Representation theory of compact Lie groups

Bruin Benthem

BSc thesis in mathematics and physics, 2006, universiteit Leiden

(supervised by Theo van den Bogaart, Gerard Nienhuis and Bas Edixhoven, translated from dutch to english and (slightly) edited in 2019 by Bas Edixhoven)*

The original version is here: https://www.universiteitleiden.nl/binaries/content/assets/science/mi/scripties/benthembachelor.pdf

Contents

*After this translation was done it was pointed out to us that Ernest B. Vinberg's [9], Chapter III already contains a lot of the material of this thesis.

1 Manifolds

1.1 Introduction

A differentiable manifold is, in short, a generalisation of Euclidean space[1], on which one can do analysis. In order to define Lie groups, we must know what differentiable manifolds are. Definitions and basic properties of groups and topological spaces are assumed to be known. We will give right away a some important examples of differentiable manifolds, such as $GL_n(\mathbb{C})$, $SO(n)$ and $SU(n)$, that will return frequently later in this thesis. Moreover we introduce tangent spaces and derivatives in order to discuss Lie algebras later.

1.2 Differentiable manifolds

Een differentiable manifold (from now on: manifold) can be seen as a generalisation of Euclidean space, on which, on which one can do analysis. An atlas is, in short, a set of homeomorphisms that identify the manifold locally with a Euclidean space.

Definition 1.1. *An n-*dimensional differentiabl atlas *$\mathcal{A}$ on a topological space X is a set of* charts *(U_i, φ_i), $i \in I$, where $U_i \subset X$ is an open subset and $\varphi_i : U_i \to U_i{}'$, with $U_i{}' \in \mathbb{R}^n$ open, a homeomorphism, such that the following 2 conditions are satisfied:*

1. $\bigcup_{i\in I} U_i = X$
2. *For all $i, j \in I$ the bijection $\varphi_j\varphi_i^{-1} : \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j) : p \mapsto \varphi_j(\varphi_i^{-1}(p))$ is C^∞ (infinitely differentiable, that is, smooth). The two charts are called* differentiably compatible.

With this notion of atlas we can define what a manifold is.

Definition 1.2. *An n-*dimensional manifold *is a pair $(X, \mathcal{A})$ (often defnoted just by "X") with X a topological space that is hausdorff and that has a countable basis for the topology, and with $\mathcal{A}$ an n-dimensional differentiable atlas on X.*

The definition of n-dimensional manifold is such that different atlases $\mathcal{A}$ and $\mathcal{B}$ can describe the same manifold; in that case we call $\mathcal{A}$ and $\mathcal{B}$ *equivalent*, meaning that $\mathcal{A} \cup \mathcal{B}$ is an n-dimensional differentiable atlas too. In order to remove this annoyance, we call an atlas $\mathcal{A}$ *maximal* if it contains all charts that are differentiably compatible with all charts from $\mathcal{A}$. Each atlas $\mathcal{A}$ is contained in a unique maximal atlas $\mathcal{D}(\mathcal{A})$ (add to $\mathcal{A}$ all charts that are differentiably compatible with all charts from $\mathcal{A}$). It is then not hard to see that

$$\mathcal{A} \sim \mathcal{B} \Leftrightarrow \mathcal{D}(\mathcal{A}) = \mathcal{D}(\mathcal{B})\,.$$

[1]Here, Euclidean space means 'finite dimensional real vector space.' In particular, no inner product is given.

And then we have the notion of submanifolds.

Definition 1.3. *Let $(X, \mathcal{A})$ be an n-dimensional manifold. A subspace $X' \subset X$ is called an k-* dimensional submanifold *of X if, for every $x \in X'$ there is a chart $(U_x, \varphi_x) \in \mathcal{D}(\mathcal{A})$ such that $x \in U_x$ and $\varphi_x(U_x \cap X') = \mathbb{R}^k \cap \varphi(U_x)$.*

Such a subspace X' is called a submanifold for a good reason: the set of charts $(U_x \cap X', \varphi_x|_{U_x \cap X'})$ is a k-dimensional atlas on X'.

Of course one can consider the product of 2 manifolds $(X, \mathcal{A})$ and $(Y, \mathcal{B})$. It is easy to see that the set $X \times Y$ with the product topology is hausdorff and second countable, and that the set

$$\mathcal{A} \times \mathcal{B} := \{(U \times V, \varphi \times \psi) : (U, \varphi) \in \mathcal{A}, (V, \psi) \in \mathcal{B}\}$$

is an $n+m$-dimensional differentiable atlas, because the charts

$$\varphi \times \psi : U \times V \to U' \times V' \subset_{\text{open}} \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$$

satisfy the required conditions.

Als certain quotients of manifolds can be made into manifolds, but this is more subtle than products, as one can see in Section 1.6 of [1]. In the chapter on Lie groups we will come back to this.

Now that we have defined the objects of our interest, manifolds, we can define the maps between them that are of our interest, the so-called *differentiable* maps.

