu^{-1} p_\mu$.

NOTE. Let $\varphi : \mathcal{S}_n \to GL(V)$ be a representation of $\mathcal{S}_n$ with character χ. Some times by abuse of notation we will write ch φ or ch V instead of ch χ.

Proposition 10.1. *The linear transformation ch is an isometry i.e.,*

$$\langle f, g\rangle_{\mathbf{CF}^n} = \langle \textit{ch } f, \textit{ch } g\rangle_{\Lambda^n}.$$

Proof. We have (using Proposition 9.18)

$$\begin{aligned}\langle \text{ch } f, \text{ch } g\rangle &= \langle \sum_{\mu} z_{\lambda}^{-1} f(\lambda) p_{\lambda}, \sum_{\mu} z_{\mu}^{-1} g(\mu) p_{\mu}\rangle \\ &= \sum_{\lambda} z_{\lambda}^{-1} f(\lambda) g(\lambda) \\ &= \langle f, g\rangle\end{aligned}$$

□

We now want to define a product on class functions that will correspond to the ordinary product of symmetric functions under the characteristic map ch . Let $f \in \mathbf{CF}^m$ and $g \in \mathbf{CF}^n$. Define the pointwise product $f \times g \in \mathbf{CF}(\mathcal{S}_m \times \mathcal{S}_n)$ by

$$(f \times g)(u, v) = f(u)g(v).$$

If f and g are characters of representations of φ and ψ, then $f \times g$ is just the character of the tensor product representation $\varphi \otimes \psi$ of $\mathcal{S}_m \times \mathcal{S}_n$. Now define the *induction product* $f \circ g$ of f and g to be the induction of $f \times g$ to $\mathcal{S}_{m+n}$ where as before $\mathcal{S}_m$ permutes $1, 2, \ldots, m$ while $\mathcal{S}_n$ permutes $m + 1, m + 2, \ldots, m + n$. in symbols

$$f \circ g = \text{ind}_{\mathcal{S}_m \times \mathcal{S}_n}^{\mathcal{S}_{m+n}} (f \times g).$$

Let $\mathbf{CF} = \mathbf{CF}^0 \oplus \mathbf{CF}^1 \oplus \cdots$, and extend the scalar product on $\mathbf{CF}^n$ to all of $\mathbf{CF}$ by setting $\langle f, g\rangle = 0$ if $f \in \mathbf{CF}^m$, $g \in \mathbf{CF}^n$, and $m \neq n$. The induction product on characters extends to all of $\mathbf{CF}$ by (bi)linearity. It is not hard to check that this takes $\mathbf{CF}$ into an associative commutative graded $\mathbb{Q}$-algebra with the identity $1 \in \mathbf{CF}^0$. Similarly we can extend the characteristic map ch to a linear transformation ch $: \mathbf{CF} \to \Lambda$.

Proposition 10.2. *The characteristic map* $\mathit{ch} : \mathbf{CF} \to \Lambda$ *is a bijective algebra homomorphism, i.e.* ch *is one-to-one and onto, and satisfies*

$$\mathit{ch}\,(f \circ g) = (\mathit{ch}\ f)(\mathit{ch}\ g)$$

Proof. Let $\text{res}_H^G f$ denote the restriction of the class function f on G to the subgroup H. We then have

