Decomposition of spin as irreducible representations of $SU(2)$

Alexandre René

Project submitted in partial fulfilment of the requirements for the degree of B.Sc.
Honours Specialisation in Mathematics

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

© Alexandre René, Ottawa, Canada, 2012

Abstract

Expressing the symmetries of a physical system in terms of groups is a very fruitful approach and leads to the rich theory of representations. In this paper, we propose to use the tools of this theory to treat the problem of quantum mechanical spin. These are developed first in the case when the group is finite, then extended to case where it is compact. We then show that the Clebsch-Gordan decomposition of spin spaces can be obtained from the irreducible representations of SU(2).

Résumé

Exprimer les symétries d'un système physique en termes de groupes est une approche souvent très profitable qui mène à la théorie des représentations. Dans cet exposé, nous utilisons les outils de cette théorie pour traiter le problème du spin quantique. Ces outils sont d'abord développés dans le context des groupes finis, puis étendus au cas des groupes compacts. Nous montrons ensuite que la décomposition de Clebsch-Gordan des états de spin peut être obtenue à partir des représentations irréductibles de SU(2).

Contents

Introduction **1**

1 Finite Groups **2**
1.1 Definitions 2
1.2 Characters 7
1.3 Schur's lemma 10
1.4 Orthogonality relations of characters 12
1.5 Regular Representation 16

2 Compact Groups **20**
2.1 Measure and integration 20
2.2 Haar measure 23
2.3 Representations of compact groups 29
2.4 The Peter-Weyl theorem 32

3 Spin and SU(2) **35**
3.1 Fundamental concepts 35
3.2 The Haar measure of SU(2) 36
3.3 The irreducible representations of $SU(2)$ 40
3.4 Application to spin 46

Conclusion **48**

Bibliography **49**

Introduction

As quantum mechanics was developed, it was discovered that particles exhibit a type of intrinsic angular momentum, which we now call *spin* (see for example [2]). This property has no analog in Newtonian physics, and thus can only be understood quantum mechanically. Partly because of this, it forms the basis of much of the novel phenoma we observe at the quantum scale: quantum computer implementation, superconduction models and the entire field of spintronics are but a few examples of effects and applications that crucially depend on its quirky properties.

Mathematically, and just like any other quantum mechanical property, the spin of a particle is simply a vector in a vector space of possible states. Determining this vector space, however, can be exceedingly difficult. A mathematical approach to the problem would be to determine the symmetries of the vector space, in other words to find a group which leaves the space invariant. By constructing a *representation* of this group onto the vector space, it is possible to obtain a wealth of information on the nature of the space. The first chapter of this paper will develop the representation theory of finite groups and the use of character functions to find decompositions of representations.

Unfortunately in the case of spin, the group in question is SU(2), which is not finite. It is however compact, and we can extend most of the results on representations of finite group to compact ones. This extension will be the subject of the second chapter. In particular, we will need the Peter-Weyl theorem, which will conclude the chapter.

The third and final chapter will analyze the problem of spin in quantum systems. The minimal physical framework to state the problem will be presented, following which we will solve the problem using the tools developed in the two preceding chapters.

Chapter 1

Representations of Finite Groups

This first chapter presents the representation theory of finite groups. Working first with finite groups allows a more intuitive presentation of the concepts; it also provides the basis upon which the more general theory over compact groups is built.

The material presented here should not be new to anyone familiar with representation theory. A more detailed presentation can be found in introductory textbooks, for example one by J-P Serre [6].

In this chapter, G will always denote a finite group.

1.1 Fundamental definitions

Definition 1.1.1. *Let G be a group and V a vector space. A* representation *of G in V is a map $\rho\colon G \to \mathrm{GL}(V)$ such that $\rho(st) = \rho(s)\rho(t)$.*

Notation 1.1.2. For a fixed group G, we will denote a representation ρ in V by (ρ, V).

Hence, a representation is a group homomorphism of G in $\mathrm{GL}(V)$. In order to maintain legibility, we will usually write $\rho(s)$ as ρ_s. Also, it is useful to note that the fact that ρ is a homomorphism implies that for all $t \in G$, $\rho_t^{-1} = \rho_{t^{-1}}$.

Remark 1.1.3. In this project, the vector space V will always be a complex vector space. In particular, since ρ_s is invertible for any s in G, its characteristic polynomial must be of same degree as the dimension n of V, and therefore yields n eigenvalues.

Definition 1.1.4. *If (ρ, V) is a representation of G, the degree of the representation ρ is the dimension of V.*

Notation 1.1.5. Given a representation (ρ, V) of degree n, it is often useful to consider the matrix representation of a map ρ_s. This will be an invertible $n \times n$ matrix, denoted $R(s) = (r_{ij}(s))$.

Example 1.1.6. Consider S_3, the group of all permutations of a set $X = \{v_1, v_2, v_3\}$ of three elements. This group contains six elements: 1 (the identity permutation), three transpositions and two cyclic permutations. A trivial representation of S_3 would be the identity representation:

$$\begin{array}{rccccccc} \rho_{\mathrm{id}}\colon & S_3 & \to & \mathrm{End}(\mathbb{C}) & & & & \\ & \sigma & \mapsto & \mathrm{id}_{\mathbb{C}}\colon & \mathbb{C} & \to & \mathbb{C} \\ & & & & v & \mapsto & v \end{array}$$

(here we view $\mathbb{C}$ as a one-dimensional vector space).

Example 1.1.7. A more interesting example occurs when we consider the elements of X to be the coordinates of an equilateral triangle in $\mathbb{R}^2$ (see Figure 1.1.7). Recall that the *dihedral group* D_n is the set of rotations and reflections of a regular n-sided polygon which conserve its vertices. It is easy to see that here S_3 is in fact isomorphic to D_3, with transpositions being equivalent to reflections about symmetry planes and cyclic permutations being equivalent to rotations of $\pm 2\pi/3$. In fact, viewing the action of S_3 this way defines a natural representation $(\rho, \mathbb{R}^2)$. To write it out explicitly, let $c = (v_2 v_3 v_1)$ and $t = (v_2 v_3)$. Then any element of S_3 can be written uniquely as either c^k or tc^k, for $0 \leqslant k < 2$, and ρ is written in matrix form as

$$\rho(c^k) = \begin{pmatrix} \cos(ka) & -\sin(ka) \\ \sin(ka) & \cos(ka) \end{pmatrix}, \qquad \rho(tc^k) = \begin{pmatrix} \cos(ka) & -\sin(ka) \\ -\sin(ka) & -\cos(ka) \end{pmatrix} \tag{1.1.1}$$

where $a = 2\pi/3$.

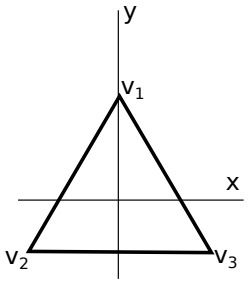

Figure 1.1: *The triangle on which the action of S_3 is defined in Example 1.1.7.*

Definition 1.1.8. *Let (ρ, V) and (ρ', V') be two representations of a group G. Then ρ and ρ' are* equivalent *or* isomorphic *if there exists a linear isomorphism $\phi\colon V \to V'$ such that $\phi \circ \rho_s = \rho'_s \circ \phi$ for all $s \in G$. Such a ϕ is called an* intertwining operator.

Definition 1.1.9. *Let (ρ, V) be a representation of G. We say that a subspace $W \subset V$ is a* subrepresentation *of V if it is stable under G.*

In other words, if W is a subrepresentation of V, we have $\rho(t)W \subset W$, $\forall t \in G$. Then the restriction ρ^W defined by $\rho^W(t) = \rho(t)|_W$ is a representation of G in W.

Definition 1.1.10. *A representation on a vector space V is said to be* irreducible *if its only subrepresentations are $\{0\}$ and V.*

Example 1.1.11. Take the group S_3 of permutations and construct the 6-dimensional vector space V by assigning each element of S_3 to a basis vector of V. In other words, the basis of V is $\{e_\sigma\}_{\sigma \in S_3}$ where $\sigma \in \{(123), (231), (312), (132), (213), (321)\}$. We can define a representation (r, V) by

$$r_\sigma e_\tau = e_{\sigma\tau} \qquad \sigma, \tau \in S_3. \tag{1.1.2}$$

In matrix form, this is given explicitly by

$$\begin{aligned} r_{(123)} &= \begin{pmatrix} A^0 & 0 \\ 0 & A^0 \end{pmatrix} \quad r_{(231)} = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \quad r_{(312)} = \begin{pmatrix} A^2 & 0 \\ 0 & A^2 \end{pmatrix} \\ r_{(132)} &= \begin{pmatrix} 0 & A^0 \\ A^0 & 0 \end{pmatrix} \quad r_{(213)} = \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix} \quad r_{(321)} = \begin{pmatrix} 0 & A^2 \\ A^2 & 0 \end{pmatrix} \end{aligned} \tag{1.1.3}$$

where $A^0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ and $A^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.

This representation r is the *regular representation* of S_3 (see Section 1.5). Note that if we consider the subspace $W = \mathrm{Span}\{(1,1,1,1,1,1)^\intercal\}$ of V, we notice that the action of S_3 is invariant on W; in fact, the restriction of r to W is exactly the identity representation defined in Example 1.1.6. Hence ρ_{id} is a subrepresentation of r. It is also irreducible, by virtue of being of dimension 1.

One of the crucial results upon which most of the representation theory of finite groups relies is that any representation can be expressed as a direct sum of irreducible representations. This is the subject of the following proposition and theorem. The equivalent result for compact groups is the Peter-Weyl theorem, which is one of the objectives of this paper.

Recall that if there exist spaces V, W and W' such that $V = W + W'$ and $W \cap W' = \{0\}$, then we write $V = W \oplus W'$ and V is said to be the *direct sum* of W and W'. Then for each vector v of V, there exist unique vectors $w \in W$, $w' \in W'$ such that $v = w + w'$. The vector space W' is also said to be the *supplementary* of W and the linear map

$$\begin{array}{rccc} p\colon & W \oplus W' & \to & W \\ & (v, v') & \mapsto & v \end{array}$$

is called the *projector* of V on W. These notions can be directly translated to representations:

Definition 1.1.12. *We say that a representation ρ of G on $V = W_1 \oplus W_2$ is the* direct sum *of representations (ρ^1, W_1) and (ρ^2, W_2) if, for all $v = w_1 + w_2$, $s \in G$, $w_i \in W_i$, we have $\rho_s(v) = \rho_s(w_1 + w_2) = \rho_s^1(w_1) + \rho_s^2(w_2)$.*

This definition of course extends to any finite sum of vector spaces. It is easy to see that the resulting representation will be block diagonal; for example, if $R^1(s)$ and $R^2(s)$ are the matrix forms of two representations (as in Notation 1.1.5), then

$$R(s) = (R^1 \oplus R^2)(s) = \begin{pmatrix} R^1(s) & 0 \\ 0 & R^2(s) \end{pmatrix} \tag{1.1.4}$$

Remark 1.1.13. We will routinely associate a representation with the target vector space. In this case, we say that the representation V is the direct sum of W_1 and W_2, and write $V = W_1 \oplus W_2$.

Proposition 1.1.14. *Let ρ be a representation of a group G in a vector space V and W a subspace of V. If W is stable under G, then there exists a supplementary space W^0 to W such that W_0 is also stable under G.*

Proof: Let $\rho\colon G \to GL(V)$ be a representation of G in V and let W' be an arbitrary supplementary subspace of W. Given $p\colon V \to W$ the projector corresponding to W', define the new projection $p^0\colon V \to W$

$$p^0 = \frac{1}{g} \sum_{t \in G} \rho_t \cdot p \cdot \rho_t^{-1}$$

where g is the order of G. Since for all t, $p\rho_t^{-1}\colon V \to W$ and $\rho_t(W) \subset W$, we see that p^0 indeed sends V onto W. Furthermore, we also have $\rho_t^{-1}(W) = \rho_{t^{-1}}(W) \subset W$. Therefore, for any v in W, $p \cdot \rho_t^{-1} v = \rho_{t^{-1}} v$ and $\rho_t p \cdot \rho_t^{-1} v = \rho_t \rho_{t^{-1}} v = v$. Thus $p^0 v = v$, which shows that p^0 is indeed a projection. Additionnaly, p^0 has the property that it commutes with ρ_s for all $s \in G$. Indeed, we have that

$$\rho_s p^0 \rho_s^{-1} = \frac{1}{g} \sum_{t \in G} \rho_s \rho_t p \rho_t^{-1} \rho_s^{-1} = \frac{1}{g} \sum_{t \in G} \rho_{st} p \rho_{st}^{-1} = \frac{1}{g} \sum_{t \in G} \rho_t p \rho_t^{-1} = p^0$$

Hence

$$\rho_s p^0 = p^0 \rho_s$$

We define the subspace W^0 as $\{v \in V \mid p^0(v) = 0\}$. Hence, given any $v \in W^0$ and $s \in G$, we have that $p^0 \rho_s v = \rho_s p^0 v = 0$, which shows that W^0 is stable under G. ∎

Note that a consequence of this proposition is that if W is a subrepresentation of V, then the representation V is the direct sum of the representations W and W^0 (where W^0 is defined as in the proposition).

Recall that a Euclidean space is a finite dimensional vector space V endowed with an inner product $(\cdot \mid \cdot)\colon V \to V$. Then a linear map U on V is unitary if $(v \mid w) = (Uv \mid Uw)$ for all v, w in V. Extending this concept to representations, we get the following definition:

Definition 1.1.15. *A representation ρ on a Euclidean space V (or, more generally, a Hilbert space) is said to be* unitary *if ρ_s is unitary for every s in G.*

It turns out that any representation of a finite group on a Euclidean space can be made unitary.

Proposition 1.1.16. *Let ρ be a representation of G on a Euclidean space V, with scalar product $(\cdot \mid \cdot)\colon V \times V \to \mathbb{C}$. There exists a scalar product $(\cdot \mid \cdot)'$ on V with respect to which ρ_s is unitary for every $s \in G$.*

Proof: If we define $(\cdot \mid \cdot)'$ as

$$\begin{array}{rccc} (\cdot \mid \cdot)'\colon & V \times V & \to & \mathbb{C} \\ & (u, v) & \mapsto & \sum_{t \in G} (\rho_t(u) \mid \rho_t(v)) \end{array}$$

then it is also a scalar product. Furthermore, for any u,v in V, s in G, we have

$$\begin{aligned} (\rho_s(u) \mid \rho_s(v))' &= \sum_{t \in G} (\rho_t(\rho_s(u)) \mid \rho_t(\rho_s(v))) \\ &= \sum_{t \in G} (\rho_{ts}(u) \mid \rho_{ts}(v)) \\ &= (u \mid v)' \end{aligned}$$

which shows that ρ_s isometric. And since ρ_s is already bijective, it is also unitary. ∎

Theorem 1.1.17. *Every representation is a direct sum of irreducible representations.*

Proof: Let V be a representation of G. We prove this by induction on the dimension of V. If $\dim(V) = 0$, there is nothing to prove. Suppose then that $\dim(V) \geqslant 1$. If V is irreducible, the proof is complete. If instead V is reducible, then by Proposition 1.1.14, there exist two G-invariant subspaces W, W' of V such that $V = W \oplus W'$. Since the dimensions of W and W' are smaller than that of V, by the induction hypothesis, they can be decomposed into irreducible represention. The same is true of V. ∎

Recall that, given vector spaces V and W of dimension n and m, we can form their tensor product $V \otimes W$. Also, if $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_m\}$ are bases for V and W, then $\{v_i \otimes w_j \mid \ 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m\}$ is a basis of $V \otimes W$. Therefore, any linear map on $V \otimes W$ defined on a basis can be extended by linearity to the whole space.

