8

Group Representations

In this chapter we want to have a first look into the representation theory of various groups with extra structure, such as algebraic or compact groups. We will use the necessary techniques from elementary algebraic geometry or functional analysis, referring to standard textbooks. One of the main points is a very tight relationship between a special class of algebraic groups, the reductive groups, and compact Lie groups. We plan to illustrate this in the classical examples, leaving the general theory to Chapter 10.

1 Characters

1.1 Characters

We want to deduce some of the basic theory of characters of finite groups and more generally, compact and reductive groups. We start from some general facts, valid for any group.

Definition. Given a linear representation $\rho : G \to GL(V)$ of a group G, where V is a finite-dimensional vector space over a field F, we define its *character* to be the following function on G:[48]

$$\chi_\rho(g) := \mathrm{tr}(\rho(g)).$$

Here tr is the usual trace. We say that a character is irreducible if it comes from an irreducible representation.

Some properties are immediate (cf. Chapter 6, §1.1).

[48] There is also a deep theory for infinite-dimensional representations. In this setting the trace of an operator is not always defined. With some analytic conditions a character may also be a distribution.

Proposition 1.

(1) $\chi_\rho(g) = \chi_\rho(aga^{-1})$, $\forall a, g \in G$. *The character is constant on conjugacy classes. Such a function is called a* ***class function****.*

(2) Given two representations ρ_1, ρ_2 *we have*

$$\chi_{\rho_1 \oplus \rho_2} = \chi_{\rho_1} + \chi_{\rho_2}, \quad \chi_{\rho_1 \otimes \rho_2} = \chi_{\rho_1} \chi_{\rho_2}. \tag{1.1.1}$$

(3) If ρ *is unitarizable, the character of the dual representation* ρ^* *is the conjugate of* χ_ρ*:*

$$\chi_{\rho^*} = \chi_{\overline{\rho}}. \tag{1.1.2}$$

Proof. Let us prove (3) since the others are clear. If ρ is unitarizable, there is a basis in which the matrices $A(g)$ of $\rho(g)$ are unitary. In the dual representation and in the dual basis the matrix $A^*(g)$ of $\rho^*(g)$ is the transposed inverse of $A(g)$. Under our assumption $A(g)$ is unitary, hence $(A(g)^{-1})^t = \overline{A(g)}$ and $\chi_{\rho^*}(g) = \mathrm{tr}(A^*(g)) = \mathrm{tr}(\overline{A(g)}) = \overline{\mathrm{tr}(A(g))} = \overline{\chi_\rho(g)}$. □

We have just seen that characters can be added and multiplied. Sometimes it is convenient to extend the operations to include the difference $\chi_1 - \chi_2$ of two characters. Of course such a function is no longer necessarily a character but it is called a *virtual character.*

Proposition 2. *The virtual characters of a group G form a commutative ring called the character ring of G.*

Proof. This follows immediately from 1.1.1. □

Of course, if the group G has extra structure, we may want to restrict the representations, hence the characters, to be compatible with the structure. For a topological group we will restrict to continuous representations, while restricting to rational ones for algebraic groups. We will thus speak of continuous or rational characters.

In each case the class of representations is closed under direct sum and tensor product. Thus we also have a character ring, made by the virtual continuous (respectively, algebraic) characters.

Example. In Chapter 7, §3.3 we have seen that for a torus T of dimension n, the (rational) irreducible characters are the elements of a free abelian group of rank n. Thus the character ring is the ring of Laurent polynomials $\mathbb{Z}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$. Inside this ring the characters are the polynomials with nonnegative coefficients.

1.2 Haar Measure

In order to discuss representations and characters for compact groups we need some basic facts from the theory of integration on groups.

The type of measure theory needed is a special case of the classical approach to the Daniell integral (cf. [DS]).

Let X be a locally compact topological space. We use the following notation: $C_0(X, \mathbb{R})$ denotes the algebra of real-valued continuous functions with compact support, while $C_0(X)$ and $C(X)$ denote the complex-valued continuous functions with compact support, respectively, all continuous functions. If X is compact, every function has compact support, hence we drop the subscript 0.

Definition. An integral on X is a nonzero linear map $I : C_0(X, \mathbb{R}) \to \mathbb{R}$, such that if $f \in C_0(X, \mathbb{R})$ and $f(x) \geq 0$, $\forall x \in X$ (a positive function) we have $I(f) \geq 0$.[49]

If $X = G$ is a topological group, we say that an integral I is left-invariant if, for every function $f(x) \in C_0(X, \mathbb{R})$ and every $g \in G$, we have $I(f(x)) = I(f(g^{-1}x))$.

Measure theory allows us to extend a Daniell integral to larger classes of functions, in particular to the characteristic functions of *measurable sets*, and hence deduce a measure theory on X in which all closed and open sets are measurable. This measure theory is essentially equivalent to the given integral. Therefore one uses often the notation dx for the measure and $I(f) = \int f(x)dx$ for the integral.

In the case of groups the measure associated to a left-invariant integral is called a *left-invariant Haar measure.*

In our treatment we will mostly use L^2 functions on X. They form a Hilbert space $L^2(X)$, containing as a dense subspace the space $C_0(X)$; the Hermitian product is $I(f(x)\overline{g}(x))$. A basic theorem (cf. [Ho]) states that:

Theorem. *On a locally compact topological group G, there is a left-invariant measure called the Haar measure.*

The Haar measure is unique up to a scale factor.

This means that if I and J are two left-invariant integrals, there is a positive constant c with $I(f) = cJ(f)$ for all functions.

Exercise. If I is a left-invariant integral on a group G and f a nonzero positive function, we have $I(f) > 0$.

When G is compact, the Haar measure is usually normalized so that the volume of G is 1, i.e., $I(1) = 1$. Of course G also has a right-invariant Haar measure. In general the two measures are not equal.

Exercise. Compute the left and right-invariant Haar measure for the two-dimensional Lie group of affine transformations of $\mathbb{R}$, $x \mapsto ax + b$.

If $h \in G$ and we are given a left-invariant integral $\int f(x)$, it is clear that $f \mapsto \int f(xh)$ is still a left-invariant integral, so it equals some multiple $c(h) \int f(x)$. The function $c(h)$ is immediately seen to be a continuous multiplicative character with values positive numbers.

Proposition 1. *For a compact group, the left and right-invariant Haar measures are equal.*

[49] The axioms of the Daniell integral in this special case are simple consequences of these hypotheses.

Proof. Since G is compact, $c(G)$ is a bounded set of positive numbers. If for some $h \in G$ we had $c(h) \neq 1$, we then have $\lim_{n\to\infty} c(h^n) = \lim_{n\to\infty} c(h)^n$ is 0 or ∞, a contradiction. □

We need only the Haar measure on Lie groups. Since Lie groups are differentiable manifolds one can use the approach to integration on manifolds using differential forms (cf. [Spi]). In fact, as for vector fields, one can find $n = \dim G$ left-invariant differential linear forms ψ_i, which are a basis of the cotangent space at 1 and so too at each point.

Proposition 2. *The exterior product $\omega := \psi_1 \wedge \psi_2 \wedge \ldots \wedge \psi_n$ is a top-dimensional differential form which is left-invariant and defines the volume form for an invariant integration.*

Proof. Take a left translation L_g. By hypothesis $L_g^*(\psi_i) = \psi_i$ for all i. Since L_g^* preserves the exterior product we have that ω is a left-invariant form. Moreover, since the ψ_i are a basis at each point, ω is nowhere 0. Hence ω defines an orientation and a measure on G, which is clearly left-invariant. □

1.3 Compact Groups

The Haar measure on a compact group allows us to average functions, thus getting projections to invariants. Recall that for a representation V of G, the space of invariants is denoted by V^G.

Proposition 1. *Let $\rho : G \to GL(V)$ be a continuous complex finite-dimensional representation of a compact group G (in particular a finite group). Then (using Haar measure),*

$$\dim_{\mathbb{C}} V^G = \int_G \chi_\rho(g)dg. \tag{1.3.1}$$

Proof. Let us consider the operator $\pi := \int \rho(g)dg$. We claim that it is the projection operator on V^G. In fact, if $v \in V^G$,

$$\pi(v) = \int_G \rho(g)(v)dg = \int_G v\, dg = v.$$

Otherwise,

$$\rho(h)\pi(v) = \int_G \rho(h)\rho(g)v\, dg = \int_G \rho(hg)v\, dg = \pi(v)$$

by left invariance of the Haar integral.

We then have $\dim_{\mathbb{C}} V^G = \mathrm{tr}(\pi) = \mathrm{tr}(\int_G \rho(g)dg) = \int_G \mathrm{tr}(\rho(g))dg = \int_G \chi_{\rho(g)}dg$ by linearity of the trace and of the integral. □

The previous proposition has an important consequence.

Theorem 1 (Orthogonality of characters). *Let χ_1, χ_2 be the characters of two irreducible representations ρ_1, ρ_2 of a compact group G. Then*

$$\int_G \chi_1(g)\overline{\chi}_2(g)dg = \begin{cases} 0 & \text{if } \rho_1 \neq \rho_2 \\ 1 & \text{if } \rho_1 = \rho_2 \end{cases}. \tag{1.3.2}$$

Proof. Let V_1, V_2 be the spaces of the two representations. Consider $\hom(V_2, V_1) = V_1 \otimes V_2^*$. As a representation $V_1 \otimes V_2^*$ has character $\chi_1(g)\overline{\chi}_2(g)$, from 1.1.1 and 1.1.2.

We have seen that $\hom_G(V_2, V_1) = (V_1 \otimes V_2^*)^G$, hence, from the previous proposition, $\dim_{\mathbb{C}} \hom_G(V_2, V_1) = \int_G \chi_1(g)\overline{\chi}_2(g)dg$. Finally, by Schur's lemma and the fact that V_1 and V_2 are irreducible, $\hom_G(V_2, V_1)$ has dimension 0 if $\rho_1 \neq \rho_2$ and 1 if they are equal. The theorem follows. □

In fact a more precise theorem holds. Let us consider the Hilbert space of L^2 functions on G. Inside we consider the subspace $L^2_c(G)$ of class functions, which is clearly a closed subspace. Then:

Theorem 2. *The irreducible characters are an orthonormal basis of $L^2_c(G)$.*

Let us give the proof for finite groups. The general case requires some basic functional analysis and will be discussed in Section 2.4. For a finite group G decompose the group algebra in matrix blocks according to Chapter 6, §2.6.3 as $\mathbb{C}[G] = \bigoplus_i^m M_{h_i}(\mathbb{C})$.

The m blocks correspond to the m irreducible representations. Their irreducible characters are the composition of the projection to a factor $M_{h_i}(\mathbb{C})$ followed by the ordinary trace.

A function $f = \sum_{g\in G} f(g)g \in \mathbb{C}[G]$ is a class function if and only if $f(ga) = f(ag)$ or $f(a) = f(gag^{-1})$, for all $a, g \in G$. This means that f lies in the center of the group algebra.

The space of class functions is identified with the center of $\mathbb{C}[G]$.

The center of a matrix algebra $M_h(\mathbb{C})$ is formed by the scalar matrices. Thus the center of $\bigoplus_i^m M_{h_i}(\mathbb{C})$ equals $\mathbb{C}^{\oplus m}$.

It follows that the number of irreducible characters equals the dimension of the space of class functions. Since the irreducible characters are orthonormal they are a basis.

As a corollary we have:

Corollary. *(a) The number of irreducible representations of a finite group G equals the number of conjugacy classes in G.*

(b) If $h_1, \dots, h_r$ are the dimensions of the distinct irreducible representations of G, we have $|G| = \sum_i h_i^2$.

Proof. (a) Since a class function is a function which is constant on conjugacy classes, a basis for class functions is given by the characteristic functions of conjugacy classes.

(b) This is just the consequence of 2.6.3 of Chapter 6. □

There is a deeper result regarding the dimensions of irreducible representations (see [CR]):

Theorem 3. *The dimension h of an irreducible representation of a finite group G divides the order of G.*

The previous theorem allows us to compute a priori the dimensions h_i in some simple cases but in general this is only a small piece of information.

We need one more general result on unitary representations which is a simple consequence of the definitions.

Proposition 2. *Let V be a Hilbert space and a unitary representation of a compact group G. If V_1, V_2 are two non-isomorphic irreducible G-submodules of V, they are orthogonal.*

Proof. The Hermitian pairing (u, v) induces a G-equivariant, antilinear map $j : V_2 \to V_1^*$, $j(u)(v) = (v, u)H$. Since G acts by unitary operators, $V_1^* = \overline{V_1}$. Thus j can be interpreted as a linear G-equivariant map between V_2 and V_1. Since these irreducible modules are non-isomorphic we have $j = 0$. □

1.4 Induced Characters

We now perform a computation on induced characters which will be useful when we discuss the symmetric group.

Let G be a finite group, H a subgroup and V a representation of H with character χ_V. We want to compute the character χ of $\mathrm{Ind}_H^G(V) = \bigoplus_{x \in G/H} xV$ (Chapter 1, §3.2.2, 3.2.3). An element $g \in G$ induces a transformation on $\bigoplus_{x \in G/H} xV$ which can be thought of as a matrix in block form. Its trace comes only from the contributions of the blocks xV for which $gxV = xV$, and this happens if and only if $gx \in xH$, which means that the coset xH is a fixed point under g acting on G/H. As usual, we denote by $(G/H)^g$ these fixed points. The condition that $xH \in (G/H)^g$ can also be expressed as $x^{-1}gx \in H$.

If $gxV = xV$, the map g on xV has the same trace as the map $x^{-1}gx$ on V; thus

$$\chi(g) = \sum_{(G/H)^g} \chi_V(x^{-1}gx). \tag{1.4.1}$$

It is useful to transform the previous formula. Let $X_g := \{x \in G | x^{-1}gx \in H\}$. The next assertions are easily verified:

(i) *The set X_g is a union of right cosets $G(g)x$ where $G(g)$ is the centralizer of g in G.*

(ii) *The map $\pi : x \mapsto x^{-1}gx$ is a bijection between the set of such cosets and the intersection of the conjugacy class C_g of g with H.*

Proof. (i) is clear. As for (ii) observe that $x^{-1}gx = (ax)^{-1}gax$ if and only if $a \in C(g)$. Thus the $G(g)$ cosets of X_g are the nonempty fibers of π. The image of π is clearly the intersection of the conjugacy class C_g of g with H. □

Decompose $C_g \cap H = \cup_i O_i$ into H-conjugacy classes. Of course if $a \in C_g$ we have $|G(a)| = |G(g)|$ since these two subgroups are conjugate. Fix an element $g_i \in O_i$ in each class and let $H(g_i)$ be the centralizer of g_i in H. Then $|O_i| = |H|/|H(g_i)|$ and finally

(1.4.2)
$$\chi(g) = \frac{1}{|H|}\sum_{x\in X}\chi_V(x^{-1}gx) = \frac{1}{|H|}\sum_i\sum_{a\in O_i}|G(a)|\chi_V(a) = \sum_i \frac{|G(g)|}{|H(g_i)|}\chi_V(g_i).$$

In particular one can apply this to the case $V = 1$. This is the example of the permutation representation of G on G/H.

Proposition. *The number of fixed points of g on G/H equals the character of the permutation representation $\mathbb{C}[G/H]$ and is*

$$\chi(g) = \frac{|C_g \cap H||G(g)|}{|H|} = \sum_i \frac{|G(g)|}{|H(g_i)|}. \tag{1.4.3}$$

2 Matrix Coefficients

2.1 Representative Functions

Let G be a topological group. We have seen in Chapter 6, §2.6 the notion of matrix coefficients for G. Given a continuous representation $\rho : G \to GL(U)$ we have a linear map $i_U : \mathrm{End}(U) \to C(G)$ given by $i_U(X)(g) := \mathrm{tr}(X\rho(g))$. We want to return to this concept in a more systematic way.

We will use the following simple fact, which we leave as an exercise. X is a set, F a field.

Lemma. *The n functions $f_i(x)$ on a set X, with values in F, are linearly independent if and only if there exist n points $p_1, \dots, p_n \in X$ with the determinant of the matrix $f_i(p_j)$ nonzero.*

Lemma-Definition. *For a continuous function $f \in C(G)$ the following are equivalent:*

(1) The space spanned by the left translates $f(gx)$, $g \in G$ is finite dimensional.
(2) The space spanned by the right translates $f(xg)$, $g \in G$ is finite dimensional.
(3) The space spanned by the bitranslates $f(gxh)$, $g, h \in G$ is finite dimensional.
(4) There is a finite expansion $f(xy) := \sum_{i=1}^k u_i(x)v_i(y)$.

A function satisfying the previous conditions is called a ***representative function****.*

(5) Moreover, in the expansion (4) the functions u_i, v_i can be taken as representative functions.