$$
\begin{aligned}
\text{ch }(f \circ g) &= \text{ch }(\text{ind}_{\mathcal{S}_m \times \mathcal{S}_n}^{\mathcal{S}_{m+n}}(f \times g)) \\
&= \langle \text{ind}_{\mathcal{S}_m \times \mathcal{S}_n}^{\mathcal{S}_{m+n}}(f \times g), \Psi \rangle && \text{by (10.1)} \\
&= \langle f \times g, \text{res}_{\mathcal{S}_m \times \mathcal{S}_n}^{\mathcal{S}_{m+n}} \Psi \rangle_{\mathcal{S}_m \times \mathcal{S}_n} && \text{(by Frobenius reciprocity)} \\
&= \frac{1}{m!n!} \sum_{u \in \mathcal{S}_m} \sum_{v \in \mathcal{S}_n} f(u)g(v)\Psi(uv) \\
&= \frac{1}{m!n!} \sum_{u \in \mathcal{S}_m} \sum_{v \in \mathcal{S}_n} f(u)g(v)\Psi(u)\Psi(v) \\
&= \langle f, \Psi \rangle_{\mathcal{S}_m} \langle g, \Psi \rangle_{\mathcal{S}_n} \\
&= (\text{ch } f)(\text{ch } g)
\end{aligned}
$$

Moreover, from the definition of ch and the fact that the power sums p_μ form a $\mathbb{Q}$-basis for Λ it follows that ch is bijective. $\square$

Note that by Equation (9.19) and the definition of ch we have

$$\text{ch } 1_{\mathcal{S}_n} = \sum_{\lambda \vdash n} z_\lambda^{-1} p_\lambda = h_n \tag{10.2}$$

Corollary 10.3. *We have ch* $1_{\mathcal{S}_\alpha}^{\mathcal{S}_n} = h_\alpha$

Proof. Since $1_{\mathcal{S}_\alpha}^{\mathcal{S}_n} = 1_{\mathcal{S}_{\alpha_1}} \circ 1_{\mathcal{S}_{\alpha_2}} \circ \cdots \circ 1_{\mathcal{S}_{\alpha_l}}$, the proof follows from Proposition 10.2 and Equation (10.2). □

Now let $\mathbf{R}^n$ denote the set of all virtual characters of $\mathcal{S}_n$, i.e. functions of $\mathcal{S}_n$ that are the difference of two characters ($=$ integer linear combinations of irreducible characters). Thus $\mathbf{R}^n$ is a lattice (discrete subgroup of maximum rank) in the vector space $\mathbf{CF}^n$. The rank of $\mathbf{R}^n$ is $p(n)$, the number of partitions of n, and a basis consists of the irreducible characters of $\mathcal{S}_n$. This basis is the unique orthonormal basis of $\mathbf{R}^n$ up to sign and order, since the transition matrix between two such bases must be a integral orthogonal matrix and hence a signed permutation. Define $\mathbf{R} = \mathbf{R}^0 \oplus \mathbf{R}^1 \oplus \cdots$.

Proposition 10.4. *The image of* $\mathbf{R}$ *under the characteristic map ch is* $\Lambda_{\mathbb{Z}}$. *Hence ch* $: \mathbf{R} \to \Lambda_{\mathbb{Z}}$ *is a ring isomorphism.*

Proof. It will suffice to find integer linear combinations of the characters η^α of the representations $1_{\mathcal{S}_\alpha}^{\mathcal{S}_n}$ that are irreducible characters of $\mathcal{S}_n$. The Jacobi-Trudi identity (Theorem 9.53) suggests we define the (possibly virtual) characters $\psi^\lambda = \det(\eta^{\lambda_i - i + j})$, where the product used in evaluating the determinant is the induction product. Then by the Jacobi-Trudi identity and Proposition 10.1 we have

$$\text{ch}\ (\psi^\lambda) = s_\lambda \tag{10.3}$$

Since ch is an isometry (Proposition 10.1) we get $\langle \psi^\lambda, \psi^\mu \rangle = \delta_{\lambda\mu}$. As pointed out above, this means that the class functions ψ^λ are, up to sign, the irreducible characters of $\mathcal{S}_n$. Hence the ψ^λ for $\lambda \vdash n$ form a $\mathbb{Z}$ basis for $\mathbf{R}^n$, and the image of $\mathbf{R}^n$ is the $\mathbb{Z}$-span of the s_λ's which is just $\Lambda_{\mathbb{Z}}^n$ as claimed. □