Definition 1.1.18. *Let G be a group and V and W be vector spaces as above. Let (ρ, V) and (ρ', W) be representations. If s is an element of G, we define a map π_s on the basis elements of $V \otimes W$ by $\pi_s(v_i \otimes w_j) = \rho_s(v_i) \otimes \rho'_s(w_j)$. Extending this map by linearity, we obtain the representation $(\rho \otimes \rho', V \otimes W)$:*

$$\begin{array}{rccc} \rho \otimes \rho' \colon & G & \to & \mathrm{GL}(V \otimes W) \\ & s & \mapsto & (\rho \otimes \rho')_s \colon \end{array} \begin{array}{ccc} V \otimes W & \to & V \otimes W \\ \sum_{i=1}^{n} \sum_{j=1}^{m} a_{i,j} v_i \otimes w_j & \mapsto & \sum_{i=1}^{n} \sum_{j=1}^{m} a_{i,j} \pi_s(v_i \otimes w_j) \end{array} . \tag{1.1.5}$$

Remark 1.1.19. If we have the matrix forms of the representations ρ and ρ':

$$\rho^1_s(v_{i_1}) = \sum_{j_1} r_{i_1 j_1}(s) v_{j_1} \qquad\qquad \rho^2_s(w_{i_2}) = \sum_{j_2} r_{i_2 j_2}(s) w_{j_2},$$

we can obtain the matrix form of the representation $\rho \otimes \rho'$:

$$\begin{aligned} (\rho \otimes \rho')_s(v_{i_1} \otimes w_{i_2}) &= \rho_s(v_{i_1}) \otimes \rho'_s(w_{i_2}) \\ &= \sum_{j_1, j_2} r_{i_1 j_1}(s) r_{i_2 j_2}(s)\, v_{j_1} \otimes w_{j_2}. \end{aligned} \tag{1.1.6}$$

Note that the representation $\rho^1 \otimes \rho^2$ is not irreducible in general. Finding its irreducible representations is one of the uses of character theory, which we develop in the following section.

1.2 The character of a representation

Recall that, when given a finite vector space V, the trace is a map $\mathrm{Tr}\colon \mathrm{End}(V) \to \mathbb{C}$, and that if we are given a matrix representation of an operator $A \in \mathrm{End}(V)$, then $\mathrm{Tr}(A)$ is simply the sum of the diagonal elements of the matrix. Therefore, for a given representation (ρ, V), we can consider the trace of each ρ_s. This turns out to be extremely helpful.

Definition 1.2.1. *Let (ρ, V) be a representation of G. We define the* character χ *of ρ as*

$$\begin{array}{rccc} \chi_\rho \colon & G & \to & \mathbb{C} \\ & s & \mapsto & \mathrm{Tr}(\rho_s) \end{array}$$

Example 1.2.2. Let D_n be the dihedral group realized as in Example 1.1.7. Let r denote the rotation by an angle $2\pi/n$ and s one of the symmetries of the polygon. Then any element of D_n can be written either as r^k or sr^k, for some integer $0 \leqslant k < n$. Furthermore, we have that

$$r^n = 1 \qquad\qquad s^2 = 1 \qquad\qquad srs = r^{-1}$$

If we let $w = e^{i2\pi/n}$ and h be an integer strictly between 0 and $\frac{n}{2}$, then it can be verified that the following defines a representation $(\rho^h, \mathbb{C}^3)$ of D_n when n is even:

$$\begin{aligned}
\rho^h_{r^k} &= \begin{pmatrix} \frac{1}{2}((-1)^k + w^{-hk}) & 0 & \frac{1}{2}((-1)^k - w^{-hk}) \\ \frac{1}{2}(-w^{hk} + w^{-hk}) & w^{hk} & \frac{1}{2}(w^{hk} - w^{-hk}) \\ \frac{1}{2}((-1)^k - w^{-hk}) & 0 & \frac{1}{2}((-1)^k + w^{-hk}) \end{pmatrix} \\
\rho^h_{sr^k} &= \begin{pmatrix} \frac{1}{2}(-1)^k - \frac{1}{2\sqrt{3}}w^{hk} & \frac{w^{hk}}{\sqrt{3}} & \frac{1}{2}(-1)^k + \frac{1}{2\sqrt{3}}w^{hk} \\ \frac{1}{2\sqrt{3}}(-w^{hk} + 3w^{-hk}) & \frac{w^{hk}}{\sqrt{3}} & \frac{1}{2\sqrt{3}}(w^{hk} - 3w^{-hk}) \\ \frac{1}{2}(-1)^k + \frac{1}{2\sqrt{3}}w^{hk} & -\frac{w^{hk}}{\sqrt{3}} & \frac{1}{2}(-1)^k - \frac{1}{2\sqrt{3}}w^{hk} \end{pmatrix}
\end{aligned} \tag{1.2.1}$$

(For instance, direct multiplications shows that $\rho^h_{r^k}\rho^h_{r^l} = \rho^h_{r^{(k+l)}}$, $(\rho^h_s)^2 = 1$ and $(\rho^h_{sr})^n = 1$.)

The character of this representation is given by

$$\begin{aligned}
\chi(\rho^h_{r^k}) &= (-1)^k + w^{hk} + w^{-hk} \\
\chi(\rho^h_{sr^k}) &= (-1)^k
\end{aligned} \tag{1.2.2}$$

Some basic but useful properties of characters are given in the following proposition.

Proposition 1.2.3. *Let (ρ, V) be a representation of degree n on a group G, and χ_ρ be its associated character. Then*

i) $\chi_\rho(1) = n$

ii) $\chi_\rho(s^{-1}) = \overline{\chi_\rho(s)}$

iii) $\chi_\rho(tst^{-1}) = \chi_\rho(s)$

(Here $\overline{z}$ denotes the complex conjugate of z.)

Proof:

i) This follows from the fact that ρ_1 is the identity matrix of size n.

ii) Since the order of G is finite, there exists $q \in \mathbb{N}$ such that $\rho^q_s = 1$. Then, for each of ρ_s's n eigenvalues λ_i (Remark 1.1.3), $\lambda^q_i = 1$. Hence $|\lambda_i| = 1$ and

$$\overline{\chi(s)} = \overline{\mathrm{Tr}(\rho_s)} = \sum_{i=1}^{n} \overline{\lambda_i} = \sum_{i=1}^{n} \lambda_i^{-1} = \mathrm{Tr}(\rho_s^{-1}) = \chi_\rho(s^{-1})$$

iii) This follows from the fact that for any square matrices A and B of same dimension, $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. Then

$$\chi_\rho(tst^{-1}) = \chi_\rho(t^{-1}ts) = \chi_\rho(s)$$

∎

Remark 1.2.4. Recall that a function $f\colon G \to \mathbb{C}$ such that $f(tst^{-1}) = f(s)$, for all s, t in G is called a *class function*. Part (iii) of the proposition above shows that the character of a representation is such a function.

Proposition 1.2.5. *Let V_1 and V_2 be vector spaces, (ρ^1, V_1) and (ρ^2, V_2) be representations and χ^1 and χ^2 be their respective characters. Then*

i) the character χ of the representation $(\rho^1 \oplus \rho^2, V_1 \oplus V_2)$ is given by $\chi^1 + \chi^2$.

ii) the character ψ of the representation $(\rho^1 \otimes \rho^2, V_1 \otimes V_2)$ is given by $\chi^1 \cdot \chi^2$.

Proof:

i) Let $R^1(s)$ and $R^2(s)$ be the matrix forms of the representations ρ^1 and ρ^2. Then as in (1.1.4), the matrix form $R(s)$ of the representation $\rho^1 \oplus \rho^2$ is

$$R(s) = \begin{pmatrix} R^1(s) & 0 \\ 0 & R^2(s) \end{pmatrix}$$

Therefore $\chi(s) = \mathrm{Tr}(R(s)) = \mathrm{Tr}(R^1(s)) + \mathrm{Tr}(R^2(s)) = \chi^1(s) + \chi^2(s)$

ii) Taking again the matrix forms $R^1(s) = (r_{i_1 j_1}(s))$ and $R^2(s) = (r_{i_2 j_2}(s))$, we have

$$\chi^1(s) = \sum_{i_1} r_{i_1 i_1}(s) \qquad\qquad \chi^2(s) = \sum_{i_2} r_{i_2 i_2}(s). \tag{1.2.3}$$

Using the notation of equation (1.1.6), we find

$$\psi(s) = \sum_{i_1, i_2} r_{i_1 i_1}(s) r_{i_2 i_2}(s) = \chi^1(s) \cdot \chi^2(s).$$

∎

1.3 Schur's lemma

Characters are a powerful tool for determining irreducible representations, this follows from the following important proposition.

Proposition 1.3.1 (Schur's lemma)**.** *Let (ρ^1, V_1) and (ρ^2, V_2) be two representations of a group G, and let $f\colon V_1 \to V_2$ be a linear map such that $\rho^2_s \circ f = f \circ \rho^1_s$, $\forall s \in G$. Then*

i) If ρ^1 and ρ^2 are not isomorphic, then $f = 0$.

ii) If $V_1 = V_2$ and $\rho^1 = \rho^2$, then f is a homothety. In other words, f is a scalar multiple of the identity transformation.

Proof:

i) The case $f = 0$ is trivial, so let us suppose $f \neq 0$. We need to show that f is an isomorphism. Let W_1 be the kernel of f. Then, for $u \in W_1$ and $s \in G$, we have $(f\rho^1_s)(u) = (\rho^2_s f)(u) = 0 \Rightarrow \rho^1_s(u) \in W_1$. Thus W_1 is stable under G. Since (ρ^1, V_1) is irreducible, we either have $W_1 = 0$ or $W_1 = V_1$. Since the latter would imply that $f = 0$, we conclude that $W_1 = 0$, i.e. that f is injective. Now let W_2 be the image of f. Then, for $v \in W_2$ and $s \in G$, there exists $u \in V_1$ such that $f(u) = v$. For such u, we have $\rho^2_s(v) = (\rho^2_s f)(u) = (f\rho^1_s)(u) \in W_2$. Thus W_2 is stable under G, and because V_2 is irreducible and $f \neq 0$, we conclude that $W_2 = V_2$, i.e. that f is surjective, and therefore an isomorphism.

ii) Now we suppose that $V_1 = V_2$ and $\rho^1 = \rho^2$. Because the scalar field is the complex numbers, f has at least one eigenvalue λ. Therefore we can define $f' = f - \lambda I$, where I is the identity map. Then we have that $\ker f' \neq 0$ while $\rho^2_s \circ f = f \circ \rho^1_s$. By the same reasoning as in part (i), we conclude that we must have $f' = 0$, and therefore $f = \lambda$.

∎

Definition 1.3.2. *Note that in case (ii), $\rho = \rho^1 = \rho^2$ and f is an* intertwining operator*.*

Corollary 1.3.3. *Let G be a finite group of order g and (ρ^1, V_1), (ρ^2, V_2) be irreducible representations of G. Given a linear map $h\colon V_1 \to V_2$, define $h^0 = \frac{1}{g}\sum_{t\in G}(\rho^2_t)^{-1} h \rho^1_t$. Then*

i) If ρ^1 and ρ^2 are not isomorphic, then $h^0 = 0$.

ii) If $V_1 = V_2$ and $\rho^1 = \rho^2$, then h^0 is a homothety of ratio $\frac{1}{n}Tr(h)$, where $n = \dim V_1$.

Proof: Note that h^0 is a linear map from V_1 to V_2 and that, for all $s \in G$,

$$\rho_s^2 \circ h^0 = \frac{1}{g} \sum_{t \in G} \rho_{st^{-1}}^2 h \rho_t^1 = \frac{1}{g} \sum_{u \in G} \rho_{u^{-1}}^2 h \rho_{us}^1 = h^0 \circ \rho_s^1 \tag{1.3.1}$$

where we have set $u = ts^{-1}$. Applying Schur's lemma in case (i) yields $h^0 = 0$, and in case (ii), it yields $h^0 = \lambda$. Moreover, in this case

$$n\lambda = \mathrm{Tr}(h^0) = \frac{1}{g} \sum_{t \in G} \mathrm{Tr}(\rho_t^{-1} \rho_t h) = \mathrm{Tr}(h),$$

and we get $\lambda = \frac{1}{n} Tr(h)$. ∎

We now explicit Corollary 1.3.3, assuming that ρ^1 and ρ^2 are given in matrix form by:

$$\rho_t^1 = (r_{ij}^1(t))_{1 \leqslant i,j \leqslant n} \quad \text{and} \quad \rho_t^2 = (r_{ij}^2(t))_{1 \leqslant i,j \leqslant m} \tag{1.3.2}$$

where $n = \dim V_1$ and $m = \dim V_2$. Then if the linear maps h and h^0 from V_1 to V_2 are similarly given by the $(m \times n)$-matrices

$$h = (h_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} \quad \text{and} \quad h^0 = (h_{ij}^0),$$

we have

$$\begin{aligned} h_{ij}^0 &= \frac{1}{g} \sum_{t \in G} \begin{pmatrix} r_{i1}^2(t^{-1}) & \dots & r_{im}^2(t^{-1}) \end{pmatrix} \begin{pmatrix} h_{11} & \dots & h_{1n} \\ \vdots & \ddots & \vdots \\ h_{m1} & \dots & h_{mn} \end{pmatrix} \begin{pmatrix} r_{1j}^1(t) & & \\ \vdots & & \\ r_{mj}^1(t) & & \end{pmatrix} \\ &= \frac{1}{g} \sum_{t \in G} \sum_{k=1}^{m} \sum_{l=1}^{n} r_{ik}^2(t^{-1}) h_{kl}\, r_{lj}^1(t) \end{aligned} \tag{1.3.3}$$

We have the following corollary.