Proof. Assume (1) and let $u_i(x),\ i = 1, \dots, m$, be a basis of the space spanned by the functions $f(gx),\ g \in G$.

Write $f(gx) = \sum_i v_i(g)u_i(x)$. This expression is continuous in g. By the previous lemma we can find m points p_j such that the determinant of the matrix with entries $u_i(p_j)$ is nonzero.

Thus we can solve the system of linear equations $f(gp_j) = \sum_i v_i(g)u_i(p_j)$ by Cramer's rule, so that the coefficients $v_i(g)$ are continuous functions, and (4) follows. (4) is a symmetric property and clearly implies (1) and (2).

In the expansion $f(xy) := \sum_{i=1}^k u_i(x)v_i(y)$ we can take the functions v_i to be a basis of the space spanned by the left translates of f. They are representative functions. We have that

$$f(xzy) := \sum_{i=1}^k u_i(xz)v_i(y) = \sum_{i=1}^k u_i(x)v_i(zy) = \sum_{i=1}^k u_i(x)\sum_{h=1}^k c_{i,h}(z)v_h(y)$$

implies $u_i(xz) = \sum_{h=1}^k u_h(x)c_{hi}(z)$ implying (5) and also (3). □

Proposition 1. *The set $\mathcal{T}_G$ of representative functions is an algebra spanned by the matrix coefficients of the finite-dimensional continuous representations of G.*

Proof. The fact that it is an algebra is obvious. Let us check the second statement. First, a continuous finite-dimensional representation is given by a homomorphism $\rho : G \to GL(n, \mathbb{C})$. The entries $\rho(g)_{i,j}$ by definition span the space of the corresponding matrix coefficients. We have that $\rho(xy) = \rho(x)\rho(y)$, which in matrix entries shows that the functions $\rho(g)_{i,j}$ satisfy (4).

Conversely, let $f(x)$ be a representative function. Clearly also $f(x^{-1})$ is representative. Let $u_i(x)$ be a basis of the space U of left translates, $f(g^{-1}) = \sum_{i=1}^k a_i u_i(g)$. U is a linear representation by the left action and $u_i(g^{-1}x) = \sum_j c_{ij}(g)u_j(x)$ where the functions $c_{i,j}(g)$ are the matrix coefficients of U in the given basis. We thus have $u_i(g^{-1}) = \sum_j c_{ij}(g)u_j(1)$ and $f(g) = \sum_{i=1}^k a_i \sum_j c_{ij}(g)u_j(1)$.

If G, K are two topological groups, we have that

Proposition 2. *Under multiplication $f(x)g(y)$ we have an isomorphism*

$$\mathcal{T}_G \otimes \mathcal{T}_K = \mathcal{T}_{G\times K}.$$

Proof. The multiplication map of functions on two distinct spaces to the product space is always an isomorphism of the tensor product of the spaces of functions to the image. So we only have to prove that the space of representative functions of $G \times K$ is spanned by the functions $\psi(x, y) := f(x)g(y),\ f(x) \in \mathcal{T}_G,\ g(y) \in \mathcal{T}_K$.

Using the property (4) of the definition of representative function we have that if $f(x_1x_2) = \sum_i u_i(x_1)v_i(x_2),\ g(y_1, y_2) = \sum_k w_k(y_1)z_k(y_2)$, then

$$\psi((x_1, y_1)(x_2, y_2)) = \sum_{i,k} u_i(x_1)w_k(y_1)v_i(x_2)z_k(y_2).$$

Conversely, if $\psi(x, y)$ is representative, writing $(x, y) = (x, 1)(1, y)$ one immediately sees that ψ is in the span of the product of representative functions. □

Finally, let $\rho : H \to K$ be a continuous homomorphism of topological groups.

Proposition 3. *If $f(k)$ is representative on K, then $f(\rho(k))$ is representative in H.*

Proof. We have $f(xy) = \sum_i u_i(x)v_i(y)$, hence $f(\rho(ab)) = f(\rho(a)\rho(b)) = \sum_i u_i(\rho(a))v_i(\rho(b))$. □

In terms of matrix coefficients what we are doing is to take a representation of K and deduce, by composition with ρ, a representation of H.

Particularly important for us will be the case of a compact group K, when all the finite-dimensional representations are semisimple. We then have an analogue of

Theorem. *The space $\mathcal{T}_K$ is the direct sum of the matrix coefficients $V_i^* \otimes V_i$ as $V_i \in \hat{K}$ runs on the set of different irreducible representations of K.*

$$\mathcal{T}_K = \bigoplus_{V \in \hat{K}} V^* \otimes V. \tag{2.1.1}$$

Proof. The proof is essentially identical to that of Chapter 7, §3.1. □

2.2 Preliminaries on Functions

Before we continue our analysis we wish to collect two standard results on function theory which will be useful in the sequel. The first is the Stone–Weierstrass theorem. This theorem is a generalization of the classical theorem of Weierstrass on approximation of continuous functions by polynomials.

In its general form it says:

Stone–Weierstrass theorem. *Let A be an algebra of real-valued continuous functions on a compact space X which separates points.*[50] *Then either A is dense in $C(X, \mathbb{R})$ or it is dense in the subspace of $C(X, \mathbb{R})$ of functions vanishing at a given point a.*[51]

Proof. Let A be such an algebra. If $1 \in A$, then we cannot be in the second case, where $f(a) = 0, \forall f \in A$. Otherwise we can add 1 to A and assume that $1 \in A$. Then let S be the uniform closure of A. The theorem can thus be reformulated as follows: if S is an algebra of continuous functions, which separates points, $1 \in S$, and S is closed under uniform convergence, then $S = C_0(X, \mathbb{R})$.

We will use only one statement of the classical theorem of Weierstrass, the fact that given any interval $[-n, n]$, the function $|x|$ can be uniformly approximated by polynomials in this interval. This implies for our algebra S that if $f(x) \in S$, then $|f(x)| \in S$. From this we immediately see that if f and $g \in S$, the two functions $\min(f, g) = (f + g - |f - g|)/2$ and $\max(f, g) = (f + g + |f - g|)/2$ are in S.

Let x, y be two distinct points in X. By assumption there is a function $a \in S$ with $a(x) \neq a(y)$. Since the function $1 \in S$ takes the value 1 at x and y, we can

[50] This means that, given $a, b \in X$, $a \neq b$, there is an $f \in A$ with $f(a) \neq f(b)$.

[51] If $X = \{p_1, \dots, p_n\}$ is a finite set, the theorem is really a theorem of algebra, a form of the Chinese Remainder Theorem.

find a linear combination g of a, 1 which takes at x, y any prescribed values. Let $f \in C_0(X, \mathbb{R})$ be a function. By the previous remark we can find a function $g_{x,y} \in S$ with $g_{x,y}(x) = f(x)$, $g_{x,y}(y) = f(y)$. Given any $\epsilon > 0$ we can thus find an open set U_y such that $g_{x,y}(z) > f(z) - \epsilon$ for all $z \in U_y$. By compactness of X we can find a finite number of such open sets U_{y_i} covering X. Take the corresponding functions g_{x,y_i}. We have that the function $g_x := \max(g_{x,y_i}) \in S$ has the property $g_x(x) = f(x)$, $g_x(z) > f(z) - \epsilon$, $\forall z \in X$. Again there is a neighborhood V_x of x such that $g_x(z) < f(z) + \epsilon$, $\forall z \in V_x$. Cover X with a finite number of these neighborhoods V_{x_j}. Take the corresponding functions g_{x_j}. We have that the function $g := \min(g_{x_j}) \in S$ has the property $|g(z) - f(z)| < \epsilon$, $\forall z \in X$. Letting ϵ tend to 0, since S is closed under uniform convergence, we find that $f \in S$, as desired. □

We will often apply this theorem to an algebra A of complex functions. In this case we easily see that the statement is:

Corollary. *If $A \subset C(X)$ is an algebra of complex functions which separates points in X, $1 \in A$, A is closed under uniform convergence and A is closed under complex conjugation, then $A = C(X)$.*

For the next theorem we need to recall two simple notions. These results can be generalized but we prove them in a simple case.

Definition. A set A of continuous functions on a space X is said to be *uniformly bounded* if there is a positive constant M such that $|f(x)| < M$ for every $f \in A, x \in X$.

A set A of continuous functions on a metric space X is said to be *equicontinuous* if, for every $\epsilon > 0$, there is a $\delta > 0$ with the property that $|f(x) - f(y)| < \epsilon, \forall (x, y)$ with $\overline{xy} < \delta$ and $\forall f \in A$.

We are denoting by $\overline{xy}$ the distance between the two points x, y.

Recall that a topological space is *first countable* if it has a dense countable subset.

Theorem (Ascoli–Arzelà). *A uniformly bounded and equicontinuous set A of continuous functions on a first countable compact metric space X is relatively compact in $C(X)$, i.e., from any sequence $f_i \in A$ we may extract one which is uniformly convergent.*

Proof. Let $p_1, p_2, \dots, p_k, \dots$ be a dense sequence of points in X. Since the functions f_i are uniformly bounded, we can extract a subsequence $s_1 := f_1^1, f_2^1, \dots, f_i^1, \dots$ from the given sequence for which the sequence of numbers $f_i^1(p_1)$ is convergent. Inductively, we construct sequences s_k where s_k is extracted from s_{k-1} and the sequence of numbers $f_i^k(p_k)$ is convergent. It follows that for the *diagonal* sequence $F_i := f_i^i$, we have that the sequence of numbers $F_i(p_j)$ is convergent for each j. We want to show that F_i is uniformly convergent on X. We need to show that F_i is a Cauchy sequence. Given $\epsilon > 0$ we can find by equicontinuity a $\delta > 0$ with the property that $|f(x) - f(y)| < \epsilon, \forall (x, y)$ with $\overline{xy} < \delta$ and $\forall f \in A$. By compactness we can find a finite number of points q_j, $j = 1, \dots, m$, from our list p_i such that for all $x \in X$, there is one of the q_j at a distance less than δ from x. Let k be such that $|F_s(q_j) - F_t(q_j)| < \epsilon, \forall j = 1, \dots, m, \ \forall s, t > k$. For each x find a q_j at a distance less than δ, then

$$\begin{aligned}|F_s(x) - F_t(x)| &= |F_s(x) - F_s(q_j) - F_t(x) + F_t(q_j) + F_s(q_j) - F_t(q_j)| \\ &< 3\epsilon, \quad \forall s, t > k.\end{aligned}$$

□

2.3 Matrix Coefficients of Linear Groups

One possible approach to finding the representations of a compact group could be to identify the representative functions. In general this may be difficult but in a special case it is quite easy.

Theorem. *Let $G \subset U(n, \mathbb{C})$ be a compact linear group. Then the ring of representative functions of G is generated by the matrix entries and the inverse of the determinant.*

Proof. Let A be the algebra of functions generated by the matrix entries and the inverse of the determinant. Clearly $A \subset \mathcal{T}_G$ by Proposition 2.2. Moreover, by matrix multiplication it is clear that the space of matrix entries is stable under left and right G action, similarly for the inverse of the determinant and thus A is $G \times G$-stable.

Let us prove now that A is dense in the algebra of continuous functions. We want to apply the Stone–Weierstrass theorem to the algebra A which is made up of complex functions and contains 1. In this case, besides verifying that A separates points, we also need to show that A is closed under complex conjugation. Then we can apply the previous theorem to the real and imaginary parts of the functions of A and conclude that they are both dense.

In our case A separates points since two distinct matrices must have two different coordinates. A is closed under complex conjugation. In fact the conjugate of the determinant is the inverse, while the conjugates of the entries of a unitary matrix X are entries of X^{-1}. The entries of this matrix, by the usual Cramer rule, are indeed polynomials in the entries of X divided by the determinants, hence are in A.

At this point we can conclude. If $A \neq \mathcal{T}_G$, since they are both $G \times G$ representations and $\mathcal{T}_G = \bigoplus_i V_i^* \otimes V_i$ is a direct sum of irreducible $G \times G$ representations, for some i we have $V_i^* \otimes V_i \cap A = 0$. By Proposition 2 of 1.3 this implies that $V_i^* \otimes V_i$ is orthogonal to A and this contradicts the fact that A is dense in $C(G)$. □

Given a compact Lie group G it is not restrictive to assume that $G \subset U(n, \mathbb{C})$. This will be proved in Section 4.3 as a consequence of the Peter–Weyl theorem.

3 The Peter–Weyl Theorem

3.1 Operators on a Hilbert Space

The representation theory of compact groups requires some basic functional analysis. Let us recall some simple definitions.

Definition 1. A *norm* on a complex vector space V is a map $v \mapsto \|v\| \in \mathbb{R}^+$, satisfying the properties:

$$\|v\| = 0 \iff v = 0, \quad \|av\| = |a|\|v\|, \quad \|v + w\| \leq \|v\| + \|w\|.$$

A vector space with a norm is called a *normed space.*

From a norm one deduces the structure of metric space setting as distance $\overline{xy} := \|x - y\|$.

Definition 2. A Banach space is a normed space complete under the induced metric.

Most important for us are Hilbert spaces. These are the Banach spaces where the norm is deduced from a positive Hermitian form $\|v\|^2 = (v, v)$. When we talk about convergence in a Hilbert space we usually mean in this norm and also speak of *convergence in mean.*[52] All our Hilbert spaces are assumed to be countable, and to have a countable orthonormal basis.

The special properties of Hilbert spaces are the *Schwarz inequality* $|(u, v)| \leq \|u\| \, \|v\|$, and the existence of orthonormal bases u_i with $(u_i, u_j) = \delta_i^j$. Then $v = \sum_{i=1}^{\infty}(v, u_i)u_i$ for every vector $v \in H$, from which $\|v\|^2 = \sum_{i=1}^{\infty} |(v, u_i)|^2$. This is called the *Parseval formula.*[53]

The other Banach space which we will occasionally use is the space $C(X)$ of continuous functions on a compact space X with norm the *uniform norm* $\|f\|_\infty := \max_{x \in X} |f(x)|$. Convergence in this norm is *uniform convergence.*

Definition 3. A linear operator $T : A \to B$ between normed spaces is *bounded* if there is a positive constant C such that $\|T(v)\| \leq C\|v\|, \forall v \in A$.

The minimum such constant is the *operator norm* $\|T\|$ of T.

By linearity it is clear that $\|T\| = \sup_{\|v\|=1} \|T(v)\|$.

Exercise.

1. The sum and product of bounded operators are bounded.

$$\|aT\| = |a|\|T\|, \, \|aT_1 + bT_2\| \leq |a|\|T_1\| + |b|\|T_2\|, \, \|T_1 \circ T_2\| \leq \|T_1\|\|T_2\|.$$

2. If B is complete, bounded operators are complete under the norm $\|T\|$.

When $A = B$ bounded operators on A will be denoted by $\mathcal{B}(A)$. They form an algebra. The previous properties can be taken as the axioms of a *Banach algebra.*

Given a bounded[54] operator T on a Hilbert space, its adjoint T^* is defined by $(Tv, w) = (v, T^*w)$. We are particularly interested in bounded Hermitian operators (or self-adjoint), i.e., bounded operators T for which $(Tu, v) = (u, Tv), \ \forall u, v \in H$.

[52] There are several other notions of convergence but they do not play a role in our work.

[53] For every n we also have $\|v\|^2 \geq \sum_{i=1}^{n} |(v, u_i)|^2$, which is called *Bessel's inequality.*

[54] We are simplifying the theory drastically.

The typical example is the Hilbert space of L^2 functions on a measure space X, with Hermitian product $(f, g) = \int_X f(x)\overline{g(x)}dx$. An important class of bounded operators are the integral operators $Tf(x) := \int_X K(x, y)f(y)\,dy$. The *integral kernel* $K(x, y)$ is itself a function on $X \times X$ with some suitable restrictions (like L^2). If $K(x, y) = \overline{K(y, x)}$ we have a self-adjoint operator.