Theorem 10.5. *Regard the functions* χ^λ *(where* $\lambda \vdash n$*) of Section 9.15 as functions on* $\mathcal{S}_n$ *given by* $\chi^\lambda(w) = \chi^\lambda(\mu)$*, where* w *has cycle type* μ*. Then the* χ^λ *are the irreducible characters of the symmetry group* $\mathcal{S}_n$*.*

By the Murnaghan-Nakayama rule (Corollary 9.59), we have

$$\text{ch }(\chi^\lambda) = \sum_\mu z_\mu^{-1} \chi^\lambda(\mu) p_\mu = s_\lambda$$

Hence by Equation (10.3), we get $\chi^\lambda = \psi^\lambda$. Since the ψ^λ, up to sign, are the irreducible characters of $\mathcal{S}_n$, it remains only to determine wether χ^λ or $-\chi^\lambda$ is a character. But $\chi^\lambda(1^n) = f^\lambda$, so χ^λ is an irreducible character. $\square$

By definition, η^λ is the character of the module M^λ. It can be shown that χ^λ is the character of the Specht module S^λ.

Proposition 10.6. *Let α be a composition of n and $\lambda \vdash n$. Then the multiplicity of the irreducible character χ^λ in the character η^α is just the Kostka number $K_{\lambda\alpha}$. In symbols*

$$\langle \eta^\alpha, \chi^\lambda \rangle = K_{\lambda\alpha}.$$

Proof. By Corollary 10.3 we have ch $\eta^\alpha = h_\alpha$. Then the proof follows from Corollary 9.35 and Theorem 10.5. □

10.2 The characters of GL$(n, \mathbb{C})$

A *linear representation* of $GL(V)$ is a homomorphism $\varphi : GL(V) \to GL(W)$, where W is a complex vector space. From now on we assume that all representations are *finite dimensional* i.e dim $(W) < \infty$. We call dim (W) the dimension of the representation φ, denoted dim (φ).

The representation φ is a *polynomial* representation if, after choosing ordered bases for V and W, the entries of $\varphi(A)$ are polynomials in the entries of $A \in GL(n, \mathbb{C})$. It is clear that the notion of polynomial representations is independent of the choice of ordered bases of V and W, since linear combinations of polynomials remain polynomials.

Fact: If φ is a polynomial representation of $GL(V)$, then there is a symmetric polynomial char φ in dim V variables such that

$$Tr\varphi(A) = \text{char } \varphi(\theta_1, \ldots, \theta_n)$$

for all $A \in GL(V)$, where $\theta_1, \ldots, \theta_n$ are the eigenvalues of A.

Theorem 10.7. *The irreducible polynomial representations* φ^λ *of GL(V) can be indexed by partitions* λ *of length at most* n *so that*

$$\text{char}\ \varphi^\lambda = s_\lambda(x_1, \ldots, x_n)$$

Examples:

- If $\varphi(A) = 1$ (the trivial representation), then char $\varphi = s_\emptyset = 1$.
- If $\varphi(A) = A$ (the defining representation), then char $\varphi = x_1 + \cdots + x_n = s_1$.
- If $\varphi(A) = (\text{det}A)^m$ for an positive integer m, then char $\varphi = (x_1 \cdots x_n)^m = s_{m^n}$.

Proof. (**Sketch**) Let V be an n-dimensional complex vector space. Then $GL(V)$ acts diagonally on the k-th tensor power $V^{\otimes k}$ i.e

$$A \cdot (v_1 \otimes \cdots v_k) = (A \cdot v_1) \otimes \cdots \otimes (A \cdot v_k), \tag{10.4}$$

and the symmetric group $\mathcal{S}_k$ acts on $V^{\otimes k}$ by permuting tensor coordinates, i.e.

$$w \cdot (v_1 \otimes \cdots v_k) = v_{w^{-1}(1)} \otimes \cdots v_{w^{-1}(k)} \tag{10.5}$$