Corollary 1.3.4. *Keeping the notation of Corollary 1.3.3 and of* (1.3.2)*, we have:*

i) If ρ^1 and ρ^2 are not isomorphic, then

$$\sum_{t \in G} r_{ik}^2(t^{-1}) r_{lj}^1(t) = 0 \quad \forall i, j, k, l \tag{1.3.4}$$

ii) If $V_1 = V_2$, $n = \dim V$ and $\rho^1 = \rho^2$, then

$$\frac{1}{g} \sum_{t \in G} r_{ik}^2(t^{-1}) r_{lj}^1(t) = \frac{1}{n} \delta_{ij} \delta_{kl} \tag{1.3.5}$$

Proof:

1. Choose i, j such that $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$. We define the linear map $h\colon V_1 \to V_2$ whose matrix representation is given by $h_{kl} = \delta_{ki}\delta_{lj}$. Then by Corollary 1.3.3, we have $h = 0$, and therefore in particular
$$0 = h^0_{ij} = \frac{1}{g}\sum_{t\in G}\sum_{k=1}^{m}\sum_{l=1}^{n} r^2_{ik}(t^{-1})\delta_{kl}\, r^1_{lj}(t)$$
$$0 = \sum_{t\in G} r^2_{ik}(t^{-1}) r^1_{lj}(t)$$

2. By Corollary 1.3.3,
$$h^0_{ij} = \delta_{ij}\lambda = \delta_{ij}\frac{1}{n}\operatorname{Tr}(h) = \delta_{ki}\delta_{lj}\frac{1}{n}\sum_{k=1}^{n} h_{kk}$$
where δ_{ij} is the Kronecker delta, equal to 1 if $i = j$, and 0 otherwise. Substituting this into the left hand side of eq. (1.3.3) yields the result.

∎

1.4 Orthogonality relations of characters

If G is a finite group, let $\mathscr{F}(G)$ denote the vector space of functions $f\colon G \to \mathbb{C}$. There is a natural isomorphism ϕ between $\mathscr{F}(G)$ and $\mathbb{C}^g$, up to a choice of ordering of $G = \{t_1, \ldots, t_g\}$, given by
$$\begin{array}{rccc} \sigma\colon & \mathscr{F}(G) & \to & \mathbb{C}^g \\ & f & \mapsto & \begin{pmatrix} f(t_1) \\ \vdots \\ f(t_g) \end{pmatrix}. \end{array} \tag{1.4.1}$$

Now the standard inner product of two vectors u and v in $\mathbb{C}^g$ is $(u|v) = \sum_{1\leqslant i\leqslant g} u_i\overline{v_i}$. The corresponding function on $\mathscr{F}(G)\times\mathscr{F}(G)$ would be
$$\begin{array}{ccc} \mathscr{F}(G)\times\mathscr{F}(G) & \to & \mathbb{C} \\ (f,h) & \mapsto & \sum_{t\in G} f(t)\overline{h(t)} \end{array}. \tag{1.4.2}$$

One easily verifies that (1.4.2) is indeed conjugate symmetric, linear in the left slot and positive definite. Thus we can define an inner product on $\mathscr{F}(G)$:

Definition 1.4.1. *Let G be a finite group of order g, and ϕ and ψ be two functions from G to $\mathbb{C}$. We define $(\phi|\psi)$ as*

$$(\phi \mid \psi) = \frac{1}{g} \sum_{t \in G} \phi(t) \overline{\psi(t)}.$$

Of particular interest is the inner product of characters. Indeed, if χ and χ' are the characters of two irreducible representations ρ and ρ', we have

$$\begin{aligned}
(\chi \mid \chi') &= \frac{1}{g} \sum_{t \in G} \chi(t) \overline{\chi'(t)} \\
&= \frac{1}{g} \sum_{t \in G} \chi(t) \chi'(t^{-1}) \quad \text{(Propositon 1.2.3, part (ii))} \\
&= \frac{1}{g} \sum_{t \in G} \chi(t^{-1}) \chi'(t) \\
&= \frac{1}{g} \sum_{t \in G} \left(\sum_i r_{ii}(t^{-1}) \right) \left(\sum_j r'_{jj}(t) \right) \\
&= \sum_{i,j} \frac{1}{g} \sum_{t \in G} r_{ii}(t^{-1}) r'_{jj}(t)
\end{aligned}$$

where we have used the matrix representation (1.2.3) of the character function.

Using Corollary 1.3.4, we then obtain the following :

Theorem 1.4.2.

i) If χ and χ' are the characters of irreducible, non-isomorphic representations, then $(\chi \mid \chi') = 0$.

ii) If χ and χ' are the characters of irreducible, isomorphic representations, then $(\chi \mid \chi') = 1$.

Proof: Part (i) is immediate. For part (ii), simply note that in this case the corollary yields $(\chi \mid \chi') = \sum_{i,j} \frac{1}{n} \delta_{ij} = 1$, where n is the dimension of the vector space. ∎

Thus we have an orthogonality relation convenient for determining whether two representations are isomorphic.

Theorem 1.4.3. *Let ϕ be the character of a representation of G on V, and suppose $V = W_1 \oplus \cdots \oplus W_k$, the W_i being irreducible representations. If W is an irreducible representation of V with character χ, then the number of W_i's isomorphic to W is equal to $(\phi \mid \chi)$.*

Proof: From Proposition 1.2.5, we know that $\phi = \sum_{1 \leqslant i \leqslant k} \chi_i$, where χ_i is the character of the representation W_i. Then,

$$(\chi \mid \phi) = \big(\chi \mid \sum_{1 \leqslant i \leqslant k} \chi_i\big) = \sum_{1 \leqslant i \leqslant k} (\chi \mid \chi_i)$$

and since $(\chi \mid \chi_i)$ is 1 if W_i is isomorphic to W, and 0 otherwise, we have the result. ∎

Corollary 1.4.4. *Given a decomposition $V = W_1 \oplus \cdots \oplus W_k$ as in Theorem 1.4.3, the number of W_i isomorphic to W does not depend on the chosen decomposition.*

Proof: This follows from the fact that $(\phi \mid \chi)$ is independent of the decomposition. ∎

We will refer to this number as *the number of W contained in V*.

Corollary 1.4.5. *Let ρ and ρ' be two representations with characters χ and χ' respectively. If $\chi = \chi'$, then ρ is isomorphic to ρ'.*

Proof: This follows from the previous corollary, since any irreducible representation is contained in both the same number of times. ∎

Theorem 1.4.6. *Let ϕ be the character of a representation V. Then $(\phi \mid \phi)$ is a positive integer and $(\phi \mid \phi) = 1$ if and only if V is irreducible.*

Proof: We know from Theorem 1.1.17 that there exist irreducible representations W_i such that $V = W_1 \oplus \cdots \oplus W_m$ for some m. If the character of W_i is χ_i, then

$$\begin{aligned}(\phi \mid \phi) &= \sum_{k,l=1}^{m} (\chi_k, \chi_l) \\ &\geqslant \sum_{k=1}^{m} (\chi_k \mid \chi_k) = m\end{aligned}$$

Thus $(\phi \mid \phi)$ is a positive integer, and can equal to 1 only if $m = 1$, i.e. if V is irreducible. The fact that it does equal one in this case is the statement of Theorem 1.4.2. ∎

Example 1.4.7. Recall the representation $(\rho^h, \mathbb{C}^3)$ of D_n for n even given in Example 1.2.2, where h is an integer strictly between 0 and $n/2$. Its character was:

$$\chi(\rho^h_{r^k}) = (-1)^k + e^{i2\pi hk/n} + e^{-i2\pi hk/n}$$
$$\chi(\rho^h_{sr^k}) = (-1)^k$$

There are exactly n rotations and n symmetries in D_n, therefore we can verify whether ρ^h is irreducible by computing $(\chi|\chi)$ (recall that ρ^h was defined for n even):

$$\begin{aligned}
(\chi|\chi) &= \frac{1}{2n}\left[\sum_{k=0}^{n-1}((-1)^k + e^{i2\pi hk/n} + e^{-i2\pi hk/n})((-1)^k + e^{-i2\pi hk/n} + e^{i2\pi hk/n}) + \sum_{k=0}^{n-1}(-1)^{2k}\right] \\
&= \frac{1}{2n}\sum_{k=0}^{n-1}\left[(-1)^{2k} + 2 + 2(-1)^k(e^{i2\pi hk/n} + e^{-i2\pi hk/n}) + e^{i4\pi hk/n} + e^{-i4\pi hk/n} + 1\right] \\
&= 2 + \frac{1}{n}\sum_{k=0}^{n-1}(-e^{i2\pi h/n})^k + \frac{1}{n}\sum_{k=0}^{n-1}(-e^{-i2\pi h/n})^k + \frac{1}{2n}\sum_{k=0}^{n-1}(e^{i4\pi h/n})^k + \frac{1}{2n}\sum_{k=0}^{n-1}(e^{-i4\pi h/n})^k \\
&= 2 + \frac{1-(-1)^n}{n(1+e^{i2\pi h/n})} + \frac{1-(-1)^n}{n(1+e^{-i2\pi h/n})} + \frac{1-(1)^n}{2n(1-e^{i4\pi h/n})} + \frac{1-(1)^n}{2n(1-e^{-i4\pi h/n})} \\
&= 2
\end{aligned}$$

Since the $(\chi|\chi) \neq 1$, we see that in fact ρ^h is reducible.

In order to decompose ρ^h in irreducible subrepresentations, we must find subspaces of $\mathbb{C}^3$ which are invariant under ρ^h. In other words, we are looking for eigenspaces shared by every matrix $\rho^h_{r^k}$ and $\rho^h_{sr^k}$. For the matrices $\rho^h_{r^k}$, we find the following eigenvectors:

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{pmatrix}.$$

Writing the representation in this basis, we find the more pleasant forms

$$\rho^h_{r^k} = \begin{pmatrix} (-1)^k & 0 & 0 \\ 0 & w^{hk} & 0 \\ 0 & 0 & w^{-hk} \end{pmatrix} \qquad \rho^h_{sr^k} = \begin{pmatrix} (-1)^k & 0 & 0 \\ 0 & 0 & w^{-hk} \\ 0 & w^{hk} & 0 \end{pmatrix}$$

which easily decompose into two representations $(\rho^1, \mathbb{C})$ and $(\rho^{2,h}, \mathbb{C}^2)$:

$$\rho^1_{r^k} = \rho^1_{sr^k} = (-1)^k$$
$$\rho^{2,h}_{r^k} = \begin{pmatrix} w^{hk} & 0 \\ 0 & w^{-hk} \end{pmatrix} \qquad \rho^{2,h}_{sr^k} = \begin{pmatrix} 0 & w^{-hk} \\ w^{hk} & 0 \end{pmatrix}.$$

One can compute the scalar product of the characters of these representations, verifying that they are both now irreducible.

1.5 The Regular Representation

Definition 1.5.1. *Let G be a finite group of order g and V a vector space of dimension g. Given a basis of V, we can index this basis by the elements of G, allowing us to write it as $\{e_s\}_{s\in G}$. We then define the left* regular representation *(R, V) as the map such that $R_s(e_t) = e_{st}$ for all $s, t \in G$.*

An example of an explicit regular representation was given in Example 1.1.11.

Remark 1.5.2. Note that using the isomorphism (1.4.1), we can define a representation on the space of finite functions $\mathscr{F}(G)$:

$$\begin{array}{rccc} \lambda\colon & G & \to & \mathrm{GL}(\mathscr{F}(G)) \\ & s & \mapsto & \sigma^{-1} \circ R_s \circ \sigma. \end{array}$$

Thus if we have $f \in \mathscr{F}(G)$, then for all u in G:

$$\begin{aligned} (\lambda_s f)(u) &= \sigma^{-1} \circ R_s \left(\sum_t f(t) e_t \right)(u) \\ &= \sigma^{-1} \left(\sum_t f(s^{-1}t) e_t \right)(u) \\ &= f(s^{-1}u). \end{aligned}$$

In other words, on the space of functions, the regular representation amounts to a left translation.

Proposition 1.5.3. *The character r_G of the regular representation of G is given by*

$$r_G(s) = \begin{cases} g & \text{if } s = 1 \\ 0 & \text{otherwise} \end{cases}$$

Proof: If $s \neq 1$, then $st \neq t \quad \forall t \in G$. Therefore all the diagonal terms of R_s are zero, and $\mathrm{Tr}(R_s) = 0$. If $s = 1$, $r_G(1) = \mathrm{Tr}(1) = \dim V = g$. ∎

For the rest of this section, we will denote $W_1, \ldots, W_h$ the distinct irreducible subrepresentations of the regular representation; $\chi_1, \ldots, \chi_h$ their characters, and $n_1, \ldots, n_h$ their degrees.

Corollary 1.5.4. *The irreducible representations W_i appear in the regular representation with a multiplicity equal to their degree n_i.*

Proof: By Propositon 1.4.3, this multiplicity will be equal to

$$(r_G \mid \chi_i) = \frac{1}{g}\sum_{s\in G} r_G(s)\overline{\chi_i(s)} = \frac{1}{g}\sum_{s\in G} r_G(s^{-1})\chi_i(s) = \chi_i(1) = n_i$$

∎

Since any representation of G is at least of degree one, this effectively implies that any irreducible representation of G also appears in its regular representation. Therefore speaking of the irreducible representations of G or of the subrepresentations of its regular representation is equivalent.

Knowing the multiplicity of each W_i, we can write r_G as a sum of the irreducible characters:

$$r_G(s) = \sum_{i=1}^{h} n_i\chi_i(s)$$

Applying this to the cases $s = 1$ and $s \neq 1$ and comparing with Proposition 1.5.3, we obtain the following.

Corollary 1.5.5.

i) $\sum_{i=1}^{h} n_i^2 = g$

ii) $\sum_{i=1}^{h} n_i\chi_i(s) = 0$ *for* $s \neq 1$

Remark 1.5.6. Given a set of nonisomorphic irreducible representations W_i of degree n_i of G, this allows us to check immediately whether it is all the irreducible representations of G. Indeed, that will be the case if and only if $\sum_{i=1}^{h} n_i^2 = g$.

Example 1.5.7. We can use the above remark to find all nonisomorphic irreducible representations of D_n, in which r denotes a rotation by $2\pi/n$ and s a symmetry. Any one-dimensional representation must send r and s to $+1$ or -1, and can do so independently. Therefore, there are four one-dimensional representations, whose characters ψ_i are given by

	r^k	sr^k
ψ_1	1	1
ψ_2	1	-1
ψ_3	$(-1)^k$	$(-1)^k$
ψ_4	$(-1)^k$	$(-1)^{k+1}$

(If n were odd, there would be only two one-dimensional representations, corresponding to the first two rows)

We have also already determined one family of irreducible representations ρ^h of degree 2 in Example 1.4.7:

$$\rho^h_{r^k} = \begin{pmatrix} w^{hk} & 0 \\ 0 & w^{-hk} \end{pmatrix} \qquad \rho^h_{sr^k} = \begin{pmatrix} 0 & w^{-hk} \\ w^{hk} & 0 \end{pmatrix}.$$

As ρ^h is isomorphic to ρ^{n-h} (just take the map which swaps coordinates as the isomorphism), the only possible values of h for nonisomorphic representations are $0 \leqslant h \leqslant \frac{n}{2}$. If $h = 0$ or $h = \frac{n}{2}$, the representation is reducible (the characters are $\psi_1 + \psi_2$ and $\psi_3 + \psi_4$, respectively). That leaves us with $\frac{n}{2} - 1$ distinct representations. The sum of the squares of their degrees is $4 \times 1^2 + (\frac{n}{2} - 1) \times 2^2 = 2n$, which is the order of D_n. Therefore these are all of the irreducible subrepresentations of D_n.

We conclude this chapter by showing that the character functions form a basis of the space of class functions. We will use this to show that the number of irreducible representations is given by the number of conjugacy classes of G.