Theorem 1. *(1) If A is a self-adjoint bounded operator, $\|A\| = \sup_{\|v\|=1} |(Av, v)|$.*
(2) For any bounded operator $\|T\|^2 = \|T^*T\|$.[55]

Proof. (1) By definition if $\|v\| = 1$ we have $\|Av\| \le \|A\|$ hence $|(Av, v)| \le \|A\|$ by the Schwarz inequality.[56] In the self-adjoint case $(A^2v, v) = (Av, Av)$. Set $N := \sup_{\|v\|=1} |(Av, v)|$. If $\lambda > 0$, we have

$$\begin{aligned}
\|Av\|^2 &= \frac{1}{4}\left[\left(A\left(\lambda v + \frac{1}{\lambda}Av\right), \lambda v + \frac{1}{\lambda}Av\right) - \left(A\left(\lambda v - \frac{1}{\lambda}Av\right), \lambda v - \frac{1}{\lambda}Av\right)\right] \\
&\le \frac{1}{4}\left[N\left\|\lambda v + \frac{1}{\lambda}Av\right\|^2 + N\left\|\lambda v - \frac{1}{\lambda}Av\right\|^2\right] \\
&= \frac{N\|v\|^2}{2}\left[\lambda^2 + \frac{1}{\lambda^2}\frac{\|Av\|^2}{\|v\|^2}\right].
\end{aligned}$$

For $Av \neq 0$ the minimum on the right-hand side is obtained when

$$\lambda^2 = \frac{\|Av\|}{\|v\|}, \quad \text{since} \quad \left(\lambda^2 + \frac{1}{\lambda^2}c^2 = \left(\lambda - \frac{c}{\lambda}\right)^2 + 2c \ge 2c\right).$$

Hence

$$\|Av\|^2 \le N\|Av\|\|v\| \implies \|Av\| \le N\|v\|.$$

Of course this holds also when $Av = 0$, hence $\|A\| \le N$.
(2) $\|T\|^2 = \sup_{\|v\|=1}(Tv, Tv) = \sup_{\|v\|=1} |(T^*Tv, v)| = \|T^*T\|$ from 1. □

Recall that for a linear operator T an eigenvector v of eigenvalue $\lambda \in \mathbb{C}$ is a vector with $Tv = \lambda v$. If T is self-adjoint, necessarily $\lambda \in \mathbb{R}$. Eigenvalues are bounded by the operator norm. If λ is an eigenvalue, from $Tv = \lambda v$ we get $\|Tv\| = |\lambda|\|v\|$, hence $\|T\| \ge |\lambda|$.

In general, actual eigenvectors need not exist, the typical example being the operator $f(x) \mapsto g(x)f(x)$ of multiplication by a continuous function on L^2 functions on $[0, 1]$.

Lemma 1. *If A is a self-adjoint operator v, w two eigenvectors of eigenvalues $\alpha \neq \beta$, we have $(v, w) = 0$.*

[55] This property is taken as an axiom for C^* algebras.

[56] This does not need self-adjointness.

Proof.

$$\alpha(v,w)=(Av,w)=(v,Aw)$$
$$=\beta(v,w)\implies(\alpha-\beta)(v,w)=0\implies(v,w)=0.\qquad\square$$

There is a very important class of operators for which the theory better resembles the finite-dimensional theory. These are the *completely continuous operators* or compact operators.

Definition 4. A bounded operator A of a Hilbert space is *completely continuous*, or *compact* if, given any sequence v_i of vectors of norm 1, from the sequence $A(v_i)$ one can extract a convergent sequence.

In other words this means that A transforms the sphere $\|v\|=1$ in a relatively compact set of vectors, where compactness is by convergence of sequences.

We will denote by $\mathcal{I}$ the set of completely continuous operators on H.

Proposition. *$\mathcal{I}$ is a two-sided ideal in $\mathcal{B}(H)$, closed in the operator norm.*

Proof. Suppose that $A=\lim_{i\to\infty}A_i$ is a limit of completely continuous operators A_i. Given a sequence v_i of vectors of norm 1 we can construct by hypothesis for each i, and by induction a sequence $s_i:=(v_{i_1(i)},v_{i_2(i)},\dots,v_{i_k(i)},\dots)$ so that s_i is a subsequence of s_{i-1} and the sequence $A_i(v_{i_1(i)}),A_i(v_{i_2(i)}),\dots,A_i(v_{i_k(i)}),\dots$ is convergent.

We then take the diagonal sequence $w_k:=v_{i_k(k)}$ and see that $A(w_k)$ is a Cauchy sequence. In fact given $\epsilon>0$ there is an N such that $\|A-A_i\|<\epsilon/3$ for all $i\geq N$, there is also an M such that $\|A_N(v_{i_k(N)})-A_N(v_{i_h(N)})\|<\epsilon/3$ for all $h,k>M$. Thus when $h\leq k>\max(N,M)$ we have that $v_{i_k(k)}=v_{i_h(t)}$ for some $t\geq k$, and so

$$\begin{aligned}\|A(w_h)-A(w_k)\|&=\|A(v_{i_h(h)})-A_h(v_{i_h(h)})+A_h(v_{i_h(h)})-A(v_{i_k(k)})\|\\&\leq\|A(v_{i_h(h)})-A_h(v_{i_h(h)})\|+\|A_h(v_{i_h(h)})-A_h(v_{i_h(t)})\|\\&\quad+\|A_h(v_{i_h(t)})-A(v_{i_h(t)})\|<\epsilon.\end{aligned}$$

The property of being a two-sided ideal is almost trivial to verify and we leave it as exercise. $\square$

From an abstract point of view, completely continuous operators are related to the notion of the complete tensor product $H\hat{\otimes}\overline{H}$, discussed in Chapter 5, §3.8. Here $\overline{H}$ is the conjugate space. We want to associate an element $\rho(u)\in\mathcal{I}$ to an element $u\in H\hat{\otimes}\overline{H}$.

The construction is an extension of the algebraic formula 3.4.4 of Chapter 5.[57] We first define the map on the algebraic tensor product as in the formula $\rho(u\otimes v)w:=u(w,v)$. Clearly the image of $\rho(u\otimes v)$ is the space generated by u; hence $\rho(H\otimes\overline{H})$ is made of operators with finite-dimensional image.

[57] We see now why we want to use the conjugate space: it is to have bilinearity of the map ρ.

Lemma 2. *The map $\rho : H \otimes \overline{H} \to \mathcal{B}(H)$ decreases the norms and extends to a continuous map $\rho : H \hat{\otimes} \overline{H} \to \mathcal{I} \subset \mathcal{B}(H)$.*

Proof. Let us fix an orthonormal basis u_i of H. We can write an element $v \in H \otimes \overline{H}$ as a finite sum $v = \sum_{i,j=1}^{m} c_{i,j} u_i \otimes u_j$. Its norm in $H \otimes \overline{H}$ is $\sqrt{\sum_{i,j} |c_{i,j}|^2}$.

$$\rho(v)\left(\sum_h a_h u_h\right) = \sum_h \left(\sum_i c_{i,h} a_h\right) u_i = \sum_i \left(\sum_h a_h c_{i,h}\right) u_i.$$

Given $w := \sum_h a_h u_h$ we deduce from the Schwarz inequality for $\|\rho(v)(w)\|$ that

$$\begin{aligned}\left\|\rho(v)\left(\sum_h a_h u_h\right)\right\| &= \sqrt{\sum_i \left|\left(\sum_h a_h c_{i,h}\right)\right|^2} \\ &\leq \sqrt{\sum_i \left(\sum_h |a_h|^2\right)\left(\sum_h |c_{i,h}|^2\right)} = \|w\|\,\|v\|.\end{aligned}$$

Since the map decreases norms it extends by continuity. Clearly bounded operators with finite range are completely continuous. From the previous proposition, limits of these operators are also completely continuous. □

In fact we will see presently that the image of ρ is dense in $\mathcal{I}$, i.e., that every completely continuous operator is a limit of operators with finite-dimensional image.

Warning. The image of ρ is not $\mathcal{I}$. For instance the operator T, which in an orthonormal basis is defined by $T(e_i) := \frac{1}{\sqrt{i}} e_i$, is not in $\mathrm{Im}(\rho)$.

The main example of the previous construction is given by taking $H = L^2(X)$ with X a space with a measure. We recall a basic fact of measure theory (cf. [Ru]). If X, Y are measure spaces, with measures $d\mu, d\nu$, one can define a product measure $d\mu \times d\nu$ on $X \times Y$. If $f(x), g(y)$ are L^1 functions on X, Y respectively, we have that $f(x)g(y)$ is L^1 on $X \times Y$ and

$$\int_{X \times Y} f(x)g(y)\,d\mu d\nu = \int_X f(x)\,d\mu \int_Y g(y)\,d\nu.$$

Lemma 3. *The map $i : f(x) \otimes g(y) \mapsto f(x)g(y)$ extends to a Hilbert space isomorphism $L^2(X)\hat{\otimes}L^2(Y) = L^2(X \times Y)$.*

Proof. We have clearly that the map i is well defined and preserves the Hermitian product; hence it extends to a Hilbert space isomorphism of $L^2(X)\hat{\otimes}L^2(Y)$ with some closed subspace of $L^2(X \times Y)$. To prove that it is surjective we use the fact that, given measurable sets $A \subset X$, $B \subset Y$ of finite measure, the characteristic function $\chi_{A \times B}$ of $A \times B$ is the tensor product of the characteristic functions of A and B. By standard measure theory, since the sets $A \times B$ generate the σ-algebra of measurable sets in $X \times Y$, the functions $\chi_{A \times B}$ span a dense subspace of $L^2(X \times Y)$. □

Proposition 1. *An integral operator $Tf(x) := \int_X K(x,y)f(y)dy$, with the integral kernel in $L^2(X \times X)$, is completely continuous.*

Proof. By the previous lemma, we can write $K(x,y) = \sum_{i,j} c_{i,j}u_i(x)\bar{u}_j(y)$ with $u_i(x)$ an orthonormal basis of $L^2(X)$. We see that $Tf(x) = \sum_{i,j} c_{i,j}u_i(x)\int_X \bar{u}_j(y)$ $f(y)dy = \sum_{i,j} c_{i,j}u_i(f,u_j)$. Now we apply Lemma 2. □

We will also need a variation of this theme. Assume now that X is a locally compact metric space with a Daniell integral. Assume further that the integral kernel $K(x,y)$ is continuous and has compact support.

Proposition 2. *The operator $Tf(x) := \int_X K(x,y)f(y)dy$ is a bounded operator from $L^2(X)$ to $C_0(X)$. It maps bounded sets of functions into uniformly bounded and equicontinuous sets of continuous functions.*[58]

Proof. Assume that the support of the kernel is contained in $A \times B$ with A, B compact. Let m be the measure of B. First, $Tf(x)$ is supported in A, and it is a continuous function. In fact if $x \in A$, by compactness and continuity of $K(x,y)$, there is a neighborhood U of x such that $|K(x,y) - K(x_0,y)| < \epsilon$, $\forall y \in B, \forall x \in U$ so that (Schwarz inequality)

(3.1.1)

$$\forall x \in U, \quad |Tf(x) - Tf(x_0)| \le \int_B |K(x,y) - K(x_0,y)|\,|f(y)|dy \le \epsilon\, m^{1/2}\,\|f\|.$$

Moreover, if $M = \max|K(x,y)|$ we have

$$\|T(f)\|_\infty = \sup\left(\left|\int_X K(x,y)f(y)dy\right|\right)$$
$$\le \sup\left(\sqrt{\int_X |K(x,y)|^2 dy}\right)\|f\| \le m^{1/2}M\|f\|.$$

Let us show that the functions $Tf(x)$, $\|f\| = 1$ are equicontinuous and uniformly bounded. In fact $|Tf(x)| \le m^{1/2}M$ where $M = \max|K(x,y)|$. The equicontinuity follows from the previous argument. Given $\epsilon > 0$ we can, by the compactness of $A \times B$, find $\eta > 0$ so that $|K(x_1,y) - K(x_0,y)| < \epsilon$ if $\overline{x_1x_0} < \eta, \forall y \in B$. Hence if $\|f\| \le M$, we have $|Tf(x_1) - Tf(x_0)| \le Mm^{1/2}\epsilon$ when $\overline{x_1x_0} < \eta$. □

Proposition 3. *If A is a self-adjoint, completely continuous operator, there is an eigenvector v with eigenvalue $\pm\|A\|$.*

Proof. By Theorem 1, there is a sequence of vectors v_i of norm 1 for which $\lim_{i\to\infty}(Av_i, v_i) = \mu = \pm\|A\|$. By hypothesis we can extract a subsequence, which

[58] If X is not compact, $C_0(X)$ is not complete, but in fact T maps into the complete subspace of functions with support in a fixed compact subset $A \subset X$.

we still call v_i, such that $\lim_{i\to\infty} A(v_i) = w$. Since $\mu := \lim_{i\to\infty}(A(v_i), v_i)$, the inequality

$$0 \leq \|Av_i - \mu v_i\|^2 = \|Av_i\|^2 - 2\mu(Av_i, v_i) + \mu^2 \leq 2\mu^2 - 2\mu(Av_i, v_i)$$

implies that $\lim_{i\to\infty}(A(v_i) - \mu v_i) = 0$. Thus $\lim_{i\to\infty} \mu v_i = \lim_{i\to\infty} A(v_i) = w$. In particular v_i must converge to some vector v such that $w = \mu v$, and $w = \lim_{i\to\infty} Av_i = Av$. Since $\mu = (Aw, w)$ if $A \neq 0$ we have $\mu \neq 0$, hence $v \neq 0$ is the required eigenvector. □

Given a Hilbert space H an orthogonal decomposition for H is a family of mutually orthogonal closed subspaces H_i, $i = 1, \dots, \infty$, such that every element $v \in H$ can be expressed (in a unique way) as a series $v = \sum_{i=1}^{\infty} v_i$, $v_i \in H_i$. An orthogonal decomposition is a generalization of the decomposition of H given by an orthonormal basis.

Definition 5. A self-adjoint operator is *positive* if $(Av, v) \geq 0, \forall v$.

Remark. If T is any operator T^*T is positive self-adjoint. The eigenvalues of a positive operator are all positive or 0.

Theorem 2. *Let A be a self-adjoint, completely continuous operator.*

If the image of A is not finite dimensional, there is a sequence of numbers λ_i and orthonormal vectors v_i such that

$$\|A\| = |\lambda_1| \geq |\lambda_2| \geq \cdots |\lambda_n| \geq \cdots$$

and

(1) Each λ_i is an eigenvalue of A, $Av_i = \lambda_i v_i$.

(2) The numbers λ_i are the only nonzero eigenvalues.

(3) $\lim_{i\to\infty} \lambda_i = 0$.

(4) The eigenspace H_i of each eigenvalue $\lambda \neq 0$ is finite dimensional with basis the v_i for which $\lambda_i = \lambda$.

(5) H is the orthogonal sum of the kernel of A and the subspace with basis the orthonormal vectors v_i.

If the image of A is finite dimensional, the sequence $\lambda_1, \lambda_2, \dots, \lambda_m$ is finite and $H = \mathrm{Ker}(A) \bigoplus_{i=1}^{m} \mathbb{C}v_i$.

Proof. Call $A = A_1$; by Proposition 3 there is an eigenvector v_1, of absolute value 1, with eigenvalue λ_1 and $|\lambda_1| = \|A_1\| > 0$. Decompose $H = \mathbb{C}v_1 \oplus v_1^{\perp}$, the operator A_1 induces on $v_1^{\perp}$ a completely continuous operator A_2 with $\|A_2\| \leq \|A_1\|$. Repeating the reasoning for A_2 there is an eigenvector v_2, $|v_2| = 1$ with eigenvalue λ_2 and $|\lambda_2| = \|A_2\| > 0$.

Continuing in this way we find a sequence of orthonormal vectors $v_1, \dots, v_k, \dots$ with $Av_i = \lambda_i v_i$ and such that, setting $H_i := (\mathbb{C}v_1 + \cdots + \mathbb{C}v_{i-1})^{\perp}$, we have that $|\lambda_i| = \|A_i\|$ where A_i is the restriction of A to H_i.

This sequence stops after finitely many steps if for some i we have that $A_i = 0$; this implies that the image of A is spanned by $v_1, \dots, v_{i-1}$, otherwise it continues indefinitely. We must then have $\lim_{i\to\infty}\lambda_i = 0$, otherwise there is a positive constant $0 < b < |\lambda_i|$, for all i. In this case no subsequence of the sequence $Av_i = \lambda_i v_i$ can be chosen to be convergent since these vectors are orthogonal and all of absolute value $> b$, which contradicts the hypothesis that A is compact. Let $\overline{H}$ be the Hilbert subspace with basis the vectors v_i and decompose $H = \overline{H} \oplus \overline{H}^{\perp}$. On $\overline{H}^{\perp}$ the restriction of A has a norm $\leq |\lambda_i|$ for all i, hence it must be 0 and $\overline{H}^{\perp}$ is the kernel of A. Now given a vector $v = \sum_i c_i v_i + u$, $u \in \overline{H}^{\perp}$, we have $Av = \sum_i c_i\lambda_i v_i$; thus v is an eigenvector if and only if either $v = u$ or $u = 0$ and all the indices i for which $c_i \neq 0$ have the same eigenvalue λ. Since $\lim \lambda_i = 0$, this can happen only for a finite number of these indices. This proves also (2), (4), and finishes the proof. □

Exercise 2. Prove that $\mathcal{I}$ is the closure in the operator norm of the operators of finite-dimensional image. *Hint*: Use the spectral theory of T^*T and Exercise 1.