The actions of $GL(V)$ and $\mathcal{S}_k$ commute, so we have an action of $\mathcal{S}_k \times GL(V)$ on $V^{\otimes k}$. A crucial fact is that the actions of $GL(V)$ and $\mathcal{S}_k$ *centralize* each other. i.e the (invertible) linear transformations $V^{\otimes k} \to V^{\otimes k}$ that commute with the $\mathcal{S}_k$ action are just those given by Eqn. (10.4), while conversely the linear transformations that commute with the $GL(V)$ actions are those generated (as a $\mathbb{C}$ algebra) by Eqn (10.5). From this it can be shown that $V^{\otimes k}$ decomposes into irreducible $\mathcal{S}_k \times GL(V)$-modules as follows

$$V^{\otimes k} = \bigoplus (M^{\lambda} \otimes F^{\lambda}) \tag{10.6}$$

where $\bigoplus$ denotes the direct sum (the "double commutant theorem").

Here the M_λ's are nonisomorphic irreducible $\mathcal{S}_k$ modules, the F^λ's are nonisomorphic irreducible $GL(V)$ modules, and λ ranges over some index set. We know (Theorem 10.5) that the irreducible representations of $\mathcal{S}_k$ are indexed by partitions λ of k, so we choose the indexing so that M^λ is the irreducible $\mathcal{S}_k$ module corresponding to $\lambda \vdash k$ via Theorem 10.5. Thus we have constructed irreducible (or possibly 0) $GL(V)$-modules F^λ. These modules afford polynomial representations φ^λ, and the nonzero ones are inequivalent.

Next we compute the character of φ^λ. Let $w \times A$ be an element of $\mathcal{S}_k \times GL(V)$, and let $tr(w \times A)$ denote the trace of $w \times A$ acting on $V^{\otimes k}$. Then by Equation (10.6) we have

$$tr(w \times A) = \sum_\lambda \chi^\lambda(w) \cdot tr(\varphi^\lambda(A)).$$

Let A have eigenvalues $\theta = (\theta_1, \ldots, \theta_n)$. A straightforward computation shows that $tr(w \times A) = p_{\rho(w)}(\theta)$, so

$$p_{\rho(w)}(\theta) = \sum_\lambda \chi^\lambda(w)(\text{char } \varphi^\lambda)(\theta)$$

But we know (Corollary 9.59) that

$$p_{\rho(w)} = \sum_\lambda \chi^\lambda(w) s_\lambda$$

Since the χ^λ's are linearly independent, we conclude char $\varphi^\lambda = s_\lambda$.

A separate argument shows that there are no other irreducible polynomial characters. $\square$

Fact: The φ^λ remain irreducible when restricted to $U(V)$ because the $(\text{dim } V)^2$ entries of a general unitary matrix are algebraically independent, and so every irreducible polynomial representation of $GL(V)$ is still irreducible when restricted to $U(V)$.

11 Eigenvalues of random matrices

For $n \in \mathbb{N}$, let M_n be a random $n \times n$ unitary matrix with distribution given by Haar measure on the unitary group. The eigenvalues of M_n lie on the unit circle $\mathbb{T}$ of the complex plane $\mathbb{C}$. Write Ξ_n for the random measure on $\mathbb{T}$ that places a unit mass at each of the eigenvalues of M_n. That is, if the eigenvalues are $\{\nu_{n1}, \ldots, \nu_{nn}\}$, then

$$\Xi_n(f) := \int_{\mathbb{T}} f \, d\Xi_n = \sum_j f(\nu_{nj})$$

Note that if $f : \mathbb{T} \to \mathbb{C}$ has Fourier expansion $f(e^{i\theta}) = \sum_{j\in\mathbb{Z}} \hat{f}_j e^{ij\theta}$, then

$$\Xi_n(f) = n\hat{f}_0 + \sum_{j=1}^{\infty} \hat{f}_j \operatorname{Tr}(M_n^j) + \sum_{j=1}^{\infty} \hat{f}_{-j} \overline{\operatorname{Tr}(M_n^j)},$$

where Tr denotes the trace.