Recall that two elements $t, t' \in G$ are *conjugate* if there exists $s \in G$ such that $t' = sts^{-1}$. Recall also that this defines an equivalence relation which partitions G into *conjugacy classes* and that, as per Remark 1.2.4, functions which are constant over conjugacy classes are called *class functions*.

We will denote the set of all class functions on G by H.

Proposition 1.5.8. *Let ρ be a representation of G on V and $f \in H$. We define the linear map ρ_f as*

$$\rho_f = \sum_{t \in G} f(t)\rho_t.$$

If V is irreducible of degree n and character χ, then ρ_f is a homothety of ratio λ, where

$$\lambda = \frac{g}{n}(f \mid \overline{\chi}).$$

Proof: In order to apply Schur's lemma, we need to show that $\rho_s \rho_f = \rho_f \rho_s$. So we compute

$$\rho_s \rho_f \rho_{s^{-1}} = \sum_{t \in G} f(t)\rho_s \rho_t \rho_{s^{-1}} = \sum_{t \in G} f(t)\rho_{sts^{-1}},$$

letting $u = sts^{-1}$,

$$\begin{aligned} &= \sum_{u \in G} f(s^{-1}us)\rho_u = \sum_{u \in G} f(u)\rho_u \\ &= \rho_f. \end{aligned}$$

Having verified the condition, we conclude by Schur's lemma that ρ_f is a homothety of ratio λ. Then,

$$\begin{aligned} \operatorname{Tr}(\lambda) = \operatorname{Tr}(\rho_f) &= \operatorname{Tr}\Big(\sum_{t\in G} f(t)\rho_t\Big) \\ n\lambda = \sum_{t\in G} f(t)\operatorname{Tr}(\rho_t) &= \sum_{t\in G} f(t)\chi(t) \\ &= g(f \mid \overline{\chi_\rho}) \end{aligned}$$

which shows that $\lambda = \frac{g}{n}(f \mid \overline{\chi})$. ∎

Theorem 1.5.9. *Let G be a group. The characters $\chi_1, \ldots, \chi_h$ of its irreducible representations form an orthonormal basis of H.*

Proof: We have already shown that the χ_i are orthonormal (Th. 1.4.2), therefore it suffices to prove that they generate H. For that, we will show that if a class function f is orthogonal to $\chi_1, \ldots, \chi_h$, then $f = 0$.

Let f be such a function and define $\rho_f = \sum_{t\in G} f(t)\rho_t$. We decompose V into irreducible representations; by Proposition 1.5.8, $\rho_f = 0$ on each of these, and therefore is also zero on V.

Now we consider the action of ρ_f on the base vector e_1 of the regular representation. We have

$$\begin{aligned} \rho_f(e_1) &= \sum_{t\in G} f(t)\rho_t(e_1) \\ 0 &= \sum_{t\in G} f(t)e_t \end{aligned}$$

which shows that $f(t) = 0$, $\forall t \in G$. ∎

Theorem 1.5.10. *The number of irreducible representations of G (up to isomorphism) is equal to the number of conjugacy classes of G.*

Proof: Let $C_1, \ldots, C_k$ be the conjugacy classes of G. Since class functions are constant over classes, any $f \in H$ can be expressed as $f = \sum_{i=1}^{k} \lambda_i \mathbb{1}_{C_i}$, where $\mathbb{1}$ is the characteristic function of C_i. Since the functions of H are exactly determined by the choice of the λ_i, $1 \leqslant i \leqslant k$, and each λ_i can be independently varied, we must have $\dim H = k$. Yet we also have, by Th. 1.5.9, that $\dim H = h$, where h is the number of non-isomorphic irreducible representations of G. Hence $h = k$. ∎

Chapter 2

Extension to Compact Groups

The objective of this chapter is to extend the previous results to the case of a compact group G. This will require the construction of the *Haar measure* on G. It will then be possible to rewrite most of the results of the previous section by replacing the sums by integrals. In particular, if G is finite, its Haar measure reduces to the normalized counting measure, and integrating over G reduces to summing over G, as before.

We will begin by showing the construction of the integral using measure theory. Although this is not the subject of this paper, as we will use the results, we include it for completeness. However we will only state the results; for proofs, one can see [3]. The same will be true for some results from functional analysis, which go beyond the scope of this presentation.

2.1 Measure and integration

Definition 2.1.1. *Let X be a set. A* σ-algebra *is a subset $\mathscr{A} \subset \mathscr{P}(X)$ such that*

i) $X \in \mathscr{A}$.

ii) If $A \in \mathscr{A}$, then $A^c \in \mathscr{A}$, where A^c denotes the complement of A in X.

iii) If $A = \bigcup_{n \geqslant 1} A_n$, and each $A_n \in \mathscr{A}$, then $A \in \mathscr{A}$.

Definition 2.1.2. *If X is a topological space, the* Borel σ-algebra *over X, noted $\mathscr{B}(X)$, is the σ-algebra generated by the open sets of X. A set $A \in \mathscr{B}(X)$ is called a* Borel set.

Definition 2.1.3. *Let X be a set and $\mathscr{A}$ be a σ-algebra on X. A* measure μ *on $(X, \mathscr{A})$ is a map $\mu \colon \mathscr{A} \to [0, \infty]$ such that*

i) $\mu(\varnothing) = 0$.

ii) If $\{A_n\}_{n\geqslant 1}$ *is a sequence of disjoints sets of* $\mathscr{A}$*, then*

$$\mu\big(\cup_{n\geqslant 1}A_n\big) = \sum_{n\geqslant 1}\mu(A_n)$$

.

Given a set X, a σ-algebra $\mathscr{A}$ over X and a measure μ, we call $(X, \mathscr{A}, \mu)$ a *measured space*.

Definition 2.1.4. *If* X *is a topological space, a measure defined over the Borel* σ*-algebra of* X *is called a* Borel measure*.*

Definition 2.1.5. *Let* $(X, \mathscr{A}, \mu)$ *be a measured space where* $\mathscr{B}(X) \subset \mathscr{A}$*. We say that the measure* μ *is* regular *if*

i) $\mu(K) < \infty$ *for all* $K \in \mathscr{A}$ *compact,*

ii) $\mu(A) = \sup\{\mu(K) \mid K \subset A, K \text{ compact}\}$ *for all* $A \in \mathscr{A}$*,*

iii) $\mu(A) = \inf\{\mu(U) \mid A \subset U, U \text{ open}\}$ *for all* $A \in \mathscr{A}$*.*

In particular, since we will require the Haar measure to be regular, it will be possible to express the measure of a measurable set A as the supremum of the measures of the compact sets contained in A.

Definition 2.1.6. *Let* $(X, \mathscr{A}, \mu)$ *be a measured space and* Y *be a topological space. A function* $f\colon X \to Y$ *is* measurable *if, for any open* $U \subset Y$*, we have* $f^{-1}(U) \in \mathscr{A}$*.*

Definition 2.1.7. *A* simple function *is a measurable function* $s\colon X \to [0, \infty]$ *which takes only a finite number of values.*

In other words, s is a simple function if and only if there exists a collection of $\{(A_i, \alpha_i) \mid A_i \in \mathscr{A}, \alpha_i \in [0, \infty]\}$, such that the A_i are two by two disjoint and $s = \sum_i \alpha_i \mathbb{1}_{A_i}$, where $\mathbb{1}_A$ is the characteristic function, equal to 1 on A and 0 elsewhere.

Definition 2.1.8. *Let* $(X, \mathscr{A}, \mu)$ *be a measured space and* s *be a simple function. We have* $s = \sum_i \alpha_i \mathbb{1}_{A_i}$ *for some* α_i*,* A_i*. Then we define the integral of* s *as*

$$\int s \,\mathrm{d}\mu \overset{def}{=} \sum_i \alpha_i \mu(A_i),$$

where we have adopted the convention that $0 \cdot \infty \overset{def}{=} 0$*.*

Now that we have defined the integral over simple functions, it is simply a matter of successively extending this definition to complex functions, as we do in the three following definitions.

Definition 2.1.9. *Let $(X, \mathscr{A}, \mu)$ be a measured space, f be a positive measurable function and $\mathscr{S}_+$ be the set of positive simple functions. We define the integral of f as*

$$\int f \,\mathrm{d}\mu \stackrel{def}{=} \sup\Big\{\int s \,\mathrm{d}\mu, s \in \mathscr{S}_+, s \leqslant f\Big\}$$

Definition 2.1.10. *Let $(X, \mathscr{A}, \mu)$ be a measured space, $f\colon X \to \mathbb{R}$ be a measurable function. Define the positive measurable functions f^+ and f^- by*

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) \geqslant 0 \\ 0 & \text{otherwise} \end{cases} \qquad f^-(x) = \begin{cases} -f(x) & \text{if } f(x) < 0 \\ 0 & \text{otherwise} \end{cases}$$

for $x \in X$. Then if $\int |f| \,\mathrm{d}\mu < \infty$, we define the integral of f as

$$\int f \,\mathrm{d}\mu \stackrel{def}{=} \int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu.$$

Definition 2.1.11. *Let $(X, \mathscr{A}, \mu)$ be a measured space, $f\colon X \to \mathbb{C}$ be a measurable function. We write $f = u + iv$, where $u, v\colon X \to \mathbb{R}$. Then if $\int |f| \,\mathrm{d}\mu < \infty$, we define the integral of f as*

$$\int f \,\mathrm{d}\mu \stackrel{def}{=} \int u \,\mathrm{d}\mu + i \int v \,\mathrm{d}\mu.$$

Example 2.1.12. Suppose we are given a measured space $(X, \mathscr{B}(X), \mu)$ and a continuous function $f\colon X \to \mathbb{R}^+$. Then f is Borel measurable and the function

$$\begin{array}{rccc} \nu_f\colon & \mathscr{B}(X) & \to & [0, \infty] \\ & A & \mapsto & \int \mathbb{1}_A f \,\mathrm{d}\mu \end{array}$$

is in fact a Borel measure.

Definition 2.1.13. *Let $(X, \mathscr{A}, \mu)$ be a measured space, $f\colon X \to \mathbb{C}$ be a measurable function. We define the* essential supremum *of f as*

$$\|f\|_\infty = \inf\{\alpha \in [0, \infty], \mu(\{x \in X \mid \alpha < |f(x)|\}) = 0\}$$

Remark 2.1.14. Note that $\|\cdot\|_\infty$ as defined above is only a semi-norm, because we do not have $\|f\|_\infty = 0 \Leftrightarrow f = 0$. We can however make it a norm by taking an appropriate quotient of the space of measurable functions, in which case, $\|\cdot\|_\infty$ is called the *sup norm*. For details on how this is done, see for example [3].

We conclude this section by stating a standard result, the Riesz representation theorem.

Definition 2.1.15. *Let f be a real valued function on a topological space X. The* support *of f, denoted* $\operatorname{supp}(f)$*, is defined as the closure of the set $\{x \in X \mid f(x) \neq 0\}$.*

Remark 2.1.16. Let X be a topological space and let $C(X)$ denote the set of continuous real (resp. complex) valued functions on X. Then $C(X)$ is a real (resp. complex) vector space. Furthermore, $C_c(X) \stackrel{\text{def}}{=} \{f \in C(X); \text{supp}(f) \text{ compact}\}$ is a subspace of $C(X)$.

For a function $f\colon X \to \mathbb{C}$, let f^* denote the function given by $f^*(x) = \overline{f(x)}$, $x \in X$. Note that if $f \in C(X)$, then $f^* \in C(X)$. As the product of two continuous functions (resp. with compact support) is a continuous function (resp. with compact support), then $C(X)$ is a commutative involutive algebra.

Definition 2.1.17. *Let X be a topological space. A linear form $I\colon C_c(X) \to \mathbb{R}$ is said to be* positive *if for every non-negative $f \in C_c(X)$, we have $I(f) \geqslant 0$.*

Theorem 2.1.18 (Riesz representation theorem)**.** *Let X be a locally compact Hausdorff space, and I be a positive linear form on $C_c(X)$. Then there is a unique regular Borel measure μ on X such that*

$$I(f) = \int f \, \mathrm{d}\mu$$

for every f in $C_c(X)$.

2.2 Construction of the Haar measure

The Lebesgue measure λ on the real line $\mathbb{R}$ is a regular Borel measure and is invariant by translation (i.e., for $A \subset \mathbb{R}$ and $x \in \mathbb{R}$, $\lambda(A + x) = \lambda(A)$, where $A + x = \{y \in \mathbb{R} | y - x \in A\}$). In fact, this property characterizes the Lebesgue measure among the class of Borel measures μ on $\mathbb{R}$ such that $\mu([0,1]) = 1$ and $\mu(B) < \infty$ for $B \in \mathscr{B}(\mathbb{R})$, B bounded.

If G is a locally compact group, then a non-zero regular Borel measure μ on G that is invariant under left translation (resp. right-translation) (i.e., $\mu(xA) = \mu(A)$, for all $x \in G, A \in \mathscr{B}(G)$) is called a *left* (resp. *right*) *Haar measure on G*. In this section, we will prove the existence and uniqueness of the Haar measure for a compact group G. The presentation will follow the one done by Robert [4]; an alternate construction of the Haar mesure using ergodic theory can also be found in [3].

Definition 2.2.1. *A group G is a* topological group *if it is endowed with a topology such that the group operations*

$$\begin{array}{ccc} G \times G & \to & G \\ (s,t) & \mapsto & st \end{array} \qquad \textit{and} \qquad \begin{array}{ccc} G & \to & G \\ s & \mapsto & s^{-1} \end{array}$$

are continuous.

Definition 2.2.2. *Let G be a compact group. We define the* left translation $\lambda_s \colon C(G) \to C(G)$ *by*

$$\lambda_s(f)(t) = f(s^{-1}t),\ t \in G.$$

Similarly, the right translation $\rho_s \colon C(G) \to C(G)$ *is defined by*

$$\rho_s(f)(t) = f(ts),\ t \in G.$$

Remark 2.2.3. Note that if the group G is compact, then $C_c(G) = \mathrm{C}(G)$. If G is finite, then λ is in fact the left regular representation of $C(G)$, as per Remark 1.5.2.

We can now state the theorem:

Theorem 2.2.4. *Let G be a compact group. There exists a unique regular probability measure $\mu \colon \mathscr{B}(G) \to [0,1]$ such that*

i) $\displaystyle\int f(t)\,\mathrm{d}\mu(t) \geqslant 0 \quad$ *if* $f \geqslant 0$

ii) $\displaystyle\int f(s^{-1}t)\,\mathrm{d}\mu(t) = \int f(t)\,\mathrm{d}\mu(t) \qquad \forall s \in G, \forall f \in \mathrm{C}(G).$

Furthermore, μ is such that

$$\int f(ts)\,\mathrm{d}\mu(t) = \int f(t)\,\mathrm{d}\mu(t) \quad \forall s \in G, \forall f \in \mathrm{C}(G).$$

Definition 2.2.5. *For a given compact group G, the unique measure μ satisfying the conditions of Theorem 2.2.4 is called the* Haar measure *of G.*

What we will prove is that there exists a unique positive continuous linear form $I \colon C(G) \to \mathbb{C}$ such that for all $s \in G$,

i) $I(\mathbb{1}_G) = 1$,

ii) I is positive,

iii) $I(\lambda_s f) = I(f)$,

and that this linear form also verifies $I(\rho_s f) = I(f)$. Then by the Riesz representation theorem, we will conclude that there exists a unique regular Borel measure μ on G such that $I(f) = \int f\,\mathrm{d}\mu$. This μ clearly satisfies all the conditions of Theorem 2.2.4, and therefore constitutes the Haar measure of G.