Let us now specialize to an integral operator $Tf(x) := \int_X K(x, y)f(y)dy$ with the integral kernel continuous and with compact support in $A \times B$ as before. Suppose further that T is self-adjoint, i.e., $K(x, y) = \overline{K(y, x)}$.

By Proposition 2, the eigenvectors of T corresponding to nonzero eigenvalues are continuous functions. Let us then take an element $f = \sum_{i=1}^{\infty} c_i u_i + u$ expanded, as before, in an orthonormal basis of eigenfunctions u_1, u with $Tu_i = \lambda_i u_i$, $Tu = 0$. The u_i are continuous functions with support in the compact set A.

Proposition 4. *The sequence* $g_k := T\left(\sum_{i=1}^{k} c_i u_i\right) = \sum_{i=1}^{k} c_i\lambda_i u_i$ *of continuous functions converges uniformly to* Tf.

Proof. By continuity g_k converges to Tf in the L^2-norm.

But g_k is also a Cauchy sequence in the uniform norm, as follows from the continuity of the operator from $L^2(X)$ to $C_0(X)$ (for the L^2 and uniform norm, respectively) so it converges uniformly to some function g. The inclusion $C_0(X) \subset L^2(X)$, when restricted to the functions with support in A, is continuous for the two norms ∞ and L^2 (since $\int_X \|f\|^2 d\mu \leq \mu(A)\|f\|_{\infty}^2$, where μ_a is the measure of A); therefore we must have $g = Tf$. □

We want to apply the theory to a locally compact group G with a left-invariant Haar measure. This measure allows us to define the *convolution* product, which is the generalization of the product of elements of the group algebra.

The convolution product is defined first of all on the space of L^1-functions by the formula

$$(3.1.2) \qquad (f * g)(x) := \int_G f(y)g(y^{-1}x)dy = \int_G f(xy)g(y^{-1})dy.$$

When G is compact we normalize Haar measure so that the measure of G is 1. We have the continuous inclusion maps

$$(3.1.3) \qquad C_0(G) \subset L^2(G) \subset L^1(G).$$

The three spaces have respectively the uniform L^∞, L^2, L^1 norms; the inclusions decrease norms. In fact the L^1 norm of f equals the Hilbert scalar product of $|f|$ with 1, so by the Schwarz inequality, $|f|_1 \leq |f|_2$ while $|f|_2 \leq |f|_\infty$ for obvious reasons.

Proposition 5. *If G is compact, then the space of L^2 functions is also an algebra under convolution.*[59]

Both algebras $L^1(G)$, $L^2(G)$ are useful. In the next section we shall use $L^2(G)$, and we will compute its algebra structure in §3.3. On the other hand, $L^1(G)$ is also useful for representation theory.

One can pursue the algebraic relationship between group representations and modules over the group algebra in the continuous case, replacing the group algebra with the convolution algebra (cf. [Ki], [Di]).

3.2 Peter–Weyl Theorem

Theorem (Peter–Weyl).

(i) The direct sum $\bigoplus_i V_i^* \otimes V_i$ equals the space $\mathcal{T}_G$ of representative functions.
(ii) The direct sum $\bigoplus_i V_i^* \otimes V_i$ is dense in $L^2(G)$.

In other words every L^2-function f on G can be developed uniquely as a series $f = \sum_i u_i$ with $u_i \in V_i^* \otimes V_i$.

Proof. (i) We have seen (Chapter 6, Theorem 2.6) that for every continuous finite-dimensional irreducible representation V of G, the space of matrix coefficients $V^* \otimes V$ appears in the space $C(G)$ of continuous functions on G. Every finite-dimensional continuous representation of G is semisimple, and the matrix coefficients of a direct sum are the sum of the respective matrix coefficients.

(ii) For distinct irreducible representations V_1, V_2, the corresponding spaces of matrix coefficients, are irreducible non-isomorphic representations of $G \times G$. We can thus apply Proposition 2 of 1.3 to deduce that they are orthogonal.

Next we must show that the representative functions are dense in $C(G)$. For this we take a continuous function $\phi(x)$ with $\phi(x) = \phi(x^{-1})$ and consider the convolution map $R_\phi : f \mapsto f * \phi := \int_G f(y)\phi(y^{-1}x)dy$. By Proposition 2 of §3.1, R_ϕ maps $L^2(G)$ in $C(G)$ and it is compact. From Proposition 4 of §3.1 its image is in the uniform closure of the space spanned by its eigenfunctions corresponding to nonzero eigenvalues.

By construction, the convolution R_ϕ is G-equivariant for the left action, hence it follows that the eigenspaces of this operator are G-stable. Since R_ϕ is a compact operator, its eigenspaces relative to nonzero eigenvalues are finite dimensional and hence in $\mathcal{T}_G$, by the definition of representative functions. Thus the image of R_ϕ is contained in the uniform closure of $\mathcal{T}_G$.

[59] One has to be careful about the normalization. When G is a finite group the usual multiplication in the group algebra is convolution, but for the normalized measure in which G has measure $|G|$ and not 1, as we usually assume for compact groups.

The next step is to show that, given a continuous function f, as ϕ varies, one can approximate f with elements in the image of R_ϕ as close as possible.

Given $\epsilon > 0$, take an open neighborhood U of 1 such that $|f(x) - f(y)| < \epsilon$ if $xy^{-1} \in U$. Take a continuous function $\phi(x)$ with support in U, positive, with integral 1 and $\phi(x) = \phi(x^{-1})$. We claim that $|f - f * \phi| < \epsilon$:

$$\begin{aligned}|f(x) - (f * \phi)(x)| &= \left|f(x)\int_G \phi(y^{-1}x)dy - \int_G f(y)\phi(y^{-1}x)dy\right| \\ &= \left|\int_{y^{-1}x\in U} (f(x) - f(y))\phi(y^{-1}x)dy\right| \\ &\le \int_{y^{-1}x\in U} |f(x) - f(y)|\phi(y^{-1}x)dy \le \epsilon. \qquad \square\end{aligned}$$

Remark. If G is separable as a topological group, for instance if G is a Lie group, the Hilbert space $L^2(G)$ is separable. It follows again that we can have only countably many spaces $V_i^* \otimes V_i$ with V_i irreducible.

We can now apply the theory developed for L^2 class functions. Recall that a class function is one that is invariant under the action of G embedded diagonally in $G \times G$, i.e., $f(x) = f(g^{-1}xg)$ for all $g \in G$.

Express $f = \sum_i f_i$ with $f_i \in V_i^* \otimes V_i$. By the invariance property and the uniqueness of the expression it follows that each f_i is invariant, i.e., a class function.

We know that in $V_i^* \otimes V_i$ the only invariant functions under the diagonal action are the multiples of the corresponding character. Hence we see that

Corollary. *The irreducible characters are an orthonormal basis of the Hilbert space of L^2 class functions.*

Example. When G is commutative, for instance if $G = S_1^k$ is a torus, all irreducible representations are 1-dimensional. Hence we have that the irreducible characters are an orthonormal basis of the space of L^2 functions.

In coordinates $G = \{(\alpha_1, \dots, \alpha_n)\} \mid |\alpha_i| = 1$, the irreducible characters are the monomials $\prod_{i=1}^n \alpha_i^{h_i}$ and we have the usual theory of Fourier series (in this case one often uses the angular coordinates $\alpha_k = e^{2\pi i\theta_k}$).

3.3 Fourier Analysis

In order to compute integrals of L^2-functions we need to know the Hilbert space structure, induced by the L^2 norm, on each space $V_i^* \otimes V_i$ into which $L^2(G)$ decomposes.

We can do this via the following simple remark. The space $V_i^* \otimes V_i = \mathrm{End}(V_i)$ is irreducible under $G \times G$ and it has two Hilbert space structures for which $G \times G$ is unitary. One is the restriction of the L^2 structure. The other is the Hermitian product on $\mathrm{End}(V_i)$, deduced by the Hilbert space structure on V_i and given by the form $\mathrm{tr}(XY^*)$. Arguing as in Proposition 2 of 3.1, every invariant Hermitian product on an irreducible representation U induces an isomorphism with the conjugate dual.

By Schur's lemma it follows that any two invariant Hilbert space structures are then proportional. Therefore the Hilbert space structure on $\mathrm{End}(V_i)$ induced by the L^2 norm equals $c\ \mathrm{tr}(XY^*)$, c a scalar.

Denote by $\rho_i : G \to GL(V_i)$ the representation. By definition (Chapter 6, §2.6) an element $X \in \mathrm{End}(V_i)$ gives the matrix coefficient $\mathrm{tr}(X\rho_i(g))$. In order to compute c take $X = 1$. We have $\mathrm{tr}(1_{V_i}) = \dim V_i$. The matrix coefficient corresponding to 1_{V_i} is the irreducible character $\chi_{V_i}(g) = \mathrm{tr}(\rho_i(g))$ and its L^2 norm is 1. Thus we deduce that $c = \dim V_i^{-1}$. In other words:

Theorem 1. *If $X, Y \in \mathrm{End}(V_i)$ and $c_X(g) = \mathrm{tr}(\rho_i(g)X), c_Y = \mathrm{tr}(\rho_i(g)Y)$ are the corresponding matrix coefficients, we have*

$$\int_G c_X(g)\overline{c_Y(g)}dg = \dim V_i^{-1}\ \mathrm{tr}(XY^*). \tag{3.3.1}$$

Let us finally understand convolution. We want to extend the basic isomorphism theorem for the group algebra of a finite group proved in Chapter 6, §2.6. Given a finite-dimensional representation $\rho : G \to GL(U)$ of G and a function $f \in L^2(G)$ we can define an operator T_f on U by the formula $T_f(u) := \int_G f(g)\rho(g)(u)\,dg$.

Lemma. *The map $f \mapsto T_f$ is a homomorphism from $L^2(G)$ with convolution to the algebra of endomorphisms of U.*

Proof.

$$\begin{aligned} T_{a*b}(u) &:= \int_G (a*b)(g)\rho(g)(u)dg = \int_G\int_G a(h)b(h^{-1}g)\rho(g)(u)\,dh\,dg \\ &= \int_G\int_G a(h)b(g)\rho(hg)(u)\,dh\,dg \\ &= \int_G a(h)\rho(h)\left(\int_G b(g)\ \rho(g)(u)\,dg\right)dh = T_a(T_b(u)). \qquad \square \end{aligned}$$

We have already remarked that convolution $f * g$ is G-equivariant for the left action on f; similarly it is G-equivariant for the right action on g. In particular it maps the representative functions into themselves. Moreover, since the spaces $V_i^* \otimes V_i = \mathrm{End}(V_i)$ are distinct irreducibles under the $G \times G$ action and isotypic components under the left or right action, it follows that under convolution, $\mathrm{End}(V_i) * \mathrm{End}(V_j) = 0$ if $i \neq j$ and $\mathrm{End}(V_i) * \mathrm{End}(V_i) \subset \mathrm{End}(V_i)$.

Theorem 2. *For each irreducible representation $\rho : G \to GL(V)$ embed* $\mathrm{End}(V)$ *in $L^2(G)$ by the map $j_V : X \mapsto \dim V\ \mathrm{tr}(X\rho(g^{-1}))$. Then on* $\mathrm{End}(V)$ *convolution coincides with multiplication of endomorphisms.*

Proof. Same proof as in Chapter 6. By the previous lemma we have a homomorphism $\pi_V : L^2(G) \to \mathrm{End}(V)$. By the previous remarks $\mathrm{End}(V) \subset L^2(G)$ is a subalgebra under convolution. Finally we have to show that $\pi_V j_V$ is the identity of $\mathrm{End}(V)$.

In fact given $X \in \mathrm{End}(V)$, we have $j_V(X)(g) = \mathrm{tr}(\rho(g^{-1})X)\dim V$. In order to prove that $\pi_V j_X(X) = (\dim V)\int_G \mathrm{tr}(\rho(g^{-1})X)\rho(g)dg = X$ it is enough to prove that for any $Y \in \mathrm{End}(V)$, we have $(\dim V)\,\mathrm{tr}(\int_G \mathrm{tr}(\rho(g^{-1})X)\rho(g)dg\, Y) = \mathrm{tr}(XY)$. We have by 3.3.1.

$$\begin{aligned}\dim V tr\left(\int_G \mathrm{tr}(\rho(g^{-1})X)\rho(g)dg\, Y\right) &= \dim V \int_G \mathrm{tr}(\rho(g^{-1})X)\,\mathrm{tr}(\rho(g)Y)dg \\ &= \dim V \int_G \mathrm{tr}(\rho(g)Y)\overline{\mathrm{tr}(\rho(g)X^*)}dg \\ &= \mathrm{tr}(YX^{**}) = \mathrm{tr}(XY).\end{aligned}$$ □

Warning. For finite groups, when we define convolution, that is, multiplication in the group algebra, we use the non-normalized Haar measure. Then for j_V we obtain the formula $j_V : X \mapsto \frac{\dim V}{|G|}\,\mathrm{tr}(X\rho(g^{-1}))$.

3.4 Compact Lie Groups

We draw some consequences of the Peter–Weyl Theorem. Let G be a compact group. Consider any continuous representation of G in a Hilbert space H.

A vector v such that the elements gv, $g \in G$, span a finite-dimensional vector space is called a *finite vector*.

Proposition 1. *The set of finite vectors is dense.*

Proof. By module theory, if $u \in H$, the set $\mathcal{T}_G u$, spanned by applying the representative functions, is made of finite vectors, but by continuity $u = 1u$ is a limit of these vectors. □

Proposition 2. *The intersection $K = \cap_i K_i$, of all the kernels K_i of all the finite-dimensional irreducible representations V_i of a compact group G is $\{1\}$.*

Proof. From the Peter–Weyl theorem we know that $\mathcal{T}_G = \bigoplus_i V_i^* \otimes V_i$ is dense in the continuous functions; since these functions do not separate the points of the intersection of kernels we must have $K = \{1\}$. □

Theorem. *A compact Lie group G has a faithful finite-dimensional representation.*

Proof. Each of the kernels K_i is a Lie subgroup with some Lie algebra L_i and we must have, from the previous proposition, that the intersection $\cap_i L_i = 0$, of all these Lie algebras is 0. This implies that there are finitely many representations $V_i, i = 1, \dots, m$, with the property that the set of elements in all the kernels K_i of the V_i is a subgroup with Lie algebra equal to 0. Thus $\cap_{i=1}^m K_i$ is discrete and hence finite since we are in a compact group. By the previous proposition we can find finitely many representations so that also the non-identity elements of this finite group are not in the kernel of all these representations. Taking the direct sum we find the required faithful representation. □

Let us make a final consideration about the Haar integral. Since the Haar integral is both left- and right-invariant it is a $G \times G$-equivariant map from $L^2(G)$ to the trivial representation. In particular if we restrict it to the representative functions $\bigoplus_i V_i^* \otimes V_i$, it must vanish on each irreducible component $V_i^* \otimes V_i$, which is different from the trivial representation which is afforded by the constant functions. Thus,

Proposition 3. *The Haar integral restricted to $\bigoplus_i V_i^* \otimes V_i$ is the projection to the constant functions, which are the isotypic component of the trivial representation, with kernel all the other nontrivial isotypic components.*

4 Representations of Linearly Reductive Groups

4.1 Characters for Linearly Reductive Groups

We have already stressed several times that we will show a very tight relationship between compact Lie groups and linearly reductive groups. We thus start to discuss characters for linearly reductive groups.

Consider the action by conjugation of G on itself. It is the restriction to G, embedded diagonally in $G \times G$, of the left and right actions.

Let $Z[G]$ denote the space of regular functions f which are invariant under conjugation.

From the decomposition of Chapter 7, §3.1.1, $F[G] = \bigoplus_i U_i^* \otimes U_i$, it follows that the space $Z[G]$ decomposes as a direct sum of the spaces $Z[U_i]$ of conjugation-invariant functions in $U_i^* \otimes U_i$. We claim that:

Lemma. *$Z[U_i]$ is 1-dimensional, generated by the character of the representation U_i.*

Proof. Since U_i is irreducible and $U_i^* \otimes U_i = \mathrm{End}(U_i)^*$ we have by Schur's lemma that $Z[U_i]$ is 1-dimensional, generated by the element corresponding to the trace on $\mathrm{End}(U_i)$.

Now we follow the identifications. An element u of $\mathrm{End}(U_i)^*$ gives the matrix coefficient $u(\rho_i(g))$ where $\rho_i : G \to GL(U_i) \subset \mathrm{End}(U_i)$ denotes the representation map.