Questions about the asymptotic behaviour of $c_n(\Xi_n(f_n) - \mathbb{E}[\Xi_n(f_n)])$ for a sequence of test functions $\{f_n\}$ and sequence of norming constants $\{c_n\}$ may therefore be placed in the larger framework of questions about the asymptotic behaviour of $\sum_{j=1}^{\infty}(a_{nj} \operatorname{Tr}(M_n^j) + b_{nj} \overline{\operatorname{Tr}(M_n^j)})$ for arrays of complex constants $\{a_{nj} : n \in \mathbb{N},\, j \in \mathbb{N}\}$ and $\{b_{nj} : n \in \mathbb{N},\, j \in \mathbb{N}\}$.

Definition 11.1. A complex random variable is said to be standard complex normal if the real and imaginary parts are independent centred (real) normal random variables with common variance $\frac{1}{2}$.

11.1 Moments of Traces

Theorem 11.2. *a) Consider* $a = (a_1, \ldots, a_k)$ *and* $b = (b_1, \ldots, b_k)$ *with* $a_j, b_j \in \{0, 1, \ldots\}$*. Let* $Z_1, Z_2, \ldots Z_k$ *be independent standard complex normal random variables. Then for* $n \geq (\sum_{j=1}^{k} j a_j) \vee (\sum_{j=1}^{k} j b_j)$*,*

$$\mathbb{E}\left[\prod_{j=1}^{k} \left(\operatorname{Tr}(M_n^j)\right)^{a_j} \overline{\left(\operatorname{Tr}(M_n^j)\right)^{b_j}}\right] = \delta_{ab} \prod_{j=1}^{k} j^{a_j} a_j!$$

$$= \mathbb{E}\left[\prod_{j=1}^{k} \left(\sqrt{j} Z_j\right)^{a_j} \overline{\left(\sqrt{j} Z_j\right)^{b_j}}\right].$$

b) For any j, k*,*

$$\mathbb{E}\left[\operatorname{Tr}(M_n^j) \overline{\operatorname{Tr}(M_n^k)}\right] = \delta_{jk}(j \wedge n).$$

Proof. (a) Define the simple power sum symmetric function p_j to be the symmetric function $p_j(x_1, \ldots, x_n) = x_1^j + \cdots + x_n^j$. Let μ be the partition $(1^{a_1}, 2^{a_2}, \ldots, k^{a_k})$ of the integer $K = 1a_1 + 2a_2 + \cdots + ka_k$ and set $p_\mu = \prod_j p_j^{a_j}$ to be the corresponding compound power sum symmetric function. Associate μ with the conjugacy class of the symmetric group on K letters that consists of permutations with a_j j–cycles for $1 \leq j \leq k$. We have the expansion

$$p_\mu = \sum_{\lambda \vdash K} \chi_\mu^\lambda s_\lambda,$$

Here the sum is over all partitions of K, the coefficient χ^λ_μ is the character of the irreducible representation of the symmetric group associated with the partition λ evaluated on the conjugacy class associated with the partition μ, and s_λ is the Schur function corresponding to the partition λ

Given an $n \times n$ unitary matrix U, write $s_\lambda(U)$ (resp. $p_\mu(U)$) for the function s_λ (resp. p_μ) applied to the eigenvalues of U. Writing $\ell(\lambda)$ for the number of parts of the partition λ (that is, the length of λ), the functions $U \mapsto s_\lambda(U)$ are irreducible characters of the unitary group when $\ell(\lambda) \leq n$ and $s_\lambda(U) = 0$ otherwise. Thus

$$\mathbb{E}\left[s_\lambda(M_n)\overline{s_\pi(M_n)}\right] = \delta_{\lambda\pi}\mathbf{1}(\ell(\lambda) \leq n),$$

Set $\nu = (1^{b_1}, 2^{b_2}, \ldots, k^{b_k})$ and $L = 1b_1 + 2b_2 + \cdots + kb_k$.