The proof requires two results from functional analysis, Ascoli's theorem and Kakutani's fixed point theorem, which we now state; proofs of these theorems may be found in [5].

Definition 2.2.6. *Let X be a topological space, (Y, d_Y) be a metric space, and $\mathscr{F} = (f_i)_{i\in I}$ be a family of functions $f_i \colon X \to Y$. Then $\mathscr{F}$ is said to be* equicontinuous *if, for all $\epsilon > 0$ and all $x \in X$, there exists an open neighborhood $V(x)$ of x such that*

$$\forall i \in I, \forall y \in V(x), d_Y(f_i(x) - f_i(y)) < \epsilon$$

Definition 2.2.7. *A subset X of a vector space is* convex *if, for all $x, y \in X$, the segment*

$$\{tx + (1-t)y \mid 0 \leqslant t \leqslant 1\}$$

is contained in X.

Remark 2.2.8. Note that if when restricted to the continuous functions $C(X, E)$ (where X is a compact topological space and E is a Banach space), the definition of the sup norm (Definition 2.1.13) is equivalent to the the following:

$$\|f\|_\infty = \sup\{\|f(x)\|\,; x \in X\}.$$

In this case, because the functions are all continuous, this defines a proper norm. Consequentely, such a $C(X, E)$ is a Banach space when endowed with the sup norm.

Theorem 2.2.9 (Ascoli's theorem)**.** *Let X be a compact topological space and E be a Banach space. Let $\mathrm{C}(X; E)$ be the set of continuous fonctions from X to E, and endow $\mathrm{C}(X; E)$ with the sup norm. Let Φ be a subset of $\mathrm{C}(X; E)$. Then Φ is relatively compact if and only if*

i) Φ is equicontinuous .

ii) the sets $\{f(x) : f \in \Phi\}$ are relatively compact in E for all $x \in X$.

(Recall that a set A in a topological space X is *relatively compact* if its closure $\overline{A}$ is compact in X.)

Theorem 2.2.10 (Kakutani's fixed point theorem)**.** *Let π be a representation of a group G on a Banach space E. If there exists a convex compact subset K of E such that $\pi_s K \subset K$ for $s \in G$ and $\{\pi_s\}_{s\in G}$ is equicontinuous, then there exists $x \in K$ such that $\pi_s x = x$, $\forall s \in G$. (Such an x is said to be a fixed point of G in K.)*

We can now construct the linear form I, which will prove the theorem.

Proof of theorem 2.2.4:

Let f be a given function in $\mathrm{C}(G;G)$ and consider the following set

$$C_f = \Big\{ \sum_{i\in I} \alpha_i \lambda_{s_i} f : \alpha_i \in \mathbb{C}, s_i \in G, I \text{ finite}, \sum_{i\in I} \alpha_i = 1 \Big\}.$$

By definition, C_f is clearly λ-invariant, i.e. for any $g \in C_f$, $s \in G$, we have $\lambda_s(g) \in C_f$. It is also clear that for such a function g, $\sup_G |g(t)| \leqslant \sup_G |f(t)|$. And since G is compact and f is continuous over G, $\sup_G |f(t)| < \infty$. Therefore $C_f(t)$, as well as its closure $\overline{C_f(t)}$, are bounded in $\mathbb{C}$. Thus $\overline{C_f(t)}$ is compact on $\mathbb{C}$, for all $t \in \mathbb{C}$. Hence condition (ii) of Theorem 2.2.9 is satisfied.

Since G is compact, f is uniformly continuous. By definition, this means that for all $\varepsilon > 0$, there exists a neighborhood Ω_ε such that, for any $t, u \in G$,

$$t^{-1}u \in \Omega_\varepsilon \Rightarrow |f(t) - f(u)| < \varepsilon$$

Now take $g \in C_f$: $g = \sum_{i\in I} \alpha_i \lambda_{s_i} f$ for some $\alpha_i \in \mathbb{C}$, $s_i \in G$. Then, for any t, u such that $t^{-1}u \in \Omega_\varepsilon$ and any $i \in I$,

$$(s_i^{-1}t)^{-1} s_i^{-1} u = t^{-1} s_i s_i^{-1} u = t^{-1}u \in \Omega_\varepsilon \Rightarrow \left| f(s_i^{-1}t) - f(s_i^{-1}u) \right| < \varepsilon.$$

Hence

$$\begin{aligned} |g(t) - g(u)| &= \sum_i \alpha_i \left| f(s_i^{-1}t) - f(s_i^{-1}u) \right| \\ &< \sum_i \alpha_i \varepsilon = \varepsilon, \end{aligned}$$

which shows that C_f is uniformly equicontinuous. Hence condition (i) of Theorem 2.2.9 is satisfied.

Then by Ascoli's theorem, C_f is relatively compact in $\mathrm{C}(G)$. Hence $K_f \overset{\text{def}}{=} \overline{C}_f$ is compact.

We now wish to apply the fixed point theorem to show that, for all $s \in G$, there exists $g \in C(G)$ such that $\lambda_s g = g$. In order to do this, we need to verify that when $C(G)$ is endowed with the sup norm (Def. 2.1.13), $\{\lambda_s\}_{s\in G}$ is equicontinuous. Explicitly, this is expressed as

$$\forall_{\varepsilon>0}, \exists_{\delta>0}, \forall_{s\in G}, \|f - g\|_\infty < \delta \Rightarrow \|\lambda_s f - \lambda_s g\|_\infty < \varepsilon.$$

Since for all $s \in G$, $\lambda_s \colon C(G) \to C(G)$ is an isometry with respect to the sup norm, we conclude immediately that $\{\lambda_s\}_{s\in G}$ is indeed equicontinuous. Moreover, as C_f is λ-invariant, K_f is also λ-invariant.

So, by Kakutani's fixed point theorem, there exists $g \in K_f$ such that $\lambda_s g = g$ for all $s \in G$. Such a g must be constant, since it satisfies $g(st) = g(t)\ \forall s, t \in G$. We will denote its value by c.

We now need to show the uniqueness of c. Since $c \in K_f$, it can be approached by elements of C_f. Therefore, for all $\varepsilon > 0$, there exists a finite set of $\alpha_i > 0$, $s_i \in G$, $\sum \alpha_i = 1$ such that

$$\left| c - \sum_i \alpha_i f(s_i t) \right| < \varepsilon \qquad \forall t \in G \tag{2.2.1}$$

We can of course repeat the previous construction, replacing left translation by right translation. Such a process yields a new constant c' and a finite set of $\beta_j > 0$, $u_j \in G$, $\sum \beta_j = 1$ such that

$$\left| c' - \sum_j \beta_j f(t u_j) \right| < \varepsilon \qquad \forall t \in G \tag{2.2.2}$$

Multiplying (2.2.2) by α_i, evaluating it at $t = s_i$ and summing over all i, we obtain

$$\left| c' - \sum_{i,j} \alpha_i \beta_j f(s_i u_j) \right| < \varepsilon \tag{2.2.3}$$

Repeating the same process with equation 2.2.1 (multipling by β_j, evaluating at u_j, summing over j), we obtain

$$\left| c - \sum_{i,j} \alpha_i \beta_j f(s_i u_j) \right| < \varepsilon \tag{2.2.4}$$

Adding equations (2.2.3) and (2.2.4), we find that

$$2\varepsilon \geqslant \left| c - \sum_{i,j} \alpha_i \beta_j f(s_i u_j) \right| + \left| \sum_{i,j} \alpha_i \beta_j f(s_i u_j) - c' \right| \geqslant |c - c'|$$

Since ε is arbitrary, we conclude that $c = c'$.

Note that this proves the uniqueness of c, since if there were two constants c_1, c_2 obtained through the construction using left translation, we would have $c_1 = c' = c_2$. The same of course applies for right translation. Therefore, as c is unique, we can define

$$I(f) \stackrel{\text{def}}{=} c.$$

A number of results are immediate from the construction

i) $I(\mathbb{1}_G) = 1$. This follows from the fact that $K_{\mathbb{1}_G} = C_{\mathbb{1}_G} = \{\mathbb{1}_G\}$,

ii) $f \geqslant 0 \Rightarrow I(f) \geqslant 0$. (i.e. I is positive),

iii) $I(af) = aI(f)$ for all $a \in \mathbb{C}$, $f \in C(G)$,

iv) $I(\lambda_s(f)) = I(f) = I(\rho_s(f))$ for all $s \in G$, $f \in C(G)$. This follows from invariance of K_f under translation. (ρ denotes here the right translation).

Now all that is left to show is the additivity of I, which will complete linearity. Take $f, g \in \mathrm{C}(G)$ and $\varepsilon > 0$. As in (2.2.1), there exist α_i, s_i such that

$$\left| I(f) - \sum_i \alpha_i f(s_i t) \right| < \varepsilon. \tag{2.2.5}$$

We define

$$h(t) = \sum_i \alpha_i g(s_i t). \tag{2.2.6}$$

Clearly $h \in C_g$, and therefore $K_h \subset K_g$. By uniqueness of the constant in K_g, we must have $I(h) = I(g)$.

Again as in (2.2.1), there exist β_j, u_j such that

$$\left| I(h) - \sum_j \beta_j h(u_j t) \right| < \varepsilon.$$

Substituting (2.2.6), we have

$$\left| I(g) - \sum_{i,j} \alpha_i \beta_j g(s_i u_j t) \right| < \varepsilon.$$

If we multiply (2.2.5) by β_j, evaluate it at $u_j t$ and sum over the j, we obtain

$$\left| I(f) - \sum_{i,j} \alpha_i \beta_j f(s_i u_j t) \right| < \varepsilon.$$

Summing the two last equations, we find that

$$\left| I(f) + I(g) - \sum_{i,j} \alpha_i \beta_j (f+g)(s_i u_j t) \right| < 2\varepsilon$$

Therefore the constant $I(f) + I(g)$ is in K_{f+g}. Since the only constant on K_{f+g} is $I(f+g)$, we conclude that

$$I(f+g) = I(f) + I(g),$$

which completes the proof. ∎

Example 2.2.11. If G is a finite group of order g, it is easy to verify that the Haar measure on G is the normalized counting measure μ:

$$\begin{array}{rccc} \mu\colon & \mathscr{B}(G) & \to & [0,1] \\ & A & \mapsto & \frac{1}{g}\#A \end{array}.$$

In this case we have

$$\int_G f(t)\,\mathrm{d}\mu(t) = \frac{1}{g}\sum_{t\in G} f(t),$$

so integrating over G is exactly equivalent to summing over G, as we have done in Chapter 1.

2.3 Representations of compact groups

In this section we extend the definitions and results of representations of finite groups to compact groups.

Definition 2.3.1. *Let E be a Banach space and consider the following subsets of* $\mathrm{GL}(E)$*, where $\phi \in \mathrm{GL}(E)$, $\epsilon > 0$, $x_1, \ldots, x_n \in E$:*

$$V(\phi; x_1, \ldots, x_n; \epsilon) = \{\psi \in \mathrm{GL}(E); \|\phi(x_i) - \psi(x_i)\| < \epsilon; 1 \leqslant i \leqslant n\}.$$

Then the collection of all such subsets of $\mathrm{GL}(E)$ *is a topological base which generates the* topology of simple strong convergence.

Definition 2.3.2. *Let G be a compact group and E be a Banach space. A* continuous representation *of G on E is a continuous homomorphism*

$$\begin{array}{rccccccc} \pi\colon & G & \to & \mathrm{GL}(E) & & & \\ & s & \mapsto & \pi_s\colon & E & \to & E \\ & & & & v & \mapsto & \pi_s v \end{array}$$

where $\mathrm{GL}(E)$ *is endowed with the topology of simple strong convergence.*

The definitions for isomorphic representations (Def. 1.1.8), subrepresentations (Def. 1.1.9), irreducible representations (Def. 1.1.10) and unitary representations (Def. 1.1.15) are exactly the same as for finite groups. It is also true, as in Proposition 1.1.16, that any representation can be made unitary. In order to prove this, we first recall the notion of equivalence of norms and the Banach-Steinhaus theorem:

Recall 2.3.3.

(a) Two norms N_1 and N_2 on a vector space E are said to be *equivalent* if there exist positive real numbers a and b such that, for all $v \in E$,

$$aN_1(v) \leqslant N_2(v) \leqslant bN_1(v).$$

When that is the case, both N_1 and N_2 induce the same topology.

(b) If E is a Banach space and F a normed vector space, then let $B(E,F) = \{f : E \to F; f \text{ a continuous linear map}\}$. For $f \in B(E,F)$, then

$$\|f\| = \sup\{\|f(x)\|\,; x \leqslant 1\}$$

defines a norm on $B(E,F)$.

Theorem 2.3.4 (Banach-Steinhaus theorem)**.** *Let E be a Banach space and F be a normed vector space whose norm is denoted by $\|\cdot\|$. Let $\mathscr{F}$ be a pointwise bounded family of continuous linear applications from E to F. Then $\mathscr{F}$ is uniformly bounded, i.e. there exists a positive real constant $M < \infty$ such that*

$$\forall f \in \mathscr{F}, \|f\| < M$$

, or equivalently, that

$$\forall v \in E, \forall f \in \mathscr{F}, \|f(v)\| < M\,\|v\|$$

The proof of this theorem can be found in [5].

Proposition 2.3.5. *Let π be a representation of a compact group G on a Hilbert space H. There exists a positive definite hermitian form (i.e. an inner product) ϕ on H that is invariant under G and which preserves the topology on H.*

Proof: Consider the following sesquilinear form ψ:

$$\begin{array}{rccc} \psi\colon & H \times H & \to & \mathbb{C} \\ & (v,w) & \mapsto & \int_G (\pi_s v \mid \pi_s w)\,\mathrm{d}s \end{array}.$$

The first thing we note is that $s \mapsto \pi_s v$ and $(v,w) \mapsto (v \mid w)$ are continuous for all $v, w \in H$. Consequently, the map $s \mapsto (\pi_s v \mid \pi_s w)$ is also continuous, and therefore measurable, for all $v, w \in H$. ψ is also positive definite, since $(v \mid v) > 0$ for $v \in H\backslash\{0\}$.