We obtain the function $\chi_i(g) = \mathrm{tr}(\rho_i(g))$ as the desired invariant element. □

Corollary. *For a linearly reductive group, the G-irreducible characters are a basis of the conjugation-invariant functions.*

We will see in Chapter 10 that any two maximal tori are conjugate and the union of all maximal tori in a reductive group G is dense in G. One of the implications of this theorem is the fact that the character of a representation M of G is determined by its restriction to a given maximal torus T. On M the group T acts as a direct sum of irreducible 1-dimensional characters in $\hat{T}$, and thus the character of M can be expressed as a sum of these characters with nonnegative coefficients, expressing their multiplicities.

After restriction to a maximal torus T, the fact that a character is a class function implies a further symmetry. Let N_T denote the normalizer of T. N_T acts on T by conjugation, and a class function restricted to T is invariant under this action. There are many important theorems about this action, the first of which is:

Theorem 1. *T equals its centralizer and N_T/T is a finite group, called the Weyl group and denoted by W.*

Under restriction to a maximal torus T, the ring of characters of G is isomorphic to the subring of W-invariant characters of T.

Let us illustrate the first part of this theorem for classical groups, leaving the general proof to Chapter 10. We always take advantage of the same idea.

Let T be a torus contained in the linear group of a vector space V.

Decompose $V := \bigoplus_\chi V_\chi$ in weight spaces under T and let $g \in GL(V)$ be a linear transformation normalizing T. Clearly g induces by conjugation an automorphism of T, which we still denote by g, which permutes the characters of T by the formula $\chi^g(t) := \chi(g^{-1}tg)$.

We thus have, for $v \in V_\chi$, $t \in T$, $tgv = gg^{-1}tgv = \chi^g(t)gv$.

We deduce that $gV_\chi = V_{\chi^g}$. In particular g permutes the weight spaces. We thus have a homomorphism from the normalizer of the torus to the group of permutations of the weight spaces. Let us now analyze this for T a maximal torus in the general linear, orthogonal and symplectic groups. We refer to Chapter 7, §4.1 for the description of the maximal tori in these three cases. First, analyze the kernel N_T^0 of this homomorphism. We prove that $N_T^0 = T$ in the four cases.

(1) **General linear group.** Let D be the group of all diagonal matrices (in the standard basis e_i). It is exactly the full subgroup of linear transformations fixing the 1-dimensional weight spaces generated by the given basis vectors.

An element in N_D^0 by definition fixes all these subspaces and thus in this case $N_D^0 = D$.

(2) **Even orthogonal group.** Again the space decomposes into 1-dimensional eigenspaces spanned by the vectors e_i, f_i giving a hyperbolic basis. One immediately verifies that a diagonal matrix g given by $ge_i = \alpha_i e_i$, $gf_i = \beta_i f_i$ is orthogonal if and only if $\alpha_i \beta_i = 1$. The matrices form a maximal torus T. Again $N_T^0 = T$.

(3) **Odd orthogonal group.** The case is similar to the previous case except that now we have an extra non-isotropic basis vector u and g is orthogonal if furthermore $gu = \pm u$. It is special orthogonal only if $gu = u$. Again $N_T^0 = T$.

(4) **Symplectic group.** Identical to (2).

Now for the full normalizer. (1) In the case of the general linear group, N_D contains the symmetric group S_n acting as permutations on the given basis.

If $a \in N_D$ we must have that $a(e_i) \in \mathbb{C}e_{\sigma(i)}$ for some $\sigma \in S_n$. Thus $\sigma^{-1}a$ is a diagonal matrix and it follows that $N_D = D \ltimes S_n$, the semidirect product.

In the case of the special linear group, we leave it to the reader to verify that we still have an exact sequence $0 \to D \to N_D \to S_n \to 0$, but this does not split, since only the even permutations are in the special linear group.

(2) In the even orthogonal case $\dim V = 2n$, the characters come in opposite pairs and their weight spaces are spanned by the vectors $e_1, e_2, \dots, e_n;\ f_1, f_2, \dots, f_n$ of a hyperbolic basis (Chapter 7, §4.1). Clearly the normalizer permutes this set of n pairs of subspaces $\{\mathbb{C}e_i, \mathbb{C}f_i\}$.

In the same way as before, we see now that the symmetric group S_n, permuting simultaneously with the same permutation the elements $e_1, e_2, \dots, e_n$ and $f_1, f_2, \dots, f_n$ consists of special orthogonal matrices.

The kernel of the map $N_T \to S_n$ is formed by the matrices diagonal of 2×2 blocks. Each 2×2 block is the orthogonal group of the 2-dimensional space spanned by e_i, f_i and it is the semidirect product of the torus part $\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ with the permutation matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

In the special orthogonal group only an even number of permutation matrices $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ can appear. It follows that the Weyl group is the semidirect product of the symmetric group S_n with the subgroup of index 2 of $\mathbb{Z}/(2)^n$ formed by the n-tuples $a_1, \dots, a_n$ with $\sum_{i=1}^n a_i = 0, \pmod 2$.

(3) The odd special orthogonal group is slightly different. We use the notations of Chapter 5. Now one has also the possibility to act on the basis $e_1, f_1, e_2, f_2, \dots, e_n, f_n, u$ by -1 on u and this corrects the fact that the determinant of an element defined on $e_1, f_1, e_2, f_2, \dots, e_n, f_n$ may be -1.

We deduce then that the Weyl group is the semidirect product of the symmetric group S_n with $\mathbb{Z}/(2)^n$.

(4) The symplectic group. The discussion starts as in the even orthogonal group, except now the 2-dimensional symplectic group is $SL(2)$. Its torus of 2×2 diagonal matrices has index 2 in its normalizer and as representatives of the Weyl group we can choose the identity and the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

This matrix has determinant 1 and again we deduce that the Weyl group is the semidirect product of the symmetric group S_n with $\mathbb{Z}/(2)^n$.

Now we have to discuss the action of the Weyl group on the characters of a maximal torus. In the case of the general linear group a diagonal matrix X with entries $x_1, \dots, x_n$ is conjugated by a permutation matrix σ which maps $\sigma e_i = e_{\sigma(i)}$ by $\sigma X \sigma^{-1} e_i = x_{\sigma(i)} e_i$; thus the action of S_n on the *characters* x_i is the usual permutation of variables.

For the orthogonal groups and the symplectic group one has the torus of diagonal matrices of the form $x_1, x_1^{-1}, x_2, x_2^{-1}, \dots, x_n, x_n^{-1}$. Besides the permutations of the variables we now have also the inversions $x_i \to x_i^{-1}$, except that for the even orthogonal group one has to restrict to products of only an even number of inversions.

This analysis suggests an interpretation of the characters of the classical groups as particular symmetric functions. In the case of the linear group the coordinate ring of the maximal torus can be viewed as the polynomial ring $\mathbb{C}[x_1, \dots, x_n][d^{-1}]$ with $d := \prod_{i=1}^n x_i$ inverted.

d is the n^{th} elementary symmetric function and thus the invariant elements are the polynomial in the elementary symmetric functions $\sigma_i(x)$, $i = 1, \dots, n-1$ and $\sigma_n(x)^{\pm 1}$.

In the case of the inversions we make a remark. Consider the ring $A[t, t^{-1}]$ of Laurent polynomials over a commutative ring A. An element $\sum_i a_i t^i$ is invariant under $t \to t^{-1}$ if and only if $a_i = a_{-i}$. We claim then that it is a polynomial in $u := t + t^{-1}$. In fact $t^i + t^{-i} = (t + t^{-1})^i + r(t)$ where $r(t)$ has lower degree and one can work by induction. We deduce

Theorem 2. *The ring of invariants of* $\mathbb{C}[x_1, x_1^{-1}, \dots, x_n, x_n^{-1}]$ *under* $S_n \ltimes \mathbb{Z}/(2)^n$ *is the polynomial ring in the elementary symmetric functions* $\sigma_i(u)$ *in the variables* $u_i := x_i + x_i^{-1}$.

Proof. We can compute the invariants in two steps. First we compute the invariants under $\mathbb{Z}/(2)^n$ which, by the previous argument, are the polynomials in the u_i. Then we compute the invariants under the action of S_n which permutes the u_i. The claim follows. $\square$

For the even orthogonal group we need a different computation since now we only want the invariants under a subgroup. Let $H \subset \mathbb{Z}/(2)^n$ be the subgroup defined by $\sum_i a_i = 0$.

Start from the monomial $M := x_1 x_2 \dots x_n$, the orbit of this monomial, under the group of inversions $\mathbb{Z}/(2)^n$ consists of all the monomials $x_1^{\epsilon_1} x_2^{\epsilon_2} \dots x_n^{\epsilon_n}$ where the elements $\epsilon_i = \pm 1$. We next define

$$E := \sum_{\prod_{i=1}^n \epsilon_i = 1} x_1^{\epsilon_1} x_2^{\epsilon_2} \dots x_n^{\epsilon_n}, \quad \overline{E} := \sum_{\prod_{i=1}^n \epsilon_i = -1} x_1^{\epsilon_1} x_2^{\epsilon_2} \dots x_n^{\epsilon_n};$$

E is clearly invariant under H and $E + \overline{E}$, $E\overline{E}$ are invariant under $\mathbb{Z}/(2)^n$.

We claim that any H-invariant is of the form $a + bE$ where a, b are $\mathbb{Z}/(2)^n$-invariants.

Consider the set of all Laurent monomials which is permuted by $\mathbb{Z}/(2)^n$. A basis of invariants under $\mathbb{Z}/(2)^n$ is clearly given by the sums of the vectors in each orbit, and similarly for the H-invariants. Now let K be the stabilizer of an element of the orbit, which thus has $\frac{2^n}{|K|}$ elements. The stabilizer in H is $K \cap H$, hence a $\mathbb{Z}/(2)^n$ orbit is either an H-orbit or it splits into two orbits, according to whether $K \not\subset H$ or $K \subset H$.

We get H-invariants which are not $\mathbb{Z}/(2)^n$-invariants from the last type of orbits.

A monomial $M = \prod x_i^{h_i}$ is stabilized by all the inversions in the variables x_i that have exponent 0. Thus the only case in which the stabilizer is contained in H is when all the variables x_i appear. In this case, in the $\mathbb{Z}/(2)^n$ orbit of M there is a unique element, which by abuse of notations we still call M, for which $h_i > 0$ for all i. Let $S^i_{h_1,\dots,h_n}$, $i = 1, 2$, be the sum on the two orbits of M under H.

Since $S^1_{h_1,\dots,h_n} + S^2_{h_1,\dots,h_n}$ is invariant under $\mathbb{Z}/(2)^n$ it is only necessary to show that $S^1_{h_1,\dots,h_n}$ has the required form. The multiplication $S^1_{h_1-1,\dots,h_n-1} S^1_{1,1,\dots,1}$ gives rise to $S^1_{h_1,\dots,h_n}$ plus terms which are lower in the lexicographic ordering of the h_i's, and

$S^1_{1,1,\dots,1} = E$. By induction we assume that the lower terms are of the required form. Also by induction $S^1_{h_1-1,\dots,h_n-1} = a + bE$, and so we have derived the required form:

$$S^1_{h_1,\dots,h_n} = (a + bE)E = (a + b(E + \overline{E}))E - b(E\overline{E}).$$

We can now discuss the invariants under the Weyl group. Again, the ring of invariants under H is stabilized by S_n which acts by permuting the elements u_i, and fixing the element E. We deduce that the ring of W-invariants is formed by elements of the form $a+bE$ where a, b are polynomials in the elementary symmetric functions in the elements u_i.

It remains to understand the quadratic equation satisfied by E over the ring of symmetric functions in the u_i. E satisfies the relation $E^2 - (E + \overline{E})E + E\overline{E} = 0$ and so we must compute the symmetric functions $E + \overline{E}$, $E\overline{E}$.

We easily see that $E+\overline{E} = \prod_{i=1}^n (x_i+x_i^{-1})$ which is the n^{th} elementary symmetric function in the u_i's. As for $E\overline{E}$, it can be easily described as a sum of monomials in which the exponents are either 2 or -2, with multiplicities expressed by binomial coefficients. We leave the details to the reader.

5 Induction and Restriction

5.1 Clifford's Theorem

We now collect some general facts about representations of groups. First, let H be a group, $\phi : H \to H$ an automorphism, and $\rho : H \to GL(V)$ a linear representation.

Composing with ϕ we get a new representation V^ϕ given by $H \xrightarrow{\phi} H \xrightarrow{\rho} GL(V)$; it is immediately verified that if ϕ is an inner automorphism, V^ϕ is equivalent to ϕ.

Let $H \subset G$ be a normal subgroup. Every element $g \in G$ induces by inner conjugation in G an automorphism $\phi_g : h \mapsto ghg^{-1}$ of H.

Let M be a representation of G and $N \subset M$ an H-submodule. Since $hg^{-1}n = g^{-1}(ghg^{-1})n$, we clearly have that $g^{-1}N \subset M$ is again an H-submodule and canonically isomorphic to N^{ϕ_g}. It depends only on the coset $g^{-1}H$.

In particular, assume that M is irreducible as a G-module and N is irreducible as an H-module. Then all the submodules gN are irreducible H-modules and $\sum_{g\in G/H} gN$ is a G-submodule, hence $\sum_{g\in G/H} gN = M$.

We want in particular to apply this when H has index 2 in $G = H \cup uH$. We shall use the canonical sign representation ϵ of $\mathbb{Z}/(2) = G/H$, $\epsilon(u) = -1$, $\epsilon(H) = 1$.

Clifford's Theorem. *(1) An irreducible representation N of H extends to a representation of G if and only if N is isomorphic to N^{ϕ_u}. In this case it extends in two ways up to the sign representation.*

(2) An irreducible representation M of G restricted to H remains irreducible if M is not isomorphic to $M \otimes \epsilon$. It splits into two irreducible representations $N \oplus N^{\phi_u}$ if M is isomorphic to $M \otimes \epsilon$.

Proof. Let $h_0 = u^2 \in H$. If N is also a G-representation, the map $u : N \to N$ is an isomorphism with N^{ϕ_u}. Conversely, let $t : N \to N = N^{\phi_u}$ be an isomorphism so that $tht^{-1} = \phi_u(h)$ as operators on N. Then $t^2ht^{-2} = h_0hh_0^{-1}$, hence $h_0^{-1}t^2$ commutes with H.

Since N is irreducible we must have $h_0^{-1}t^2 = \lambda$ is a scalar. We can substitute t with $t\sqrt{\lambda}^{-1}$ and can thus assume that $t^2 = h_0$ (on N).

It follows that mapping $u \mapsto t$ gives the required extension of the representation. It also is clear that the choice of $-t$ is the other possible choice changing the sign of the representation.

(2) From our previous discussion if $N \subset M$ is an irreducible H-submodule, then $M = N + uM$, $uM = u^{-1}N \cong N^{\phi_u}$, and we clearly have two cases: $M = N$ or $M = N \oplus uN$.

In the first case, tensoring by the sign representation changes the representation. In fact if we had an isomorphism t between N and $N \otimes \epsilon$ this would also be an isomorphism of N to N as H-modules. Since N is irreducible over H, t must be a scalar, but then the identity is an isomorphism between N and $N \otimes \epsilon$, which is clearly absurd.

In the second case, $M = N \oplus u^{-1}N$; on $n_1 + u^{-1}n_2$ the action of H is by $hn_1 + u^{-1}\phi_u(h)n_2$, while $u(n_1 + u^{-1}n_2) = n_2 + u^{-1}h_0n_1$.

On $M \otimes \epsilon$ the action of u changes to $u(n_1 + u^{-1}n_2) = -n_2 - u^{-1}h_0n_1$. Then it is immediately seen that the map $n_1 + u^{-1}n_2 \mapsto n_1 - u^{-1}n_2$ is an isomorphism between M and $M \otimes \epsilon$. $\square$

One should compare this property of the possible splitting of irreducible representations with the similar feature for conjugacy classes.

Exercise (same notation as before). A conjugacy class C of G contained in H is either a unique conjugacy class in H or it splits into two conjugacy classes permuted by exterior conjugation by u. The second case occurs if and only if the stabilizer in G of an element in the conjugacy class is contained in H. Study $A_n \subset S_n$ (the alternating group).

5.2 Induced Characters

Now let G be a group, H a subgroup and N a representation of H (over some field k).

In Chapter 1, §3.2 we have given the notion of induced representation. Let us rephrase it in the language of modules. Consider $k[G]$ as a left $k[H]$-module by the right action. The space $\hom_{k[H]}(k[G], N)$ is a representation under G by the action of G deduced from the left action on $k[G]$.

$$\hom_{k[H]}(k[G], N) := \{f : G \to N \mid f(gh^{-1}) = hf(g)\},\ (gf)(k) := f(g^{-1}k).$$

We recover the notion given $\mathrm{Ind}_H^G(N) = \hom_{k[H]}(k[G], N)$. We already remarked that this may be called coinduced.