We have

$$
\begin{aligned}
&\mathbb{E}\left[\prod_{j=1}^{k}\left(\operatorname{Tr}(M_n^j)\right)^{a_j}\overline{\left(\operatorname{Tr}(M_n^j)\right)^{b_j}}\right] \\
&= \mathbb{E}\left[p_\mu(M_n)\overline{p_\nu(M_n)}\right] \\
&= \mathbb{E}\left[\left(\sum_{\lambda\vdash K}\chi_\mu^\lambda s_\lambda(M_n)\right)\overline{\left(\sum_{\pi\vdash L}\chi_\nu^\pi s_\pi(M_n)\right)}\right] \\
&= \delta_{KL}\sum_{\lambda\vdash K}\chi_\mu^\lambda\overline{\chi_\nu^\lambda}\mathbf{1}(\ell(\lambda)\le n).
\end{aligned}
\tag{11.1}
$$

When $K \le n$, all partitions of K are necessarily of length at most n, and so, by the second orthogonality relation for characters of the symmetric group, the rightmost term of (11.1) becomes

$$\delta_{KL}\delta_{\mu\nu}\prod_{j=1}^{k} j^{a_j} a_j! = \delta_{ab}\prod_{j=1}^{k} j^{a_j} a_j!,$$

which coincides with the claimed mixed moment of $\sqrt{j}Z_j$, $1 \le j \le k$.

(b) We have from (11.1) that

$$\mathbb{E}\left[\operatorname{Tr}(M_n^j)\overline{\operatorname{Tr}(M_n^k)}\right] = \delta_{jk}\sum_{\lambda \vdash j}\left|\chi^\lambda_{(j)}\right|^2 \mathbf{1}(\ell(\lambda) \le n),$$

where (j) is the partition of j consisting of a single part of size j. Now $\chi^\lambda_{(j)} = 0$ unless λ is a hook partition (that is, a partition with at most one part of size greater than 1), in which case

$$\chi^\lambda_{(j)} = (-1)^{\ell(\lambda)-1}$$

Since there are $j \wedge n$ hook partitions of j of length at most n, part (b) follows. □

11.2 Linear combination of Traces

Theorem 11.3. *Consider an array of complex constants* $\{a_{nj} : n \in \mathbb{N}, j \in \mathbb{N}\}$*. Suppose there exists* σ^2 *such that*

$$\lim_{n\to\infty} \sum_{j=1}^{\infty} |a_{nj}|^2 (j \wedge n) = \sigma^2.$$

Suppose also that there exists a sequence of positive integers $\{m_n : n \in \mathbb{N}\}$ *such that*

$$\lim_{n\to\infty} m_n/n = 0$$

and

$$\lim_{n\to\infty} \sum_{j=m_n+1}^{\infty} |a_{nj}|^2 (j \wedge n) = 0.$$

Then $\sum_{j=1}^{\infty} a_{nj} \, \mathrm{Tr}\,(M_n^j)$ *converges in distribution as* $n \to \infty$ *to* σZ*, where* Z *is a complex standard normal random variable.*

Proof. Recall from Theorem 11.2 that $\mathbb{E}[\operatorname{Tr}(M_n^j)] = 0$ and $\mathbb{E}[\operatorname{Tr}(M_n^j)\overline{\operatorname{Tr}(M_n^k)}] = \delta_{jk}(j \wedge n)$. Consequently, the series $\sum_{j=1}^{\infty} a_{nj} \operatorname{Tr}(M_n^j)$ converges in L^2 for each n and $\lim_{n\to\infty} \mathbb{E}[|\sum_{j=m_n+1}^{\infty} a_{nj} \operatorname{Tr}(M_n^j)|^2] = 0.$

It therefore suffices to show that $\sigma^{-1} \sum_{j=1}^{m_n} a_{nj} \operatorname{Tr}(M_n^j)$ converges in distribution as $n \to \infty$ to a complex standard normal random variable. Let $Z_0, Z_1, Z_2, \ldots$ be a sequence of independent complex standard normals.