To conclude that ψ is the desired sesquilinear form ϕ, we still need to show that it is invariant, and that the topology induced on H by ψ is the same as that generated

by $\|\cdot\| = (\cdot \mid \cdot)^{1/2}$. The invariance of ψ follows immediately from the invariance of the Haar measure with respect to left translation:

$$\begin{aligned}\psi(\pi_t v, \pi_t w) &= \int_G (\pi_{st} v \mid \pi_{st} w)\,\mathrm{d}s \\ &= \int_G (\pi_s v \mid \pi_s w)\,\mathrm{d}s \\ &= \psi(v, w)\end{aligned}$$

As per the above recall, the topologies generated by $\|\cdot\|$ and ψ will be the same if the two norms are equivalent. Now, for any given $v \in H$, the map $s \mapsto \|\pi_s v\|$ is a continuous function over a compact; therefore we have

$$\forall v \in H, \quad \sup_{s \in G}\{\|\pi_s v\|\} < \infty.$$

In other words, the family of functions $\{\pi_s\}_{s \in G}$ is point-wise bounded. Therefore, by the Banach-Steinhaus theorem, it is also uniformly bounded and there exists a positive real constant $M < \infty$ such that

$$\forall v \in H, \forall s \in G, \|\pi_s v\| < M \|v\|$$

Therefore, for any $v \in H$, $s \in G$, we have the following:

$$\begin{array}{llll} \|v\| = \|\pi_{s^{-1}}\pi_s v\| & \leqslant M\|\pi_s v\| & \leqslant M^2\|v\| \\ M^{-1}\|v\| & \leqslant \|\pi_s v\| & \leqslant M\|v\| \\ M^{-2}\|v\|^2 & \leqslant (\pi_s v \mid \pi_s v) & \leqslant M^2\|v\|^2 \\ M^{-2}\|v\|^2 \displaystyle\int_G \mathrm{d}s & \leqslant \displaystyle\int_G (\pi_s v \mid \pi_s v)\,\mathrm{d}s & \leqslant M^2\|v\|^2 \displaystyle\int_G \mathrm{d}s \\ M^{-2}\|v\|^2 & \leqslant \psi(v, v) & \leqslant M^2\|v\|^2 \end{array} \tag{2.3.1}$$

Equation (2.3.1) shows that $\psi(v, v)$ is equivalent to $\|v\|^2$. We also have from (2.3.1) that

$$\psi(v, v) = 0 \Rightarrow \|v\|^2 \leqslant 0 \leqslant \|v\|^2 \Rightarrow \|v\| = 0 \Rightarrow v = 0$$

which shows that ψ is positive-definite and concludes the proof. ∎

Definition 2.3.6. *If G is a compact group and f_1 and f_2 are functions from G to $\mathbb{C}$, we define $(f_1|f_2)$ as*

$$(f_1|f_2) = \int_G f_1(s)\overline{f_2(s)}\,\mathrm{d}s.$$

This definition of course extends the scalar product of Definition 1.4.1 to compact groups.

2.4 The Peter-Weyl theorem

In Section 1.5, we studied the decomposition of the left regular representation of a finite group. These results were generalize to compact groups in 1927, by Peter and Weyl, in the so-called Peter-Weyl theorem.

Although the demonstration of this theorem requires a more functional analysis than can reasonably be done here, its importance is such that we wish to at the least state it. Moreover, we will require parts of its statement in our treatment of the irreducible representations of $SU(2)$. We will follow the presentation done by Sternberg in [7], which states the theorem in terms of representative functions.

Definition 2.4.1. *Let f be a continuous function on a compact group G. We say that f is a* representative function *if the space generated by its left translations is of finite dimension, i.e. if* $\dim(\mathrm{Span}\{\lambda_t f \mid t \in G\}) < \infty$. *The set of all representative functions of G will be denoted* $\mathrm{Rep}(G)$.

Notation 2.4.2. Let G be a compact group, and (π, V) be a representation of degree n. We call *coefficients* the functions corresponding to the matrix elements r_{ij} (see Notation 1.1.2), where we have chosen a basis $\{e_1, \ldots, e_n\}$ of V. Note that for $1 \leqslant i, j \leqslant n$, $s \in G$, we can also write

$$r_{ij}(s) = (\pi_s e_j | e_i)$$

where $(\cdot|\cdot)$ is the usual scalar product on V. We will denote the span of all such coefficient functions, of all finite dimensional representations of G, as $\mathrm{Coeff}(G)$. ($\mathrm{Coeff}(G)$ itself does not form a vector space. This can be seen by trying to construct representations of the cyclic group C_2 whose set of coefficients is closed under scalar multiplication.)

Proposition 2.4.3. *The set of representative functions on a compact group G is precisely the span of the coefficients of finite dimensional representations of G.*

Proof: Let G be a compact group. We must show that $\mathrm{Rep}(G) = \mathrm{Coeff}(G)$.

Let π be a finite, unitary representation of degree n, H be the space on which it acts and $\{e_1, \ldots, e_n\}$ be a basis of H. For all $s, t \in G$, $1 \leqslant i, j \leqslant n$, we have

$$\begin{aligned} \lambda_s r_{ij}(t) &= \lambda_s(\pi_t e_j \mid e_i) \\ &= (\pi_{s^{-1}t} e_j \mid e_i). \end{aligned} \tag{2.4.1}$$

Since each π_s is invertible, we know there exist $\{a_k \in \mathbb{C}\}_{k \in G}$ such that

$$e_i = \sum_{k \in G} a_k \pi_s^* e_k,$$

where π_s^* is the adjoint of π_s. Then (2.4.1) becomes

$$\begin{aligned}\lambda_t r_{ij}(t) &= (\pi_{s^{-1}t} e_j \mid \sum_{k\in G} a_k \pi_s^*\ e_k) \\ &= \sum_{k\in G} a_k (\pi_s \pi_{s^{-1}t} e_j \mid\ e_k) \\ &= \sum_{k\in G} a_k (\pi_t e_j \mid\ e_k) \\ &= \sum_{k\in G} a_k r_{kj}(t)\end{aligned}$$

Thus the left translations of r_{ij} are contained in the span of $\{r_{kj} \mid 1 \leqslant k \leqslant n\}$, which has finite dimension. This shows that every coefficient of a finite representation is a representative function, i.e. that $\mathrm{Coeff}(G) \subset \mathrm{Rep}(G)$.

Suppose now that we are given a representative function f. We denote by W the space spanned by left translations of f and choose a basis $\{f_i\}$ of W, where the $f_i = \lambda_{t_i} f$ for some $t_i \in G$. Then there exist functions $h_i : G \to \mathbb{C}$ such that

$$\lambda_s f = \sum h_i(s) f_i \quad \text{for all } s \in G$$

Now we wish to determine the coefficients of $\lambda_{s^{-1}}$ in the space spanned by the $\{f_i\}$. To do this, we apply $\lambda_{s^{-1}}$ to an element f_j of the basis, and write the result as a linear combination of basis elements f_i:

$$\begin{aligned}\lambda_{s^{-1}} f_j = \lambda_{s^{-1}t_j} f &= \sum_i h_i(s^{-1}t_j) f_i \\ (\lambda_{s^{-1}} f_j | f_i) &= h_i(s^{-1}t_j).\end{aligned}$$

Of course, f can be written as a linear combination of the coefficient functions h_i:

$$f(s) = \lambda_{s^{-1}} f(e) = \sum_j \lambda_{s^{-1}t_j} f_j(e) = \sum_{i,j} h_i(s^{-1}t_j) f_i(e)$$

where the $f_i(e)$ appear as constants in the linear combination. Thus $\mathrm{Rep}(G) \subset \mathrm{Coeff}(G)$ and

$$\begin{gathered}\mathrm{Rep}(G) \subset \mathrm{Coeff}(G) \subset \mathrm{Rep}(G), \\ \mathrm{Rep}(G) = \mathrm{Coeff}(G),\end{gathered}$$

which completes the proof. ∎

Remark 2.4.4. Since the set of coefficients of irreducible finite dimensional representations is the same as the one of coefficients of finite representations, Proposition 2.4.3 remains valid when we consider only *irreducible* finite representations.

We can now finally state the Peter-Weyl theorem

Theorem 2.4.5. *Let G be a compact group. Then*

i) The set of representative functions is dense in $L^2(G)$

ii) $L^2(G)$ decomposes into a direct sum of finite irreducible representations on Hilbert spaces: $L^2(G) = \oplus_{i \in I} H_i$

iii) Every irreducible representation of G is finite dimensional.

iv) Each irreducible representation of G occurs in $L^2(G)$ with multiplicity equal to its degree.

v) If π is a unitary representation of G on a Hilbert space, then it decomposes into a direct sum of finite irreducible representations on Hilbert spaces.

A result equivalent to theorem 1.5.9 is also true for compact groups:

Theorem 2.4.6. *Let $\{\chi_i\}$ be the set of distinct characters of the irreducible representations of a compact group G, and $\mathscr{C} \subset L^2(G)$ be the set of square integrable class functions. Then $\{\chi_i\}$ is an orthonormal basis of $\mathscr{C}$.*

Proofs of theorems 2.4.5 and 2.4.6 are done by Sternberg [7] and Robert [4].

Chapter 3

Describing the dynamics of spin by the representations of SU(2)

We are now in a position to analyze spin using the tools of representation theory. This will be done in three parts. A first brief section will sketch the physical framework, showing how quantum mechanics can be expressed in the language of group theory. We will then construct the Haar measure on $SU(2)$, in order to finally determine its irreducible representations.

We should note that the treatment we will do of spin is actually valid in general for any angular momentum; we simply chose the simplest type in the context of the example. Also, although the presentation has been adapted and the physical realisation explicited by the author, the treatment is largely based on that of Sternberg [7].

3.1 Fundamental concepts

In quantum mechanics, the "state" ψ of a system is the set of all its properties which are of interest. For example, one might be interested in describing the positions and velocities of two electrons and a nucleus within an atom. Unfortunately, such a system is already surprisingly difficult, and in general does not admit an analytical solution. Therefore, for this example, we will study just about the simplest of systems: a single particle in vacuum (placing the electron in vacuum saves us from having to consider interactions with its environment). Furthermore, we will consider only one of its properties: its spin.

A useful, although not entirely correct, way of looking at spin is to consider it as the amount by which a particle is spinning on itself (hence the name). Its value has physical effects and can be measured, so we call spin an "observable".

Mathematically, a quantum mechanical system is expressed as a separable Hilbert space $\mathscr{H}$ and an observable is a self-adjoint operator on $\mathscr{H}$; for spin, we will denote

this operator as S. We usually describe the space using an orthonormal basis $\{\phi_i\}$ and call "pure states" the one-dimensional spans of each basis vector. A state is any one-dimensional subspace of $\mathscr{H}$, and can be expressed as a linear combination of pure states (in what follows, we will designate by ϕ both a state and its associated unit vector). Each state assigns a probability distribution to the possible values of a given operator. For example, in the case of an electron's spin, there are only two possible values: $s_+ = +\frac{1}{2}$ and $s_- = -\frac{1}{2}$. So an arbitrary state ψ might assign the probability distribution

$$P_{S,\psi}(\frac{1}{2}) = \frac{2}{3}$$
$$P_{S,\psi}(-\frac{1}{2}) = \frac{1}{3}.$$

s_+ and s_- are in fact eigenvalues of S. And because S is self-adjoint, it admits corresponding orthogonal eigenvectors ϕ_i, for whom $S\phi_i = s_i\phi_i$, $i = +, -$. The probability distribution for such eigenvectors is of course $P_{S,\phi_i}(s_j) = \delta_{i,j}$. Therefore, we could write the above state ψ as a linear combination of two eigenstates ϕ_+ and ϕ_-:

$$\psi = \sqrt{\frac{2}{3}}\phi_+ + \sqrt{\frac{1}{3}}\phi_-$$

This particular choice of coefficients allows us to write the expected value of S on the state ψ using the inner product on $\mathscr{H}$:

$$\begin{aligned}(S\psi,\psi) &= \frac{2}{3}(S\phi_+,\phi_+) + \frac{\sqrt{2}}{3}(S\phi_+,\phi_-) + \frac{\sqrt{2}}{3}(S\phi_-,\phi_+) + \frac{1}{3}(S\phi_-,\phi_-) \\ &= \frac{2}{3}\left(\frac{1}{2}\right)\underbrace{(\phi_+,\phi_+)}_{=1} + \frac{\sqrt{2}}{3}\left(\frac{1}{2}\right)\underbrace{(\phi_+,\phi_-)}_{=0} + \frac{\sqrt{2}}{3}\left(-\frac{1}{2}\right)\underbrace{(\phi_-,\phi_+)}_{=0} + \frac{1}{3}\left(-\frac{1}{2}\right)\underbrace{(\phi_-,\phi_-)}_{=1} \\ &= \frac{1}{6}\end{aligned}$$

where we've used the fact that ϕ_+ and ϕ_- are both orthogonal and normalized.

As with most physical systems, there is a group of symmetries under which the space $\mathscr{H}$ should be invariant. For reasons into which we do not wish to delve, when $\mathscr{H}$ is a space corresponding to spin states, it must be invariant under $SU(2)$. In other words, any space of spin states $\mathscr{H}$ appears as direct sum of irreducible representations of $SU(2)$. Finding these irreducible representations would therefore exactly determine the possible state spaces for spin.

3.2 The Haar measure of SU(2)

Recall that $SU(2)$ is the set of all 2×2 complex matrices whose determinant is 1. Therefore a given element t of $SU(2)$ is determined by two complex numbers

$z = x_1 + iy_1$ and $w = x_2 + iy_2$ and can be written as

$$t = \begin{pmatrix} \overline{z} & -w \\ \overline{w} & z \end{pmatrix}$$

where $|z|^2 + |w|^2 = x_1^2 + x_2^2 + y_1^2 + y_2^2 = 1$. There is of course a natural isomorphism between $SU(2)$ and S^3, thanks to which we can identify functions on $SU(2)$ to functions on S^3; we will do so implicitly where it simplifies the notation.

What we want to find is a regular measure μ such that, for any function $f\colon SU(2) \to \mathbb{C}$, $\int_{SU(2)} f \,\mathrm{d}\mu = \int_{SU(2)} \lambda_s f \,\mathrm{d}\mu$, where λ_s is a left translation. It will then suffice to normalize μ to obtain the Haar measure (Theorem 2.2.4). We will need to go a little beyond what was properly defined in Section 2.1 and use standard multi-variable calculus techniques; however the reader should be familiar with these, even if he or she is not familiar with their construction from measure theory.