To be precise, this construction is the induced representation only when H has finite index in G. Otherwise one has a different construction which we leave to the reader to compare with the one presented:

Consider $k[G] \otimes_{k[H]} N$, where now $k[G]$ is thought of as a right module under $k[H]$. It is a representation under G by the left action of G on $k[G]$.

Exercise.

(1) If $G \supset H \supset K$ are groups and N is a K-module we have

$$\mathrm{Ind}_H^G(\mathrm{Ind}_K^H N)) = \mathrm{Ind}_K^G N.$$

(2) The representation $\mathrm{Ind}_H^G N$ is in a natural way described by $\bigoplus_{g \in G/H} gN$ where, by $g \in G/H$, we mean that g runs over a choice of representatives of cosets. The action of G on such a sum is easily described.

The definition we have given of induced representation extends in a simple way to algebraic groups and rational representations. In this case $k[G]$ denotes the space of regular functions on G. If H is a closed subgroup of G, one can define $\hom_{k[H]}(k[G], N)$ as the set of regular maps $G \to N$ which are H-equivariant (for the right action on G).

The regular maps from an affine algebraic variety V to a vector space U can be identified to $A(V) \otimes U$ where $A(V)$ is the ring of regular functions on V. Hence if V has an action under an algebraic group H and U is a rational representation of H, the space of H-equivariant maps $V \to U$ is identified with the space of invariants $(A(V) \otimes U)^H$.

Assume now that G is linearly reductive and let us invoke the decomposition 3.1.1 of Chapter 7, $k[G] = \bigoplus_i U_i^* \otimes U_i$. Since by right action H acts only on the factor U_i,

$$\hom_{k[H]}(k[G], N) = \bigoplus_i U_i^* \otimes \hom_H(U_i, N).$$

Finally, if N is irreducible (under H), and H is also linearly reductive, it follows from Schur's Lemma that the dimension of $\hom_H(U_i, N)$ is the multiplicity with which N appears in U_i. We thus deduce

Theorem (Frobenius reciprocity for coinduced representations). *The multiplicity with which an irreducible representation V of G appears in* $\hom_{k[H]}(k[G], N)$ *equals the multiplicity with which N appears in V^* as a representation of H.*

5.3 Homogeneous Spaces

There are several interesting results of Fourier analysis on homogeneous spaces which are explained easily by the previous discussion. Suppose we have a finite-dimensional complex unitary or real orthogonal representation V of a compact group K. Let $v \in V$ be a vector and consider its orbit Kv, which is isomorphic to the homogeneous space K/K_v where K_v is the stabilizer of v. Under the simple condition that $\overline{v} \in Kv$ (no condition in the real orthogonal case) the polynomial functions on V, restricted to Kv, form an algebra of functions satisfying the properties of the Stone–Weierstrass theorem. The Euclidean space structure on V induces on the manifold Kv a K-invariant metric, hence also a measure and a unitary representation of K on the space of L^2 functions on Kv. Thus the same analysis as in 3.2 shows that we can decompose the restriction of the polynomial functions to Kv into an orthogonal direct sum of irreducible representations. The whole space $L^2(K/K_v)$ then decomposes in Fourier series obtained from these irreducible blocks. One method to understand which representations appear and with which multiplicity is to apply Frobenius reciprocity. Another is to apply methods of algebraic geometry to the associated action of the associated linearly reductive group, see §9. A classical example comes from the theory of *spherical harmonics* obtained restricting the polynomial functions to the unit sphere.

6 The Unitary Trick

6.1 Polar Decomposition

There are several ways in which linearly reductive groups are connected to compact Lie groups. The use of this (rather strict) connection goes under the name of the unitary trick. This is done in many different ways. Here we want to discuss it with particular reference to the examples of classical groups which we are studying.

We start from the remark that the unitary group $U(n, \mathbb{C}) := \{A | AA^* = 1\}$ is a bounded and closed set in $M_n(\mathbb{C})$, hence it is compact.

Proposition 1. *$U(n, \mathbb{C})$ is a maximal compact subgroup of $GL(n, \mathbb{C})$. Any other maximal compact subgroup of $GL(n, \mathbb{C})$ is conjugate to $U(n, \mathbb{C})$.*

Proof. Let K be a compact linear group. Since K is unitarizable there exists a matrix g such that $K \subset gU(n, \mathbb{C})g^{-1}$. If K is maximal this inclusion is an equality. □

The way in which $U(n, \mathbb{C})$ sits in $GL(n, \mathbb{C})$ is very special and common to maximal compact subgroups of linearly reductive groups. The analysis passes through the polar decomposition for matrices and the Cartan decomposition for groups.

Theorem. *(1) The map $B \to e^B$ establishes a diffeomorphism between the space of Hermitian matrices and the space of positive Hermitian matrices.*

(2) Every invertible matrix X is uniquely expressible in the form

$$X = e^B A \qquad \textbf{\textit{(polar decomposition)}} \tag{6.1.1}$$

where A is unitary and B is Hermitian.

Proof. (1) We leave it as an exercise, using the eigenvalues and eigenspaces.

(2) Consider $XX^* := X\overline{X}^t$ which is clearly a positive Hermitian matrix.

If $X = e^B A$ is decomposed as in 6.1.1, then $XX^* = e^B AA^* e^B = e^{2B}$. So B is uniquely determined. Conversely, by decomposing the space into eigenspaces, it is clear that a positive Hermitian matrix is uniquely of the form e^{2B} with B Hermitian. Hence there is a unique B with $XX^* = e^{2B}$. Setting $A := e^{-B}X$ we see that A is unitary and $X = e^B A$. □

The previous theorem has two corollaries, both of which are sometimes used as unitary tricks, the first of algebro-geometric nature and the second topological.

Corollary. *(1) $U(n, \mathbb{C})$ is Zariski dense in $GL(n, \mathbb{C})$.*

(2) $GL(n, \mathbb{C})$ is diffeomorphic to $U(n, \mathbb{C}) \times \mathbb{R}^{n^2}$ via $\phi(A, B) = e^B A$. In particular $U(n, \mathbb{C})$ is a deformation retract of $GL(n, \mathbb{C})$.

Proof. The first part follows from the fact that one has the exponential map $X \to e^X$ from complex $n \times n$ matrices to $GL(n, \mathbb{C})$. In this holomorphic map the two subspaces $i\mathcal{H}$ and $\mathcal{H}$ of anti-Hermitian and Hermitian matrices map to the two factors of the polar decomposition, i.e., unitary and positive Hermitian matrices.

Since $M_n(\mathbb{C}) = \mathcal{H} + i\mathcal{H}$, any two holomorphic functions on $M_n(\mathbb{C})$ coinciding on $i\mathcal{H}$ necessarily coincide. So by the exponential and the connectedness of $GL(n, \mathbb{C})$, the same holds in $GL(n, \mathbb{C})$: two holomorphic functions on $GL(n, \mathbb{C})$ coinciding on $U(n, \mathbb{C})$ coincide. □

There is a partial converse to this analysis.

Proposition 2. *Let $G \subset GL(n, \mathbb{C})$ be an algebraic group. Suppose that $K := G \cap U(n, \mathbb{C})$ is Zariski dense in G. Then G is self-adjoint.*

Proof. Let us consider the antilinear map $g \mapsto g^*$. Although it is not algebraic, it maps algebraic varieties to algebraic varieties (conjugating the equations). Thus G^* is an algebraic variety in which K^* is Zariski dense. Since $K^* = K$ we have $G^* = G$. □

6.2 Cartan Decomposition

The polar decomposition induces, on a self-adjoint group $G \subset GL(n, \mathbb{C})$ of matrices, a *Cartan decomposition*, under a mild topological condition.

Let $u(n, \mathbb{C})$ be the anti-Hermitian matrices, the Lie algebra of $U(n, \mathbb{C})$. Then $iu(n, \mathbb{C})$ are the Hermitian matrices. Let $\mathfrak{g} \subset gl(n, \mathbb{C})$ be the Lie algebra of G.

Theorem (Cartan decomposition). *Let $G \subset GL(n, \mathbb{C})$ be a self-adjoint Lie group with finitely many connected components, $\mathfrak{g}$ its Lie algebra.*

(i) For every element $A \in G$ in polar form $A = e^B U$, we have that $U \in G$, $B \in \mathfrak{g}$. Let $K := G \cap U(n, \mathbb{C})$ and let $\mathfrak{k}$ be the Lie algebra of K.

(ii) We have $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, $\mathfrak{p} = \mathfrak{g} \cap iu(n, \mathbb{C})$. The map $\phi : K \times \mathfrak{p} \to G$ given by $\phi : (u, p) \mapsto e^p u$ is a diffeomorphism.

(iii) If $\mathfrak{g}$ is a complex Lie algebra we have $\mathfrak{p} = i\mathfrak{k}$.

Proof. If G is a self-adjoint group, clearly (taking 1-parameter subgroups) also its Lie algebra is self-adjoint. Since $X \mapsto X^*$ is a linear map of order 2, by self-adjointness $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, with $\mathfrak{k}$ the space of anti-Hermitian and $\mathfrak{p}$ of Hermitian elements of $\mathfrak{g}$. We have that $\mathfrak{k} := \mathfrak{g} \cap u(n, \mathbb{C})$ is the Lie algebra of $K := G \cap U(n, \mathbb{C})$. $K \times \mathfrak{p}$ is a submanifold of $U(n, \mathbb{C}) \times iu(n, \mathbb{C})$. The map $\phi : K \times \mathfrak{p} \to G$, being the restriction to a submanifold of a diffeomorphism, is a diffeomorphism with its image. Thus the key to the proof is to show that its image is G. In other words that if $A = e^B U \in G$ is in polar form, we have that $U \in K$, $B \in \mathfrak{p}$.

Now $e^{2B} = AA^* \in G$ by hypothesis, so it suffices to see that if B is an Hermitian matrix with $e^B \in G$, we have $B \in \mathfrak{g}$. Since $e^{nB} \in G$, $\forall n \in \mathbb{Z}$, the hypothesis that G has finitely many connected components implies that for some n, $e^{nB} \in G_0$, where G_0 denotes the connected component of the identity. We are reduced to the case G connected. In the diffeomorphism $U(n, \mathbb{C}) \times iu(n, \mathbb{C}) \to GL(n, \mathbb{C})$, $(U, B) \mapsto e^B U$, we have that $K \times \mathfrak{p}$ maps diffeomorphically to a closed submanifold of $GL(n, \mathbb{C})$ contained in G. Since clearly this submanifold has the same dimension as G and G is connected we must have $G = K \times e^{\mathfrak{p}}$, the Cartan decomposition for G.

Finally, if $\mathfrak{g}$ is a complex Lie algebra, multiplication by i maps the Hermitian to the anti-Hermitian matrices in $\mathfrak{g}$, and conversely. □

Exercise. See that the condition on finitely many components cannot be dropped.

Corollary. *The homogeneous space G/K is diffeomorphic to* $\mathfrak{p}$.

It is useful to make explicit the action of an element of G, written in its polar decomposition, on the homogeneous space G/K. Denote by $P := e^{\mathfrak{p}}$. We have a map $\rho : G \to P$ given by $\rho(g) := gg^*$. ρ is a G-equivariant map if we act with G on G by left multiplication and on P by gpg^*. ρ is an orbit map, P is the orbit of 1, and the stabilizer of 1 is K. Thus ρ identifies G/K with P and the action of G on P is gpg^*.[60]

Theorem 2. *Let G be as before and let M be a compact subgroup of G. Then M is conjugate to a subgroup of K.*

K is maximal compact and all maximal compact subgroups are conjugate in G.

The second statement follows clearly from the first. By the fixed point principle (Chapter 1, §2.2), this is equivalent to proving that M has a fixed point on G/K. This may be achieved in several ways. The classical proof is via Riemannian geometry, showing that G/K is a Riemannian symmetric space of constant negative curvature.[61] We follow the more direct approach of [OV]. For this we need some preparation.

We need to study an auxiliary function on the space P and its closure $\overline{P}$, the set of all positive semidefinite Hermitian matrices. Let $G = GL(n, \mathbb{C})$. Consider the two-variable function $\mathrm{tr}(xy^{-1})$, $x \in \overline{P}$, $y \in P$. Since $(gxg^*)(gyg^*)^{-1} = gxy^{-1}g^{-1}$,

---

[60] Observe that, restricted to P, the orbit map is $p \mapsto p^2$.

[61] The geometry of these Riemannian manifolds is a rather fascinating part of mathematics; it is the proper setting to understand non-Euclidean geometry in general; we refer to [He].

this function is invariant under the G action on $\overline{P} \times P$. Since x, y are Hermitian, $\operatorname{tr}(xy^{-1}) = \operatorname{tr}(\overline{x}\overline{y}^{-1}) = \operatorname{tr}((\overline{y}^{-1})^t\overline{x}^t) = \operatorname{tr}(xy^{-1})$, so $\operatorname{tr}(xy^{-1})$ is a real function.

Let $\Omega \subset P$ be a compact set. We want to analyze the function

$$(6.2.1) \qquad \rho_\Omega(x) := \max_{a\in\Omega} \operatorname{tr}(xa^{-1}).$$

Remark. If $g \in G$, we have

$$\rho_\Omega(gxg^*) := \max_{a\in\Omega} \operatorname{tr}(gxg^*a^{-1}) = \max_{a\in\Omega} \operatorname{tr}(xg^*a^{-1}g) = \rho_{g^{-1}\Omega}(x).$$

Lemma 1. *The function $\rho_\Omega(x)$ is continuous, and there is a positive constant b such that if $x \neq 0$, $\rho_\Omega(x) > b\|x\|$, where $\|x\|$ is the operator norm.*

Proof. Since Ω is compact, $\rho_\Omega(x)$ is obviously well defined and continuous. Let us estimate $\operatorname{tr}(xa^{-1})$. Fix an orthonormal basis e_i in which x is diagonal, with eigenvalues $x_i \geq 0$. If a^{-1} has matrix a_{ij}, we have $\operatorname{tr}(xa^{-1}) = \sum_i x_i a_{ii}$. Since a is positive Hermitian, $a_{ii} > 0$ for all i and for all orthonormal bases. Since the set of orthonormal bases is compact, there is a positive constant $b > 0$, independent of a and of the basis, such that $a_{ii} > b, \forall i, \forall a \in \Omega$. Hence, if $x \neq 0$, $\operatorname{tr}(xa^{-1}) > \max_i x_i b = \|x\|b$. □

Lemma 2. *Given $C > 0$, the set P_C of matrices $x \in P$ with* $\det(x) = 1$ *and* $\|x\| \leq C$ *is compact.*

Proof. P_C is stable under conjugation by unitary matrices. Since this group is compact, it is enough to see that the set of diagonal matrices in P_C is compact. This is the set of n-tuples of numbers x_i with $\prod_i x_i = 1$, $C \geq x_i > 0$. This is the intersection of the closed set $\prod_i x_i = 1$ with the compact set $C \geq x_i \geq 0, \forall i$. □

From the previous two lemmas it follows that:

Lemma 3. *The function $\rho_\Omega(x)$ admits an absolute minimum on the set of matrices $x \in P$ with* $\det(x) = 1$.

Proof. Let $x_0 \in P$, $\det(x_0) = 1$ and let $c := \rho_\Omega(x_0)$. From Lemma 1, if $x \in P$ is such that $\|x\| > cb^{-1}$, then $\rho_\Omega(x) > c$. Thus the minimum is taken on the set of elements x such that $\|x\| \leq cb^{-1}$ which is compact by Lemma 2. Hence an absolute minimum exists. □

Recall that an element $x \in P$ is of the form $x = e^A$ for a unique Hermitian matrix A. Therefore the function of the real variable u, $x^u := e^{uA}$ is well defined. The key geometric property of our functions is:

Proposition 3. *Given $x, y \in P, x \neq 1$, the two functions of the real variable u, $\phi_{x,y}(u) :=$* $\operatorname{tr}(x^u y^{-1})$ *and $\rho_\Omega(x^u)$, are strictly convex.*

Proof. One way to check convexity is to prove that the second derivative is strictly positive. If $x = e^A \neq 1$, we have that $A \neq 0$ is a Hermitian matrix. The same proof as in Lemma 1 shows that $\ddot{\phi}_{x,y}(u) = \operatorname{tr}(A^2 e^{Au} y^{-1}) > 0$, since $0 \neq A^2 e^{Au} \in \overline{P}$.

Now for $\rho_\Omega(x^u) = \max_{a \in \Omega} \operatorname{tr}(x^u a^{-1}) = \max_{a \in \Omega} \phi_{x,a}(u)$ it is enough to remark that if we have a family of strictly convex functions depending on a parameter in a compact set, the maximum is clearly a strictly convex function. □

Now revert to a self-adjoint group $G \subset GL(n, \mathbb{C}) \subset GL(2n, \mathbb{R})$, its associated P and $\Omega \subset P$ a compact set. Assume furthermore that $G \subset SL(2n, R)$.