From Theorem 11.2 we know that

$$\mathbb{E}\left[\left\{\sum_{j=1}^{m_n} a_{nj}\,\mathrm{Tr}\left(M_n^j\right)\right\}^{\alpha}\overline{\left\{\sum_{j=1}^{m_n} a_{nj}\,\mathrm{Tr}\left(M_n^j\right)\right\}^{\beta}}\right]$$

$$=\mathbb{E}\left[\left\{\sum_{j=1}^{m_n} a_{nj}\sqrt{j}Z_j\right\}^{\alpha}\overline{\left\{\sum_{j=1}^{m_n} a_{nj}\sqrt{j}Z_j\right\}^{\beta}}\right]$$

$$=\mathbb{E}\left[\left\{\left(\sum_{j=1}^{m_n}|a_{nj}|^2 j\right)^{1/2} Z_0\right\}^{\alpha}\overline{\left\{\left(\sum_{j=1}^{m_n}|a_{nj}|^2 j\right)^{1/2} Z_0\right\}^{\beta}},\right]$$

provided that $\alpha m_n \leq n$ and $\beta m_n \leq n$. The result now follows by convergence of moments for complex normal distributions and the assumption that $m_n/n \to 0$.

Theorem 11.4. *Consider arrays of complex constants $\{a_{nj} : n \in \mathbb{N},\ j \in \mathbb{N}\}$ and $\{b_{nj} : n \in \mathbb{N},\ j \in \mathbb{N}\}$. Suppose there exist σ^2, τ^2, and γ such that*

$$\lim_{n\to\infty} \sum_{j=1}^{\infty} |a_{nj}|^2 (j \wedge n) = \sigma^2,$$

$$\lim_{n\to\infty} \sum_{j=1}^{\infty} |b_{nj}|^2 (j \wedge n) = \tau^2,$$

and

$$\lim_{n\to\infty} \sum_{j=1}^{\infty} a_{nj} b_{nj} (j \wedge n) = \gamma.$$

Suppose also that there exists a sequence of positive integers $\{m_n : n \in \mathbb{N}\}$ such that

$$\lim_{n\to\infty} m_n/n = 0$$

and

$$\lim_{n\to\infty} \sum_{j=m_n+1}^{\infty} (|a_{nj}|^2 + |b_{nj}|^2)(j \wedge n) = 0.$$

Then $\sum_{j=1}^{\infty}(a_{nj}\,\mathrm{Tr}\,(M_n^j) + b_{nj}\,\overline{\mathrm{Tr}\,(M_n^j)})$ *converges in distribution as* $n \to \infty$ *to* $X + iY$*, where* (X, Y) *is a pair of centred jointly normal real random variables with*

$$\mathbb{E}[X^2] = \frac{1}{2}(\sigma^2 + \tau^2 + 2\Re\gamma),$$

$$\mathbb{E}[Y^2] = \frac{1}{2}(\sigma^2 + \tau^2 - 2\Re\gamma),$$

and

$$\mathbb{E}[XY] = \Im\gamma.$$

Given $f \in L^2(\mathbb{T})$ (where we define $L^2(\mathbb{T})$ to be the space of *real–valued* square–integrable functions), write

$$\hat{f}_j := \frac{1}{2\pi} \int e^{-ij\theta} f(\theta)\, d\theta, \quad j \in \mathbb{Z},$$

for the Fourier coefficients of f.