Let f be a function on S^3 and define

$$D = \left\{ \begin{pmatrix} z \\ w \end{pmatrix} \in \mathbb{C}^2 : 0 < |z|^2 + |w|^2 \leqslant 1 \right\}.$$

We extend f to D by defining

$$\begin{array}{rccc} \tilde{f}\colon & D & \to & \mathbb{C} \\ & \begin{pmatrix} z \\ w \end{pmatrix} & \mapsto & \sqrt{|z|^2 + |w|^2} f\left(\frac{z}{\sqrt{|z|^2+|w|^2}}, \frac{w}{\sqrt{|z|^2+|w|^2}} \right) \end{array}$$

Let $s = \begin{pmatrix} \overline{a} & -b \\ \overline{b} & a \end{pmatrix} \in SU(2)$. Writing $z = x_1 + iy_1$ and $w = x_2 + iy_2$ as above, the group operation $t \mapsto t' = st$ can be extended to a function on $\mathbb{R}^4$:

$$\begin{array}{rccc} s\colon & \mathbb{R}^4 & \to & \mathbb{R}^4 \\ & \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix} & \mapsto & \begin{pmatrix} \overline{a}x_1 - bx_2 \\ \overline{a}y_1 - by_2 \\ \overline{b}x_1 + ax_2 \\ \overline{b}y_1 + ay_2 \end{pmatrix} = \begin{pmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \end{pmatrix} \end{array}$$

The Jacobian of s is then

$$J_s = \begin{pmatrix} \frac{\partial}{\partial x_1'}x_1' & \frac{\partial}{\partial x_1'}y_1' & \frac{\partial}{\partial x_1'}x_2' & \frac{\partial}{\partial x_1'}y_2' \\ \frac{\partial}{\partial y_1'}x_1' & \frac{\partial}{\partial y_1'}y_1' & \frac{\partial}{\partial y_1'}x_2' & \frac{\partial}{\partial y_1'}y_2' \\ \frac{\partial}{\partial x_2'}x_1' & \frac{\partial}{\partial x_2'}y_1' & \frac{\partial}{\partial x_2'}x_2' & \frac{\partial}{\partial x_2'}y_2' \\ \frac{\partial}{\partial y_2'}x_1' & \frac{\partial}{\partial y_2'}y_1' & \frac{\partial}{\partial y_2'}x_2' & \frac{\partial}{\partial y_2'}y_2' \end{pmatrix} = \begin{pmatrix} \overline{a} & 0 & -b & 0 \\ 0 & \overline{a} & 0 & -b \\ \overline{b} & 0 & a & 0 \\ 0 & \overline{b} & 0 & a \end{pmatrix}$$

and its determinant is

$$\det(J_s) = (|a|^2 + |b|^2)^2 = 1.$$

Hence we have, when integrating over $\mathbb{R}^4$ using the usual Lebesgue measure and writing $\boldsymbol{x} = (x_1, y_1, x_2, y_2)^\intercal$,

$$\begin{aligned}\int_D \lambda_s \tilde{f}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_D \tilde{f}(s^{-1}\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\ &= \int_D \tilde{f}(s\circ s^{-1}\boldsymbol{x})\,|\det(J_s)|\,\mathrm{d}\boldsymbol{x} \\ &= \int_D \tilde{f}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Which shows that $\tilde{f}$ is $SU(2)$ invariant.

Now we perform the change of variables

$$\begin{aligned}z &= r(\cos\alpha + i\sin\alpha\cos\beta) \\ w &= r(\sin\alpha\sin\beta e^{i\gamma})\end{aligned}$$

where $\begin{pmatrix}\alpha\\ \beta\\ \gamma\end{pmatrix} \in E \overset{\text{def}}{=} [0,\pi]\times[0,\pi]\times[0,2\pi]$. This can be described by the following map:

$$\begin{array}{rccc}\tilde{\tau}\colon & \mathbb{R}\times E & \to & \mathbb{R}^4 \\ & \begin{pmatrix}r\\ \alpha\\ \beta\\ \gamma\end{pmatrix} & \mapsto & \begin{pmatrix}r\cos\alpha\\ r\sin\alpha\cos\beta\\ r\sin\alpha\sin\beta\cos\gamma\\ r\sin\alpha\sin\beta\sin\gamma\end{pmatrix}\end{array}$$

Writing $\boldsymbol{\theta} = (r,\alpha,\beta,\gamma)^\intercal$,

$$\tilde{f}(\tilde{\tau}\boldsymbol{\theta}) = rg(\alpha,\beta,\gamma),$$

for some function g on $\mathbb{R}^3$.

We can compute $J_{\tilde{\tau}}$ just as we did J_s, and find that $|\det J_{\tilde{\tau}}| = r^3\sin^2\alpha\sin\beta$. Therefore the left invariant measure becomes $\mathrm{d}\eta = r^3\sin^2\alpha\sin\beta\,\mathrm{d}\boldsymbol{\theta}$ (we will not bother explicitly defining the measure itself, since what we want is really its differential form). We define two measures ν' and μ' on $\mathbb{R}$ and $\mathbb{R}^3$ respectively, such that $\mathrm{d}\nu' = r^3\,\mathrm{d}r$ and $\mathrm{d}\mu' = \sin^2\alpha\sin\beta\,\mathrm{d}\alpha\,\mathrm{d}\beta\,\mathrm{d}\gamma$. Then $\eta = \nu'\times\mu'$ is the product measure of ν' and μ', and we have

$$\begin{aligned}\int_D \tilde{f}\circ\tilde{\tau}\,\mathrm{d}\eta &= \int_D rg\,\mathrm{d}\eta \\ &= \int_0^1 r\,\mathrm{d}\nu' \int_E g\,\mathrm{d}\mu'\end{aligned}$$

It is easy to verify that action of $SU(2)$ on a vector $\boldsymbol{\theta}$ does not change its coordinate r. Therefore, $\int_D \tilde{f}\circ\tilde{\tau}\,\mathrm{d}\eta$ and $\int_0^1 r\,\mathrm{d}\nu'$ are both $SU(2)$ invariant, which implies

that $\int_E g \, \mathrm{d}\mu'$ is also $SU(2)$ invariant. Note now that g is precisely $f \circ \tau$, where τ is the change of variables

$$\begin{array}{rccc} \tau\colon & E & \to & \mathbb{C}^2 \\ & \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} & \mapsto & \begin{pmatrix} \cos\alpha + i \sin\alpha \cos\beta \\ \sin\alpha \sin\beta e^{i\gamma} \end{pmatrix} . \end{array}$$

Thus μ' is the required invariant measure. We simply need to normalize it to obtain the Haar measure μ:

$$\int_E \sin^2\alpha \sin\beta \, \mathrm{d}\alpha \, \mathrm{d}\beta \, \mathrm{d}\gamma = 2\pi^2$$

$$\mathrm{d}\mu = \frac{1}{2\pi^2} \sin^2\alpha \sin\beta \, \mathrm{d}\alpha \, \mathrm{d}\beta \, \mathrm{d}\gamma,$$

where an element $\begin{pmatrix} \overline{z} & -w \\ \overline{w} & z \end{pmatrix}$ of $SU(2)$ is expressed by

$$\begin{aligned} z &= \cos\alpha + i \sin\alpha \cos\beta \\ w &= \sin\alpha \sin\beta e^{i\gamma}, \end{aligned} \tag{3.2.1}$$

with $(\alpha, \beta, \gamma)^\intercal \in E$.

Before moving on to determining the irreducible representations of $SU(2)$, we make an important remark on class functions:

Remark 3.2.1. Let f be a class function on $SU(2)$. Recall that this implies that for all s, t in $SU(2)$,

$$f(s^{-1}ts) = f(t).$$

We wish to see what conditions this imposes on f. If we write t and s as

$$t = \begin{pmatrix} \overline{z_1} & -w_1 \\ \overline{w_1} & z_1 \end{pmatrix} \quad s = \begin{pmatrix} \overline{z_2} & -w_2 \\ \overline{w_2} & z_2 \end{pmatrix},$$

then the product is given by

$$s^{-1}ts = \begin{pmatrix} \mathrm{Re}(z_1) - 2i \, \mathrm{Im}(w_1 z_2 \overline{w_2}) & 2i z_2 w_2 \, \mathrm{Im}(z_1) - w_1 z_2^2 - \overline{w_1} w_2^2 \\ \quad - i(|z_2|^2 - |w_2|^2) \, \mathrm{Im}(z_1) & \\ 2i \overline{z_2 w_2} \, \mathrm{Im}(z_1) + \overline{w_1 z_2}^2 - w_1 \overline{w_2}^2 & \mathrm{Re}(z_1) + 2i \, \mathrm{Im}(w_1 z_2 \overline{w_2}) \\ & \quad + i(|z_2|^2 - |w_2|^2) \, \mathrm{Im}(z_1) \end{pmatrix} \tag{3.2.2}$$

Comparing the expressions for $s^{-1}ts$ and t, we see that in order for f to be independent over classes, a necessary condition is that it depend only on $\mathrm{Re}(z_1) =$

$\cos\alpha$; we will show in the next section that this condition is also sufficient. Thus the conjugacy classes of $SU(2)$ are determined by the value of $\cos\alpha$, and the scalar product of two class functions simplifies to

$$\begin{aligned}(f_1 \mid f_2) &= \frac{1}{2\pi^2}\int_G f_1(\alpha,\beta,\gamma)\overline{f_2(\alpha,\beta,\gamma)}\sin^2\alpha\ \sin\beta\,\mathrm{d}\alpha\,\mathrm{d}\beta\,\mathrm{d}\gamma \\ &= \frac{2}{\pi}\int_0^\pi f_1(\alpha)\overline{f_2(\alpha)}\sin^2\alpha\,\mathrm{d}\alpha. \end{aligned} \tag{3.2.3}$$

We are now ready to determine the irreducible representations of $SU(2)$.

3.3 The irreducible representations of $SU(2)$

Following the approach in [7], we will determine the irreducible representations of $SU(2)$ in two steps: first we will find the possible forms of its characters, then we will show that each of these characters indeed corresponds to a representation. We will conserve the notation of (3.2.1) for the elements of $SU(2)$.

Theorem 3.3.1. *If (ρ, V) is an irreducible representation of $SU(2)$, then its character χ must be of the form*

$$\chi(\alpha) = \sum_{k=0}^{2s} e^{i\,2(-s+k)\alpha},$$

where $s = \frac{n-1}{2}$, $n = 1, 2, \ldots$.

Our proof will require the following result:

Lemma 3.3.2. *Let g be a continuous function from $\mathbb{R}$ to $\mathbb{R}$. If*

$$g(x+y) = g(x) + g(y)\ \forall x, y \in \mathbb{R},$$

then g is linear and can be written $g(x) = cx$ for some $c \in \mathbb{R}$.

Proof: Let $x \in \mathbb{R}$. Then for all $n \in \mathbb{N}$,

$$\begin{aligned} g(x) &= g\left(\frac{1}{n}x\right) + \cdots + g\left(\frac{1}{n}x\right) \quad (n \text{ times}) \\ &= n \cdot g\left(\frac{1}{n}x\right) \\ g\left(\frac{1}{n}x\right) &= \frac{1}{n}g(x). \end{aligned}$$

This result can easily be extended to conclude that for any $c \in \mathbb{Q}$,

$$g(cx) = cg(x). \tag{3.3.1}$$

But then the facts that $\mathbb{Q}$ is dense in $\mathbb{R}$ and that g is continuous imply that (3.3.1) is in fact true for all $c \in \mathbb{R}$. ∎

Proof of theorem 3.3.1:

Define $A(\alpha) \in SU(2)$ and an arbitrary element $u(\alpha', \beta', \gamma') \in SU(2)$ as

$$A(\alpha) = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \tag{3.3.2}$$

$$u(\alpha', \beta', \gamma') = \begin{pmatrix} \cos\alpha' - i\sin\alpha'\cos\beta' & -\sin\alpha'\sin\beta' e^{i\gamma'} \\ \sin\alpha'\sin\beta' e^{-i\gamma'} & \cos\alpha' + i\sin\alpha'\cos\beta' \end{pmatrix}. \tag{3.3.3}$$

We begin by proving that for $0 \leqslant \alpha < \pi$,

$$\{s^{-1}A(\alpha)s : s \in SU(2)\} = \{u(\alpha, \beta', \gamma') \in SU(2) : (\beta', \gamma') \in [0, \pi] \times [0, 2\pi]\}. \tag{i}$$

Substituting $A(\alpha)$ for t in (3.2.2), we find that $s^{-1}A(\alpha)s = u(\alpha', \beta', \gamma')$ if and only if

$$\cos\alpha + i(|z_2|^2 - |w_2|^2)\sin\alpha = \cos\alpha' + i\sin\alpha'\cos\beta' \tag{3.3.4}$$

and

$$-2iz_2w_2\sin\alpha = \sin\alpha'\sin\beta' e^{i\gamma'} \tag{3.3.5}$$

The first equation yields $\cos(\alpha') = \cos(\alpha)$. Since α is in the interval $[0, \pi)$, this effectively implies that $\alpha = \alpha'$. The choice of interval also prevents ambiguities, as $A(\alpha)$ and $A(-\alpha)$ would belong to the same class. To find now the conditions on β' and γ', we express z_2 and w_2 as

$$z_2 = \cos\theta e^{i\phi}$$
$$w_2 = \sin\theta e^{i\psi},$$

where $0 \leqslant \theta \leqslant \frac{\pi}{2}$ and $0 \leqslant \phi, \psi \leqslant 2\pi$. (This is equivalent to (3.2.1)). Then (3.3.4) becomes

$$\begin{aligned}(|z_2|^2 - |w_2|^2)\sin\alpha = \cos^2\theta - \sin^2\theta &= \cos\beta' \\ \cos 2\theta &= \cos\beta'\end{aligned} \tag{3.3.6}$$

while (3.3.5) becomes

$$\begin{aligned}-2iz_2w_2 = -2i\sin\theta\cos\theta e^{i(\theta+\psi)} &= \sin\beta' e^{i\gamma'} \\ \sin 2\theta e^{i(\theta+\phi+3\pi/2)} &= \sin\beta' e^{i\gamma'}.\end{aligned} \tag{3.3.7}$$

From (3.3.6) and (3.3.7), it is clear that for any given α, β', γ', there exists an s in $SU(2)$ such that $s^{-1}A(\alpha)s = u(\alpha, \beta', \gamma')$. This proves the assertion (i), which implies that the set $\{A(\alpha) : 0 \leqslant \alpha \leqslant \pi\}$ forms a set of representatives of the conjugacy classes of $SU(2)$. We will denote by C_α the conjugacy class represented by $A(\alpha)$.

Now operators $\{\rho_{A(\alpha}\}_{0\leqslant\alpha\leqslant\pi}$ are unitary, and therefore diagonalisable. Furthermore, the matrices $A(\alpha)$ all commute since they are diagonal. Therefore the operators $\{\rho_{A(\alpha}\}_{0\leqslant\alpha\leqslant\pi}$ must also commute, implying that they are in fact *simultaneously* diagonalisable and that they share a basis $\mathscr{B}$ of eigenvectors of V. Then for $v_j \in \mathscr{B}$ and $0 \leqslant \alpha < \pi$,

$$\rho_{A(\alpha)}v_i = \lambda_i(\alpha)v_i.$$

Because ρ is unitary, we have

$$(v_j \mid v_j) = (\rho_{A(\alpha)}v_j \mid \rho_{A(\alpha)}v_j) = \overline{\lambda_j(\alpha)}\lambda_j(\alpha)(v_j \mid v_j)$$

and therefore $\lambda(\alpha) = e^{ig(\alpha)}$, for some real function g of α. Also, since $A(\theta)$ is continuous in θ and $\rho_{A(\theta)}v$ is continous in both $A(\theta)$ and v, g is also continuous.