Lemma 4. *$\rho_\Omega(x)$ has a unique minimum on P.*

Proof. First, the hypothesis that the matrices have determinant 1 implies from Lemma 3 that an absolute minimum exists. Assume by contradiction that we have two minima in A, B. By the first remark, changing Ω, since G acts transitively on P we may assume $A = 1$. Furthermore, $\lim_{u \to 0} B^u = 1$ (and it is a curve in P). By convexity and the fact that B is a minimum we have that $\rho_\Omega(B^u)$ is a strictly decreasing function for $u \in (0, 1]$, hence $\rho_\Omega(1) = \lim_{u \to 0} \rho_\Omega(B^u) > \rho_\Omega(B)$, a contradiction. □

Proof of Theorem 2. We will apply the fixed point principle of Chapter 1, §2.2, to M acting on $P = G/K$. Observe that $GL(n, \mathbb{C}) \subset GL(2n, \mathbb{R}) \subset GL^+(4n, \mathbb{R})$, the matrices of positive determinant. Thus embed $G \subset GL^+(4n, \mathbb{R})$. The determinant is then a homomorphism to $\mathbb{R}^+$. Any compact subgroup of $GL^+(m, \mathbb{R})$ is contained in the subgroup of matrices with determinant 1, and we can reduce to the case $G \subset SL(2n, \mathbb{R})$.

Let $\Omega := M1$ be the orbit of 1 in $G/K = P$. The function $\rho_{M1}(x)$ on P, by Lemma 4, has a unique minimum point p_0. We claim that $\rho_{M1}(x)$ is M-invariant. In fact, by the first remark, we have for $k \in M$ that $\rho_{M1}(kxk^*) = \rho_{k^{-1}M1}(x) = \rho_{M1}(x)$. It follows that p_0 is necessarily a fixed point of M. □

Exercise. Let G be a group with finitely many components and G_0 the connected component of 1. If G_0 is self-adjoint with respect to some positive Hermitian form, then G is also self-adjoint (under a possibly different Hermitian form).

The application of this theory to algebraic groups will be proved in Chapter 10, §6.3:

Theorem 3. *If $G \subset GL(n, \mathbb{C})$ is a self-adjoint Lie group with finitely many connected components and complex Lie algebra, then G is a linearly reductive algebraic group.*

Conversely, given a linearly reductive group G and a finite-dimensional linear representation of G on a space V, there is a Hilbert space structure on V such that G is self-adjoint.

If V is faithful, the unitary elements of G form a maximal compact subgroup K and we have a canonical polar decomposition $G = Ke^{i\mathfrak{k}}$ where $\mathfrak{k}$ is the Lie algebra of K.

All maximal compact subgroups of G are conjugate in G.

Every compact Lie group appears in this way in a canonical form.

In fact, as Hilbert structure, one takes any one for which a given maximal compact subgroup is formed of unitary elements.

6.3 Classical Groups

For the other linearly reductive groups that we know, we want to make the Cartan decomposition explicit. We are dealing with self-adjoint complex groups, hence with a complex Lie algebra $\mathfrak{g}$. In the notation of §6.2 we have $\mathfrak{p} = i\mathfrak{k}$. We leave some simple details as exercise.

1. First, the diagonal group $T = (\mathbb{C}^*)^n$ decomposes as $U(1,\mathbb{C})^n \times (\mathbb{R}^+)^n$ and the multiplicative group $(\mathbb{R}^+)^n$ is isomorphic under logarithm to the additive group of $\mathbb{R}^n$. It is easily seen that this group does not contain any nontrivial compact subgroup, hence if $K \subset T$ is compact, by projecting to $(\mathbb{R}^+)^n$ we see that $K \subset U(1,\mathbb{C})^n$.

   The *compact torus* $U(1,\mathbb{C})^n = (S^1)^n$ is the unique maximal compact subgroup of T.
2. The orthogonal group $O(n,\mathbb{C})$. We have $O(n,\mathbb{C}) \cap U(n,\mathbb{C}) = O(n,\mathbb{R})$; thus $O(n,\mathbb{R})$ is a maximal compact subgroup of $O(n,\mathbb{C})$.

Exercise. Describe the orbit map XX^*, $X \in O(n,\mathbb{C})$.

3. The symplectic group and quaternions: We can consider the quaternions $\mathbb{H} := \mathbb{C} + j\mathbb{C}$ with the commutation rules $j^2 = -1$, $j\alpha := \overline{\alpha}j$, $\forall \alpha \in \mathbb{C}$, and set

$$\overline{\alpha + j\beta} := \overline{\alpha} - \overline{\beta}j = \overline{\alpha} - j\beta.$$

Consider the right vector space $\mathbb{H}^n = \bigoplus_{i=1}^n e_i\mathbb{H}$ over the quaternions, with basis e_i. As a right vector space over $\mathbb{C}$ this has as basis $e_1, e_1 j, e_2, e_2 j, \dots, e_n, e_n j$. For a vector $u := (q_1, q_2, \dots, q_n) \in \mathbb{H}^n$ define $\|u\| := \sum_{i=1}^n q_i\overline{q}_i$. If $q_i = \alpha_i + j\beta_i$, we have $\sum_{i=1}^n q_i\overline{q}_i = \sum_{i=1}^n |\alpha_i|^2 + |\beta_i|^2$. Let $Sp(n,\mathbb{H})$ be the group of quaternionic linear transformations preserving this norm. It is easily seen that this group can be described as the group of $n \times n$ matrices $X := (q_{ij})$ with $X^* := \overline{X}^t = X^{-1}$ where X^* is the matrix with $\overline{q}_{ji}$ in the ij entry. This is again clearly a closed bounded group, hence compact.

$$Sp(n,\mathbb{H}) := \{A \in M_n(\mathbb{H}) \mid AA^* = 1\}.$$

On $\mathbb{H}^n = \mathbb{C}^{2n}$, right multiplication by j induces an antilinear transformation, with matrix a diagonal matrix J of 2×2 blocks of the form

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Since a complex $2n \times 2n$ matrix is quaternionic if and only if it commutes with j, we see that the group $Sp(n,\mathbb{H})$ is the subgroup of the unitary group $U(2n,\mathbb{C})$ commuting with the operator j.

If, on a complex vector space, we have a linear operator X with matrix A and an antilinear operator Y with matrix B, it is clear that both XY and YX are antilinear with matrices AB and $B\overline{A}$, respectively. In particular the two operators commute if and only if $AB = B\overline{A}$. We apply this now to $Sp(n, \mathbb{H})$. We see that it is formed by those matrices X in $U(2n, \mathbb{C})$ such that $XJ = J\overline{X} = J(X^{-1})^t$. Its Lie algebra $\mathfrak{k}$ is formed by the anti-Hermitian matrices Y with $YJ = J\overline{Y}$.

Taking $Sp(2n, \mathbb{C})$ to be the symplectic group associated to this matrix J, we have $X \in Sp(2n, \mathbb{C})$ if and only if $X^t J = JX^{-1}$ or $XJ = J(X^{-1})^t$. Thus we have that

$$Sp(n, \mathbb{H}) = U(2n, \mathbb{C}) \cap Sp(2n, \mathbb{C}). \tag{6.3.1}$$

We deduce again that $Sp(n, \mathbb{H})$ is maximal compact in $Sp(2n, \mathbb{C})$.

Exercise. Describe the orbit XX^*, $X \in Sp(2n, \mathbb{C})$.

Although this is not the theme of this book, there are other *real forms* of the groups we studied. For instance, the orthogonal groups or the unitary groups for indefinite forms are noncompact non-algebraic but self-adjoint. We have as further examples:

Proposition. *$O(n, \mathbb{R})$ is maximal compact both in $GL(n, \mathbb{R})$ and in $O(n, \mathbb{C})$.*

7 Hopf Algebras and Tannaka–Krein Duality

7.1 Reductive and Compact Groups

We use the fact, which will be proved in Chapter 10, §7.2, that a reductive group G has a Cartan decomposition $G = Ke^{i\mathfrak{k}}$. Given two rational representations M, N of G we consider them as continuous representations of K.

Lemma. *(1)* $\hom_G(M, N) = \hom_K(M, N)$.

(2) An irreducible representation V of G remains irreducible under K.

Proof. (1) It is enough to show that $\hom_K(M, N) \subset \hom_G(M, N)$.

If $A \in \hom_K(M, N)$, the set of elements $g \in G$ commuting with A is clearly an algebraic subgroup of G containing K. Since K is Zariski dense in G, the claim follows.

(2) is clearly a consequence of (1). □

The next step is to understand:

Proposition. *Let G be a linearly reductive group and K a maximal compact subgroup of G. The restriction map, from the space of regular functions on G to the space of continuous functions on K, is an isomorphism to the space of representative functions of K.*

Proof. First, since the compact group K is Zariski dense in G, the restriction to K of the algebraic functions is injective. It is also clearly equivariant with respect to the left and right action of K.

Since $GL(n,k)$ can be embedded in $SL(n+1,k)$ we can choose a specific faithful representation of G as a self-adjoint group of matrices of determinant 1. In this representation K is the set of unitary matrices in G. The matrix coefficients of this representation as functions on G generate the algebra of regular functions. By Theorem 2.3 the same matrix coefficients generate, as functions on K, the algebra of representative functions. □

Corollary. *The category of finite-dimensional rational representations of G is equivalent to the category of continuous representations of K.*

Proof. Every irreducible representation of K appears in the space of representative functions, while every algebraic irreducible representation of G appears in the space of regular functions. Since these two spaces coincide algebraically the previous lemma (2) shows that all irreducible representations of K are obtained by restriction from irreducible representations of G. The first part of the lemma shows that the restriction is an equivalence of categories. □

In fact we can immediately see that the two canonical decompositions, $\mathcal{T}_K = \bigoplus_{V\in\hat{K}} V^* \otimes V$ (formula 2.1.1) and $k[G] = \bigoplus_i U_i^* \otimes U_i$ of Chapter 7, §3.1.1, coincide under the identification between regular functions on G and representative functions on K.

7.2 Hopf Algebras

We want now to discuss an important structure, the Hopf algebra structure, on the space of representative functions $\mathcal{T}_K$. We will deduce some important consequences for compact Lie groups. Recall that in 2.2 we have seen:

If $f_1(x), f_2(x)$ are representative functions of K, then also $f_1(x)f_2(x)$ is representative.

If $f(x)$ is representative $f(xy)$ is representative as a function on $K \times K$, and it is obvious that $f(x^{-1})$ is representative. Finally

$$\mathcal{T}_{K\times K} = \mathcal{T}_K \otimes \mathcal{T}_K.$$

In the case of a compact group,

$$\mathcal{T}_K = \bigoplus_{i\in\hat{K}} (V_i^* \otimes V_i),$$

$$\mathcal{T}_{K\times K} = \mathcal{T}_K \otimes \mathcal{T}_K = \bigoplus_{i,j} (V_i^* \otimes V_i) \otimes (V_j^* \otimes V_j) = \oplus (V_i \otimes V_j)^* \otimes (V_i \otimes V_j).$$

$\hat{K}$ denotes the set of isomorphism classes of irreducible representations of K.

For simplicity set $\mathcal{T}_K = A$. We want to extract, from the formal properties of the previous constructions, the notion of a (commutative) Hopf algebra.[62]

[62] Hopf algebras appear in various contexts in mathematics. In particular Hopf used them to compute the cohomology of compact Lie groups.

This structure consists of several operations on A. In the general setting A need not be commutative as an algebra.

(1) A is a (commutative and) associative algebra under multiplication with 1. We set $m : A \otimes A \to A$ to be the multiplication.

(2) The map $\Delta : f \to f(xy)$ from A to $A \otimes A$ is called a *coalgebra structure.* It is a homomorphism of algebras and *coassociative* $f((xy)z) = f(x(yz))$ or equivalently, the diagram

$$\begin{tikzcd} A \arrow[r, "\Delta"] \arrow[d, "\Delta"'] & A \otimes A \arrow[d, "1_A \otimes \Delta"] \\ A \otimes A \arrow[r, "\Delta \otimes 1_A"] & A \otimes A \otimes A \end{tikzcd}$$

is commutative. In general Δ is not cocommutative, i.e., $f(xy) \neq f(yx)$.

(3) $(fg)(xy) = f(xy)g(xy)$, that is, Δ is a morphism of algebras. Since $m(f(x) \otimes g(y)) = f(x)g(x)$ we see that also m is a morphism of coalgebras, i.e., the diagram

$$\begin{tikzcd} A \otimes A \arrow[r, "m"] \arrow[d, "\Delta \otimes \Delta"'] & A \arrow[d, "\Delta"] \\ (A \otimes A) \otimes (A \otimes A) \arrow[r, "m \otimes m \circ 1_A \otimes \tau \otimes 1_A"] & A \otimes A \end{tikzcd}$$

is commutative. Here $\tau(a \otimes b) = b \otimes a$.

(4) The map $S : f(x) \to f(x^{-1})$ is called an *antipode.*
Clearly S is a homomorphism of the algebra structure. Also $f(x^{-1}y^{-1}) = f((yx)^{-1})$, hence *$S$ is an anti-homomorphism of the coalgebra structure.*
When A is not commutative the correct axiom to use is that *S is also an anti-homomorphism of the algebra structure.*

(5) It is convenient to also think of the unit element as a map $\eta : \mathbb{C} \to A$ satisfying

$$m \circ (1_A \otimes \eta) = 1_A = m \circ (\eta \otimes 1_A), \ \epsilon\eta = 1_{\mathbb{C}}.$$

(6) We have the *counit* map $\epsilon : f \mapsto f(1)$, an algebra homomorphism $\epsilon : A \to \mathbb{C}$. With respect to the coalgebra structure, we have $f(x) = f(x1) = f(1x)$ or

$$1_A \otimes \epsilon \circ \Delta = \epsilon \otimes 1_A \circ \Delta = 1_A.$$

Also $f(xx^{-1}) = f(x^{-1}x) = f(1)$ or

$$\eta \circ \epsilon = m \circ 1_A \otimes S \circ \Delta = m \circ S \otimes 1_A \circ \Delta.$$

All the previous properties except for the axioms on commutativity or co-commutativity can be taken as the axiomatic definition of a Hopf algebra.[63]

[63] Part of the axioms are dropped by some authors. For an extensive treatment one can see [Ab], [Sw].

Example. When $A = k[x_{i,j}, d^{-1}]$ is the coordinate ring of the linear group we have

$$\Delta(x_{i,j}) = \sum_h x_{i,h} \otimes x_{h,j}, \ \Delta(d) = d \otimes d, \quad \sum x_{i,h} S(x_{h,j}) = \delta_{i,j}. \tag{7.2.1}$$

One clearly has the notion of homomorphism of Hopf algebras, ideals, etc. We leave it to the reader to make explicit what we will use. The way we have set the definitions implies:

Theorem 1. *Given a topological group G, the algebra $\mathcal{T}_G$ is a Hopf algebra. The construction that associates to G the algebra $\mathcal{T}_G$ is a contravariant functor, from the category of topological groups, to the category of commutative Hopf algebras.*

Proof. Apart from some trivial details, this is the content of the propositions of 2.1. □

A commutative Hopf algebra A can be thought of abstractly as a *group in the opposite category of commutative algebras*, due to the following remark.

Given a commutative algebra B let $G_A(B) := \{\phi : A \to B\}$ be the set of homomorphisms.

Exercise. The operations:

$$\phi * \psi(a) := \sum_i \phi(u_i)\psi(v_i), \quad \Delta(a) = \sum_i u_i \otimes v_i,$$

$$\phi^{-1}(a) := \phi(S(a)), \quad 1(a) := \eta(a)$$

are the multiplication, inverse and unit of a group law on $G_A(B)$.

In fact, in a twisted way, these are the formulas we have used for representative functions on a group! The *twist* consists of the fact that when we consider the homomorphisms of A to B as points we should also consider the elements of A as functions. Thus we should write $a(\phi)$ instead of $\phi(a)$. If we do this, all the formulas become the same as for representative functions.

This allows us to go back from Hopf algebras to topological groups. This is best done in the abstract framework by considering Hopf algebras over the real numbers. In the case of groups we must change the point of view and take only real representative functions.

When we work over the reals, the abstract group $G_A(\mathbb{R})$ can be naturally given the *finite topology* induced from the product topology $\prod_{a \in A} \mathbb{R}$ of functions from A to $\mathbb{R}$.

The abstract theorem of Tannaka duality shows that under a further restriction, which consists of axiomatizing the notion of Haar integral for Hopf algebras, we have a duality.