Recall that a positive sequence $\{c_k\}_{k\in\mathbb{N}}$ is said to be *slowly varying* if

$$\lim_{k\to\infty} \frac{c_{\lfloor \lambda k \rfloor}}{c_k} = 1, \quad \lambda > 0,$$

Theorem 11.5. *Suppose that* $f \in L^2(\mathbb{T})$ *is such that the sequence* $\{\sum_{j=-k}^{k} |\hat{f}_j|^2 j\}_{k\in\mathbb{N}}$ *is slowly varying. Then*

$$\frac{\Xi_n(f) - \mathbb{E}[\Xi_n(f)]}{\sqrt{\sum_{j=-n}^{n} |\hat{f}_j|^2 |j|}}$$

converges in distribution to a standard normal random variable as $n \to \infty$*.*

Let $H_2^{\frac{1}{2}}$ denote the space of functions $f \in L^2(\mathbb{T})$ such that

$$\|f\|_{\frac{1}{2}}^2 := \sum_{j\in\mathbb{Z}} |\hat{f}_j|^2 |j| < \infty,$$

and define an inner product on $H_2^{\frac{1}{2}}$ by

$$\langle f, g\rangle_{\frac{1}{2}} := \sum_{j\in\mathbb{Z}} \hat{f}_j \, \overline{\hat{g}_j} |j|.$$

Alternatively, $H_2^{\frac{1}{2}}$ is the space of functions $f \in L^2(\mathbb{T})$ such that

$$\frac{1}{16\pi^2} \iint \frac{\left(f(\phi) - f(\theta)\right)^2}{\sin^2\left(\frac{\phi-\theta}{2}\right)} \, d\theta \, d\phi < \infty, \tag{11.2}$$

Moreover,

$$\langle f, g \rangle_{\frac{1}{2}} = \frac{1}{16\pi^2} \iint \frac{\left(f(\phi) - f(\theta)\right)\left(g(\phi) - g(\theta)\right)}{\sin^2\left(\frac{\phi-\theta}{2}\right)} \, d\theta \, d\phi$$

Theorem 11.6. *If* $f_1, \ldots, f_k \in H_2^{\frac{1}{2}}$ *with* $\mathbb{E}[\Xi_n(f_h)] = n \int f_j(\theta)\, d\theta = 0$ *for* $1 \leq h \leq k$, *then the random vector* $(\Xi_n(f_1), \ldots, \Xi_n(f_k))$ *converges in distribution to a jointly normal, centred random vector* $(\Xi(f_1), \ldots, \Xi(f_k))$ *with* $\mathbb{E}[\Xi(f_h)\Xi(f_\ell)] = \langle f_h, f_\ell \rangle_{\frac{1}{2}}$.

For $0 \leq \alpha < \beta < 2\pi$ write $N_n(\alpha, \beta)$ for the number of eigenvalues of M_n of the form $e^{i\theta}$ with $\theta \in [\alpha, \beta]$. That is, $N_n(\alpha, \beta) = \Xi_n(f)$ where f is the indicator function of the arc $\{e^{i\theta} : \theta \in [\alpha, \beta]\}$. Note that $\mathbb{E}[N_n(\alpha, \beta)] = n(\beta - \alpha)/2\pi$.

Theorem 11.7. *As $n \to \infty$, the finite–dimensional distributions of the processes*

$$\frac{N_n(\alpha,\beta) - \mathbb{E}[N_n(\alpha,\beta)]}{\frac{1}{\pi}\sqrt{\log n}}, \quad 0 \le \alpha < \beta < 2\pi,$$

converge to those of a centred Gaussian process $\{Z(\alpha,\beta) : 0 \le \alpha < \beta < 2\pi\}$ *with the covariance structure*

$$\mathbb{E}[Z(\alpha,\beta)Z(\alpha',\beta')] = \begin{cases} 1, & \text{if } \alpha = \alpha' \text{ and } \beta = \beta', \\ \frac{1}{2}, & \text{if } \alpha = \alpha' \text{ and } \beta \neq \beta', \\ \frac{1}{2}, & \text{if } \alpha \neq \alpha' \text{ and } \beta = \beta', \\ -\frac{1}{2}, & \text{if } \beta = \alpha', \\ 0, & \text{otherwise.} \end{cases}$$