Now from the fact that ρ is a homomorphism,

$$\rho_{A(\alpha_1)}\rho_{A(\alpha_2)}v_j = \rho_{A(\alpha_1)A(\alpha_2)}v_j,$$

we have that

$$\begin{aligned}\lambda_j(\alpha_1)\lambda_j(\alpha_2) &= \lambda_j(\alpha_1 + \alpha_2)\\ e^{i(g(\alpha_1)+g(\alpha_2))} &= e^{ig(\alpha_1+\alpha_2)}\\ g(\alpha_1) + g(\alpha_2) &= g(\alpha_1 + \alpha_2) + 2n(\alpha_1, \alpha_2)\pi.\end{aligned}$$

where $n(\alpha_1, \alpha_2)$ takes values in $\mathbb{Z}$. Since g is continuous, the same is true of n, which implies that n is in fact constant. Therefore we can then consider $g' = g + 2n\pi$, for which the equality becomes

$$g'(\alpha_1) + g'(\alpha_2) = g'(\alpha_1 + \alpha_2), \tag{3.3.8}$$

which implies, by Lemma 3.3.2, that g' is linear on $[0, \pi]$ and is simply a multiplication by a constant c_j:

$$\lambda_j(\alpha) = e^{ic_j\alpha}.$$

Since π is irreducible, part (iii) of the Peter-Weyl theorem states that V must be finite; therefore we can speak in a meaningful way of its character. And since the character evaluated at a group element is the sum of the eigenvalues of the corresponding operator, for any $t \in C_\alpha$, we have

$$\chi(t) = \sum e^{ic_j\alpha}. \tag{3.3.9}$$

All that is left now is to determine the values of these c_j.

Since class functions depend only on $\cos(\alpha)$, we must have $\chi(A(\alpha)) = \chi(A(-\alpha))$. This implies that for every c_j that occurs, $-c_j$ also occurs with the same multiplicity. Additionally, part of the statement of Theorem 2.4.6 is that the characters of irreducible representations have unit length. Then using (3.2.3), we have

$$\begin{aligned} 1 = (\chi \mid \chi) &= \frac{2}{\pi}\int_0^{\pi} \chi(\alpha)\overline{\chi(\alpha)}\sin^2\alpha\,\mathrm{d}\alpha \\ &= \frac{1}{\pi}\int_0^{2\pi} \chi(\alpha)\overline{\chi(\alpha)}\sin^2\alpha\,\mathrm{d}\alpha \end{aligned} \tag{3.3.10}$$

Now we write

$$\sin^2\alpha = \frac{1-\cos 2\alpha}{2} = \frac{2 - e^{2i\alpha} - e^{-2i\alpha}}{4} \tag{3.3.11}$$

and the integral (3.3.10) becomes

$$\begin{aligned} 1 &= \frac{1}{4\pi}\sum_{j,k}^{n}\int_0^{2\pi} e^{ic_j\alpha}e^{-ic_k\alpha}(2 - e^{2i\alpha} - e^{-2i\alpha})\,\mathrm{d}\alpha \\ &= \frac{1}{4\pi}\sum_{j,k}^{n}\left[\int_0^{2\pi} 2e^{i\alpha(c_j-c_k)}\,\mathrm{d}\alpha - \int_0^{2\pi} e^{i\alpha(c_j-c_k+2)}\,\mathrm{d}\alpha - \int_0^{2\pi} e^{i\alpha(c_j-c_k-2)}\,\mathrm{d}\alpha\right]. \end{aligned} \tag{3.3.12}$$

Noting that

$$\int e^{im\alpha}\,\mathrm{d}\alpha = \begin{cases} 0 & \text{if } m \neq 0 \\ 2\pi & \text{if } m = 0 \end{cases},$$

we only keep the non-zero terms in the integral (3.3.12). This can happen in one of two ways:

i) if $c_j = c_k$. In this case the integral yields $\frac{4\pi}{4\pi} = 1$.

ii) if $|c_j - c_k| = 2$, in which case the integral yields $-\frac{2\pi}{4\pi} = -\frac{1}{2}$.

We obviously have (i) whenever $j = k$, which occurs n times. (i) can also happen if some of the eigenvalues have multiplicity greater than 1; we'll say this happens $2a$ times (we multiply by 2 because each pairing with $j \neq k$ appears twice in the product). Similarly, we'll say that (ii) occurs $2b$ times.

Putting all these terms together, we get

$$(\chi \mid \chi) = (n + 2a)\cdot 1 + 2b\cdot\left(-\frac{1}{2}\right) = n + 2a - b$$

$$1 = n + 2a - b \tag{3.3.13}$$

In order to satisfy (3.3.13), we seek to maximize b. Therefore, we require each c_j to be paired with at least another term c_k such that $|c_j - c_k| = 2$. We can visualize the situation as placing the c_j into "bins" beginning at some arbitrary number z, with multiplicity m_j in each bin.

$\left.\begin{matrix} c_1 \\ \vdots \\ c_1 \end{matrix}\right\} m_1$	$\left.\begin{matrix} c_2 \\ \vdots \\ c_2 \end{matrix}\right\} m_2$	$\dots$	$\left.\begin{matrix} c_l \\ \vdots \\ c_l \end{matrix}\right\} m_l$
z	$z+2$	$\dots$	$z+2(l-1)$

b will be given by the sum of the number of possible pairings between each bin:

$$b = \sum_{j=1}^{l-1} m_j m_{j+1}.$$

Therefore, higher multiplicities lead to a higher value for b. However, higher multiplicities also raise the parameter a, which is a sum of combinatorials:

$$2a = 2\sum_{j=1}^{l} \binom{m_j}{2}$$

so any gain in b made by raising multiplicities will be removed by a greater raising of a. The minimum value for the right hand side of (3.3.13) is therefore obtained when each "bin" contains exactly one c_j, in which case $b = n - 1$. Since this unique minimum solution verifies (3.3.13), there can be no other solution to (3.3.13). Thus the c_j form an increasing sequence with $c_{j+1} - c_j = 2$.

We can deduce the values of the c_j from the fact that they are symmetric about the origin. Indeed, this implies that

$$\{c_j\}_{j=1}^{n} = \{-n+1, -n+3, \dots, n-3, n-1\}.$$

Typically, the characters are indexed by $s = (n-1)/2$, so we can rewrite (3.3.9) in the form used in the theorem's statement:

$$\chi_s(\alpha) = \sum_{k=0}^{2s} e^{i\,2(-s+k)\alpha}$$

∎

At this point we have shown the possible forms of the characters of $SU(2)$; we must still show that each of these indeed corresponds to an irreducible representation.

Proposition 3.3.3. *To each character of Theorem 3.3.1 corresponds an irreducible representation V_s of $SU(2)$*

Proof: Recall that $C(\mathbb{C}^2)$ is the vector space of all continuous functions on $\mathbb{C}^2$. The action of the regular representation on $C(\mathbb{C}^2)$ can be written as $\lambda_A(f)(\boldsymbol{z}) = f(A^{-1}\boldsymbol{z})$, where f in a function in $C(\mathbb{C}^2)$ and $\boldsymbol{z}$ a vector in $\mathbb{C}^2$. We define the subspace V_s of $C(\mathbb{C}^2)$ as the set of homogeneous polynomials of degree $2s$. Therefore, a basis of V_s is

$$\{z_1^{2s}, z_1^{2s-1}z_2, z_1^{2s-2}z_2^2, \ldots, z_2^{2s}\}.$$

Note that V_s is a subrepresentation of $C(\mathbb{C}^2)$. Indeed, if A is in $SU(2)$, f is in V_s and $\boldsymbol{z} = \left(\begin{smallmatrix} z_1 \\ z_2 \end{smallmatrix}\right)$ is in $\mathbb{C}^2$, then

$$A^{-1}\boldsymbol{z} = \begin{pmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{pmatrix}$$

for some a, b, c, d in $\mathbb{C}$. Therefore $f(A^{-1}z)$ will be a linear combination of terms of the form $(az_1 + bz_2)^{2s-k}(cz_1 + dz_2)^k$, each of which has degree $2s$. Hence $\lambda_A f$ is again an element of V_s.

Finally, if $A(\alpha)$ is a class representative as defined in (3.3.2), the character ϕ of the representation $\rho^s = \lambda|_{V_s}$ is given by

$$\phi(\alpha) = \mathrm{Tr}(\rho^s_{A(\alpha)}) = \sum \lambda_i v_i$$

where the v_i are eigenvectors of $\rho^s_{A(\alpha)}$. As each basis monomial is an eigenvector:

$$\begin{aligned} \rho^s_{A(\alpha)}(z_1^{2s-k}z_2^k) &= (e^{i\alpha}z_1)^{2s-k}(e^{-i\alpha}z_2)^k \\ &= e^{i2(s-k)\alpha}z_1^{2s-k}z_2^k, \end{aligned}$$

ϕ is simply the sum of their eigenvalues:

$$\begin{aligned} \phi(\alpha) &= \sum_{k=0}^{2s} e^{i2(s-k)\alpha} = \sum_{k=0}^{2s} e^{i2(-s+k)\alpha} \\ &= \chi_s(\alpha) \end{aligned}$$

Since $(\phi \mid \phi) = (\chi_s \mid \chi_s) = 1$, we know that V_s is an irreducible representation, and because the χ_s are the only possible characters, all of the irreducible representations of V must be isomorphic to some V_s. ∎

3.4 Application to spin

Lets see how this applies to our original problem of spin. Recall that the spin state spaces $\mathscr{H}$ of a particle are given by the irreducible representations of $SU(2)$, which we have denoted V_s. The immediate result from our analysis is that there are only countably many possible values for spin (one for each s); we say in this case that spin is *quantized*. This stunning fact is central to most of quantum physics, and it arises naturally from the analysis of the purely mathematical object $SU(2)$. However, single particles are not that difficult to describe, and in general complete analytical solutions for them have been known for decades. The real problem is describing systems with multiple particles, which can become intractable very quickly.

As an example of how we can treat these kinds of systems, we consider the problem of defining the spin state space of a system composed of two particles whose spins are s_1 and s_2, and whose state spaces are V_{s_1} and V_{s_2}. The possible states of one particle are not affected by the presence of the other (the probabilities associated to each might change, however). Therefore, all combinations of states are possible, and the system's spin space is expressed as a tensor product:

$$\mathscr{H}_{s_1+s_2} \cong V_{s_1} \otimes V_{s_2}.$$

To determine the structure of $\mathscr{H}_{s_1+s_2}$, we use Proposition 1.2.5 to compute its character:

$$\chi_{s_1+s_2}(\alpha) = \chi_{s_1}(\alpha) \cdot \chi_{s_2}(\alpha).$$

Now if we let $\zeta = e^{i2\alpha}$, each of the characters on the right hand side becomes

$$\chi_s(\alpha) = \zeta^{-s} + \zeta^{-s+1} + \cdots + \zeta^s = \frac{\zeta^{s+1} - \zeta^{-s}}{\zeta - 1}.$$

So for $s_1 \geqslant s_2$, we have

$$\begin{aligned} \chi_{s_1}(\alpha) \cdot \chi_{s_2}(\alpha) &= \frac{\zeta^{s_1+1} - \zeta^{-s_1}}{\zeta - 1} \cdot (\zeta^{-s_2} + \cdots + \zeta^{s_2}) \\ &= \frac{\zeta^{s_1+s_2+1} - \zeta^{-s_1-s_2}}{\zeta + 1} + \cdots + \frac{\zeta^{s_1-s_2+1} - \zeta^{-(s_1-s_2)}}{\zeta + 1}, \end{aligned} \tag{3.4.1}$$

where we have reordered the terms to obtain the final form. We see from (3.4.1) that for arbitrary s_1 and s_2,

$$\chi_{s_1} \cdot \chi_{s_2} = \chi_{|s_1-s_2|} + \chi_{|s_1-s_2|+1} + \cdots + \chi_{s_1+s_2}. \tag{3.4.2}$$

Therefore, we finally conclude that

$$\mathscr{H}_{s_1+s_2} \cong V_{s_1} \otimes V_{s_2} = V_{|s_1-s_2|} + V_{|s_1-s_2|+1} + \cdots + V_{s_1+s_2}. \tag{3.4.3}$$

This is what we call the Clebsch-Gordan decomposition of $SU(2)$.

The consequence of this decomposition is that the total spin of a system composed of two particles is *not* simply the sum of the spins, but can instead take on multiple values, the maximum of which is the sum of the spins. For example, if $s_1 = s_2 = \frac{1}{2}$, we obtain the well-known "singlet-triplet" structure, so named because the tensor product space decomposes into two spaces of dimensions one and three, corresponding to spins of 0 and 1:

$$V_{\frac{1}{2}} \otimes V_{\frac{1}{2}} = V_0 \oplus V_1.$$

Conclusion

In this paper, we have built the concepts necessary to gain some understanding of quantum mechanical spin through representation theory. Concretely, this meant finding the irreducible representations of the group of symmetries of spin, which in this case is $SU(2)$; we call this its Clebsch-Gordan decomposition. Although $SU(2)$ is only compact, we began by defining representations on finite groups; this was done because in that case proofs are not only more intuitive, but also do not require familiarity with measure theory or functional analysis. We demonstrated a number of crucial results related to characters and the regular representation, most of which were later directly extended to compact groups.

This extension was obtained by defining the Haar measure of a group. We showed that this measure is unique, and that it shares many of the same natural properties as the Lebesgue measure on the reals. Most importantly, the Haar measure over compact groups is always finite, allowing one to compute the average of a function on a group. The chapter on representations of compact groups concluded with the presentation of the Peter-Weyl theorem, which formed the basis of our treatment of $SU(2)$.

The final chapter concerned itself with demonstrating how one might use the results developped previously in order to determine the irreducible represesentations of a compact group. Thus we deduced the explicit form of the Haar measure of $SU(2)$, and determined the possible forms of its characters using the Peter-Weyl theorem. It was then shown that each of these characters indeed corresponds to an irreducible representation. Finally, we briefly explained how this decomposition can be used to analyze the properties of spin, in particular in the case of systems of many particles. Of course, the real interest of this approach is that it scales relatively well when the system (and therefore, the group of symmetries) becomes more complex. For this reason, representation theory is the basis of much of the current fundamental physics research.

Bibliography

[1] J. Baez and J. P. Muniain. *Gauge Fields, Knots and Gravity*, volume 4 of *Knots and Everything*. World Scientific Publishing Co. Pte. Ltd., 5 Toh Tuck Link, Singapore 569224, 1994.

[2] C. Cohen-Tannoudji, F. Diu, and F. Laloë. *Mécanique quantique*, volume I. Hermann, éditeurs des sciences et des arts, 23 rue Lecourbe, 75015 Paris, 2nd edition, 1998.

[3] D. L. Cohn. *Measure Theory*. Birkhäuser Boston, 675 Massachussetts Avenue, Cambridge, MA, 1980.

[4] A. Robert. Représentation des groupes compacts. January 1979. Course notes from a PhD course taught at Lausanne between October and December 1978.

[5] Walter Rudin. *Functional Analysis*. McGraw-Hill Book Company, 1991.

[6] J.-P. Serre. *Représentations linéaires des groupes finis*. Hermann, éditeurs des sciences et des arts, 23 rue Lecourbe, 75015 Paris, 3rd edition, 1978.

[7] S. Z. Sternberg. *Group theory and physics*. The Press Syndicate of the University of Cambridge, The Pitt Building, Trumpington Street, Cambridge CB2 1RP, United Kingdom, 1994.