Formally a Haar integral on a real Hopf algebra A is defined by mimicking the group properties $\int f(xy)dy = \int f(xy)dx = \int f(x)dx$:

$$\int : A \to \mathbb{R}, \quad \forall a \in A,$$

$$\Delta(a) = \sum_i u_i \otimes v_i \implies \int a = \sum_i a_i \int v_i = \sum_i u_i \int v_i.$$

One also imposes the further positivity condition: if $a \neq 0$, $\int a^2 > 0$. Under these conditions one has:

Theorem 2. *If A is a real Hopf algebra, with an integral satisfying the previous properties, then $G_A(\mathbb{R})$ is a compact group and A is its Hopf algebra of representative functions.*

The proof is not particularly difficult and can be found for instance in [Ho]. For our treatment we do not need it but rather, in some sense, we need a refinement. This establishes the correspondence between compact Lie groups and linearly reductive algebraic groups.

The case of interest to us is when A, as an algebra, is the coordinate ring of an affine algebraic variety V, i.e., A is finitely generated, commutative and without nilpotent elements.

Recall that giving a morphism between two affine algebraic varieties is equivalent to giving a morphism in the opposite direction between their coordinate rings. Since $A \otimes A$ is the coordinate ring of $V \times V$, it easily follows that the given axioms translate on the coordinate ring the axioms of an algebraic group structure on V.

Also the converse is true. If A is a finitely generated commutative Hopf algebra without nilpotent elements over an algebraically closed field k, then by the correspondence between affine algebraic varieties and finitely generated reduced algebras we see that A is the coordinate ring of an algebraic group. In characteristic 0 the condition to be reduced is automatically satisfied (Theorem 7.3).

Now let K be a linear compact group (K is a Lie group by Chapter 3, §3.2). We claim:

Proposition. *The ring $\mathcal{T}_K$ of representative functions is finitely generated. $\mathcal{T}_K$ is the coordinate ring of an algebraic group G, the* ***complexification*** *of K.*

Proof. In fact, by Theorem 2.2, $\mathcal{T}_K$ is generated by the coordinates of the matrix representation and the inverse of the determinant. Since it is obviously without nilpotent elements, the previous discussion implies the claim. □

We know (Proposition 3.4 and Chapter 4, Theorem 3.2) that linear compact groups are the same as compact Lie groups, hence:

Theorem 3. *To any compact Lie group K there is canonically associated a reductive linear algebraic group G, having the representative functions of K as regular functions.*

G is linearly reductive with the same representations as K. K is maximal compact and Zariski dense in G.

If V is a faithful representation of K, it is a faithful representation of G. For any K-invariant Hilbert structure on V, G is self-adjoint.

Proof. Let G be the algebraic group with coordinate ring $\mathcal{T}_K$. By definition its points correspond to the homomorphisms $\mathcal{T}_K \to \mathbb{C}$. In particular evaluating the functions of $\mathcal{T}_K$ in K we see that $K \subset G$ is Zariski dense. Therefore, by the argument in 7.1, every K-submodule of a rational representation of G is automatically a G-submodule. Hence the decomposition $\mathcal{T}_K = \bigoplus_i V_i^* \otimes V_i$ is in $G \times G$-modules, and G is linearly reductive with the same irreducible representations as K.

Let $K \subset H \subset G$ be a larger compact subgroup. By definition of G, the functions $\mathcal{T}_K$ separate the points of G and hence of H. $\mathcal{T}_K$ is closed under complex conjugation so it is dense in the space of continuous functions of H. The decomposition $\mathcal{T}_K = \bigoplus_i V_i^* \otimes V_i$ is composed of irreducible representations of K and G; it is also composed of irreducible representations of H. Thus the Haar integral performed on H is 0 on all the nontrivial irreducible summands. Thus if we take a function $f \in \mathcal{T}_K$ and form its Haar integral either on K or on H we obtain the same result. By density this then occurs for all continuous functions. If $H \neq K$ we can find a nonzero, nonnegative function f on H, which vanishes on K, a contradiction.

The matrix coefficients of a faithful representation of K generate the algebra $\mathcal{T}_K$. So this representation is also faithful for G. To prove that $G = G^*$ notice that although the map $g \mapsto g^*$ is not algebraic, it is an antilinear map, so it transforms affine varieties into affine varieties (conjugating the coefficients in the equations), and thus G^* is algebraic and clearly K^* is Zariski dense in G^*. Since $K^* = K$ we must have $G = G^*$. □

At this point, since G is algebraic, it has a finite number of connected components; using the Cartan decomposition of 6.1 we have:

Corollary. *(i) The Lie algebra $\mathfrak{g}$ of G is the complexification of the Lie algebra $\mathfrak{k}$ of K.*

(ii) One has the Cartan decomposition $G = K \times e^{i\mathfrak{k}}$.

7.3 Hopf Ideals

The definition of Hopf algebra is sufficiently general so that it does not need to have a base coefficient field. For instance, for the general linear group we can work over $\mathbb{Z}$, or even any commutative base ring. The corresponding Hopf algebra is $A[n] := \mathbb{Z}[x_{i,j}, d^{-1}]$, where $d = \det(X)$ and X is the *generic matrix* with entries $x_{i,j}$. The defining formulas for Δ, S, η are the same as in 7.2.1. One notices that by Cramer's rule, the elements $d\, S(x_{i,j})$ are the *cofactors*, i.e., the entries of $\bigwedge^{n-1} X$. These are all polynomials with integer coefficients.

To define a Hopf algebra corresponding to a subgroup of the linear group one can do it by constructing a *Hopf ideal.*

Definition. A *Hopf ideal* of a Hopf algebra A is an ideal I such that

$$(7.3.1) \qquad \Delta(I) \subset I \otimes A + A \otimes I, \qquad S(I) \subset I, \qquad \eta(I) = 0.$$

Clearly, if I is a Hopf ideal, A/I inherits a structure of a Hopf algebra such that the quotient map, $A \twoheadrightarrow A/I$ is a homomorphism of Hopf algebras.

As an example let us see the orthogonal and symplectic group over $\mathbb{Z}$. It is convenient to write all the equations in an intrinsic form using the generic matrix X. We do the case of the orthogonal group, the symplectic being the same. The ideal I of the orthogonal group by definition is generated by the entries of the equation $XX^t - 1 = 0$. We have

$$\begin{aligned}
&(7.3.2) \qquad \Delta(XX^t - 1) = XX^t \otimes XX^t - 1 \otimes 1 \\
&\qquad\qquad = (XX^t - 1) \otimes XX^t + 1 \otimes (XX^t - 1) \\
&(7.3.3) \qquad S(XX^t - 1) = S(X)S(X^t) - 1 = d^{-2} \bigwedge\nolimits^{n-1}(X) \bigwedge\nolimits^{n-1}(X^t) - 1 \\
&\qquad\qquad = d^{-2} \bigwedge\nolimits^{n-1}(XX^t) - 1 \\
&(7.3.4) \qquad \eta(XX^t - 1) = \eta(X)\eta(X^t) - 1 = 1 - 1 = 0.
\end{aligned}$$

Thus the first and last conditions for Hopf ideals are verified by 7.3.2 and 7.3.4. To see that $S(I) \subset I$ notice that modulo I we have $XX^t = 1$, hence $d^2 = 1$ and $\bigwedge^{n-1}(XX^t) = \bigwedge^{n-1}(1) = 1$ from which it follows that modulo I we have $S(XX^t - 1) = 0$.

Although this discussion is quite satisfactory from the point of view of Hopf algebras, it leaves open the geometric question whether the ideal we found is really the full ideal vanishing on the geometric points of the orthogonal group. By the general theory of correspondence between varieties and ideal this is equivalent to proving that $A[n]/I$ has no nilpotent elements.

If instead of working over $\mathbb{Z}$ we work over $\mathbb{Q}$ and we can use a very general fact [Sw]:

Theorem 1. *A commutative Hopf algebra A over a field of characteristic* 0 *has no nilpotent elements (i.e., it is reduced).*

Proof. Let us see the proof when A is finitely generated over $\mathbb{C}$. It is possible to reduce the general case to this. By standard facts of commutative algebra it is enough to see that the localization A_m has no nilpotent elements for every maximal ideal $\mathfrak{m}$. Let G be the set of points of A, i.e., the homomorphisms to $\mathbb{C}$. Since G is a group we can easily see (thinking that A is like a ring of functions) that G acts as group of automorphisms of A, transitively on the points. In fact the analogue of the formula for $f(xg)$ when $g : A \to \mathbb{C}$ is a point is the composition $R_g : A \xrightarrow{\Delta} A \otimes A \xrightarrow{1 \otimes g} A \otimes \mathbb{C} = A$.

It follows from axiom (5) that $g = \epsilon \circ R_g$, as desired. Thus it suffices to see that A, localized at the maximal ideal $\mathfrak{m}$, kernel of the counit ϵ (i.e., at the point 1) has no nilpotent elements. Since the intersection of the powers of the maximal ideal is 0, this is equivalent to showing that $\bigoplus_{i=1}^{\infty} \mathfrak{m}^i/\mathfrak{m}^{i+1}$ has no nilpotent ideals.[64] If $m \in \mathfrak{m}$

[64] One thinks of this ring as the coordinate ring of the tangent cone at 1.

and $\Delta(m) = \sum_i x_i \otimes y_i$ we have $m = \sum_i \epsilon(x_i) y_i = \sum_i x_i \epsilon(y_i),\ 0 = \sum_i \epsilon(x_i)\epsilon(y_i)$. Hence

$$\Delta(m) = \sum_i x_i \otimes y_i - \sum_i \epsilon(x_i) \otimes y_i + \sum_i \epsilon(y_i) \otimes x_i - \sum_i \epsilon(x_i)\epsilon(y_i)$$

$$(7.3.5) \qquad = \sum_i (x_i - \epsilon(x_i)) \otimes y_i + \sum_i \epsilon(y_i) \otimes (x_i - \epsilon(x_i)) \in \mathfrak{m} \otimes 1 + 1 \otimes \mathfrak{m}.$$

Similarly, $S(\mathfrak{m}) \subset \mathfrak{m}$. It follows easily that $B := \bigoplus_{i=1}^{\infty} \mathfrak{m}^i/\mathfrak{m}^{i+1}$ inherits the structure of a commutative graded Hopf algebra, with $B_0 = \mathbb{C}$. Graded Hopf algebras are well understood; in fact in a more general settings they were originally studied by Hopf as the cohomology algebras of Lie groups. In our case the theorem we need says that B is a polynomial ring, hence an integral domain, proving the claim. □

The theorem we need to conclude is an extremely special case of a general theorem of Milnor and Moore [MM]. Their theory generalizes the original theorem of Hopf, which was only for finite-dimensional graded Hopf algebras and treats several classes of algebras, in particular, the ones which are generated by their primitive elements (see the end of the next section).

We need a special case of the characterization of graded connected commutative and co-commutative Hopf algebras. Graded commutative means that the algebra satisfies $ab = (-1)^{|a||b|}ba$ where $|a|, |b|$ are the degrees of the two elements. The condition to be *connected* is simply $B_0 = \mathbb{C}$. In case the algebra is a cohomology algebra of a space X it reflects the condition that X is connected. The usual commutative case is obtained when we assume that all elements have even degree. In our previous case we should consider $\mathfrak{m}/\mathfrak{m}^2$ as in degree 2. In this language one unifies the notions of symmetric and exterior powers: one thinks of a usual symmetric algebra as being generated by elements of even degree and an extrerior algebra is still called by abuse a symmetric algebra, but it is generated by elements of odd degree. In more general language one can talk of the *symmetric algebra*, $S(\underline{V})$ of a graded vector space $\underline{V} = \sum V_i$, which is $S\big(\sum_i V_{2i}\big) \otimes \bigwedge\big(\sum_i V_{2i+1}\big)$.

One of the theorems of Milnor and Moore. *Let B be a finitely generated*[65] *positively graded commutative and connected; then B is the symmetric algebra over the space $P := \{u \in B \mid \Delta(u) = u \otimes 1 + 1 \otimes u\}$ of* ***primitive elements.***

We need only a special case of this theorem, so let us show only the very small part needed to finish the proof of Theorem 1.

Finishing the proof. In the theorem above $B := \bigoplus_{i=1}^{\infty} \mathfrak{m}^i/\mathfrak{m}^{i+1}$ is a graded commutative Hopf algebra generated by the elements of lowest degree $\mathfrak{m}/\mathfrak{m}^2$ (we should give to them degree 2 to be compatible with the definitions). Let $x \in \mathfrak{m}/\mathfrak{m}^2$. We have $\Delta x = a \otimes 1 + 1 \otimes b,\ a, b \in \mathfrak{m}/\mathfrak{m}^2$ by the minimality of the degree. Applying axiom (5) we see that $a = b = x$ and x is primitive. What we need to prove is thus that if

[65] This condition can be weakened.

$x_1, \dots, x_n$ constitute a basis of $\mathfrak{m}/\mathfrak{m}^2$, then the x_i are algebraically independent. Assume by contradiction that $f(x_1, \dots, x_n) = 0$ is a homogeneous polynomial relation of minimum degree h. We also have

$$0 = \Delta f(x_1, \dots, x_n) = f(x_1 \otimes 1 + 1 \otimes x_1, \dots, x_n \otimes 1 + 1 \otimes x_n) = 0.$$

Expand $\Delta f \in \sum_{i=0}^{h} B_{h-i} \otimes B_i$ and consider the term $T_{h-1,1}$ of bidegree $h-1, 1$. This is really a polarization and in fact it is $\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_1, \dots, x_n) \otimes x_j$. Since the x_i are linearly independent the condition $T_{h-1,1} = 0$ implies $\frac{\partial f}{\partial x_j}(x_1, \dots, x_n) = 0, \forall j$. Since we are in characteristic 0, at least one of these equations is nontrivial and of degree $h-1$, a contradiction. $\square$

As a consequence, a Hopf ideal of the coordinate ring of an algebraic group in characteristic 0 is always the defining ideal of an algebraic subgroup.

Exercise. Let G be a linear algebraic group, $\rho : G \to GL(V)$ a linear representation and $v \in V$ a vector. Prove that the ideal of the stabilizer of v generated by the equations $\rho(g)v - v$ is a Hopf ideal.

It is still true that the algebra modulo the ideal I generated by the entries of the equations $XX^t = 1$ has no nilpotent ideals when we take as coefficients a field of characteristic $\neq 2$.

The proof requires a little commutative algebra (cf. [E]). Let k be a field of characteristic $\neq 2$. The matrix $XX^t - 1$ is a symmetric $n \times n$ matrix, so the equations $XX^t - 1 = 0$ are of dimension $\binom{n+1}{2}$, while the dimension of the orthogonal group is $\binom{n}{2}$ (this follows from Cayley's parametrization in any characteristic $\neq 2$) and $\binom{n+1}{2} + \binom{n}{2} = n^2$ the number of variables. We are thus in the case of a *complete intersection*, i.e., the number of equations equals the codimension of the variety. Since a group is a smooth variety we must then expect that the Jacobian of these equations has everywhere maximal rank. In more geometric language let $S_n(k)$ be the space of symmetric $n \times n$ matrices. Consider the mapping $\pi : M_n(k) \to S_n(k)$ given by $X \to XX^t$. In order to show that for some $A \in S_n(k)$ the equations $XX^t = A$ generate the ideal of definition of the corresponding variety, it is enough to show that the differential $d\pi$ of the map is always surjective on the points X such that $XX^t = A$. The differential can be computed by substituting for X a matrix $X + Y$ and saving only the linear terms in Y, getting the formula $YX^t + XY^t = YX^t + (YX^t)^t$.

Thus we have to show that given any symmetric matrix Z, we can solve the equation $Z = YX^t + (YX^t)^t$ if $XX^t = 1$. We set $Y := ZX/2$ and have $Z = 1/2(ZXX^t + (ZXX^t)^t)$.

In characteristic 2 the statement is simply not true since

$$\sum_j x_{i,j}^2 - 1 = \Big(\sum_j x_{i,j} - 1\Big)^2.$$

So $\sum_j x_{i,j} - 1$ vanishes on the variety but it is not in the ideal.

Exercise. Let L be a Lie algebra and U_L its universal enveloping algebra. Show that U_L is a Hopf algebra under the operations defined on L as

$$(7.3.6) \qquad \Delta(a) = a \otimes 1 + 1 \otimes a, \quad S(a) = -a, \quad \eta(a) = 0, \qquad a \in L.$$

Show that $L = \{u \in U_L \mid \Delta(u) = u \otimes 1 + 1 \otimes u\}$, the set of *primitive elements*. Study the Hopf ideals of U_L.

Remark. One of the theorems of Milnor and Moore is the characterization of universal enveloping algebras of Lie algebras as suitable primitively generated Hopf algebras.