Definition 1.4. *Let $(X, \{(U_i, \varphi_i) : i \in I\})$ and $(Y, \{(V_j, \psi_j) : j \in J\})$ be, respectively, an n-dimensional and an m-dimensional manifold, and let x be in X. A continuous map $f : X \to Y$ is called* differentiable in x *if for all (i, j) such that $x \in U_i$ en $f(x) \in V_j$ the map $\psi_j f \varphi_i^{-1}$ from $\varphi_i((f^{-1}V_j) \cap U_i) \subset \mathbb{R}^n$ to $\mathbb{R}^m$ is differentiable in $\varphi_i(x)$. The map f is called* differentiable, *or a* morphism of manifolds *if f is differentiable in all $x \in X$. If f is bijective and both f and f^{-1} are differentiable then f is called a* diffeomorphism.

If $f : X \to Y$ and $g : Y \to Z$ are differentiable then so is $g \circ f : X \to Z$. As composition is associative, and identity maps are morphisms, this gives us the *category of manifolds.*

Examples

1. An easy example of a manifold is $\mathbb{R}^n$. This is an n-dimensional manifold with atlas $\mathcal{A} = \{(\mathbb{R}^n, \mathrm{Id}_{\mathbb{R}^n})\}$.

2. Another easy example is $GL(V)$, the group of invertible linear transformations of a finite dimensional real vector space V, also denoted as $\mathrm{Aut}(V)$. In this thesis V will always be a real or complex vector space of (finite) dimension n, and a choice of a basis of V then gives an isomorphism from $GL(V)$ to $GL_n(K)$, the *general linear group*, that is, the group of invertible $n \times n$-matrices with coefficients in $K = \mathbb{R}$ or $\mathbb{C}$. The set $GL_n(\mathbb{R})$ is a manifold because it is a subset of $\mathbb{R}^{n^2}$ (the coordinates are the matrix coefficients). The fact that a square matrix is invertible if and only if the determinant is non-zero shows that $GL_n(\mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$. This embedding is the only chart in the atlas and makes $GL_n(\mathbb{R})$ into an n^2-dimensional manifold. The same procedure works for $GL_n(\mathbb{C})$, the group of invertible

$n \times n$-matrices with complex coefficients, and as $\mathbb{C} \cong \mathbb{R}^2$ (as real vector spaces) $GL_n(\mathbb{C})$ is a $2n^2$-dimensional manifold.

For many interesting subsets of manifolds it is not so easy to check that they are submanifolds directly from the definition of submanifold. Often it is easier to use the regular value theorem as in §1.4 of [1]; I will not discuss this here. With that theorem one can show that the following examples are submanifolds of $GL_n(\mathbb{R})$ or of $GL_n(\mathbb{C})$:

3. $SL_n(\mathbb{C})$, the *special linear group*, the group of complex $n \times n$-matrices with determinant 1, is a $2n^2 - 2$-dimensional submanifold of $GL_n(\mathbb{C})$.

4. $O(n)$, the *orthogonal group*, the set of real orthogonal $n \times n$-matrices (here we omit the field in which the matrixcoefficients take their values because the word "orthogonaal" implies that this concerns real matrices). The subset $O(n)$ is a $\frac{1}{2}n(n-1)$-dimensional submanifold van $GL_n(\mathbb{R})$.

5. $SO(n)$, the *special orthogonal group*, the set of real orthogonal $n \times n$-matrices with determinant 1. This is a $\frac{1}{2}n(n-1)$-dimensional submanifold of $GL_n(\mathbb{R})$. In physics and in the sequel of this thesis $SO(3)$ plays an important role because it is the group of rotations of the Euclidean 3-space.

6. $SU(n)$, the *special unitary group*, the set of unitary $n \times n$-matrices with determinant 1. By definition: $SU(n) := \{x \in M_n(\mathbb{C}) : \overline{x}^t x = 1, \quad \det(x) = 1\}$. This is an n^2-1-dimensional submanifold of $GL_n(\mathbb{C})$.

1.3 Tangent spaces

In order to define Lie algebras it is important that we define what tangent spaces are. For X an n-dimensional manifold and $p \in X$ we want that the tangent space at p, denoted $T_X(p)$, is the "linear approximation" of X at p. There are several ways to make this precise, and they are all equivalent, see also chapter 2 of [1]. We will use *one* of them.

Definition 1.5. *Let $(X, \mathcal{A})$ be an n-dimensional manifold and $p \in X$ a point. Let*

$$\mathcal{K}_X(p) := \{\alpha : (-\epsilon, \epsilon) \subset \mathbb{R} \to X \mid \epsilon > 0, \quad \alpha(0) = p\}$$

be the set of morphisms of manifolds from $(-\epsilon, \epsilon)$ to X. This set $\mathcal{K}_X(p)$ is called the set of differentiable curves on X through p. *The tangent space $T_X(p)$ is then defined as*

$$T_X(p) = \mathcal{K}_X(p)/\sim$$

for the equivalence relation[2] *$\alpha \sim \beta \Leftrightarrow (\varphi \circ \alpha)^{\cdot}(0) = (\varphi \circ \beta)^{\cdot}(0) \in \mathbb{R}^n$ for one (and therefore for any) chart $(U, \varphi) \in \mathcal{A}$ with $p \in U$. We denote the equivalence class of α by $[\alpha]$.*

This is a rather intuitive definition where we see tangent vectors in a point as derivatives of curves through the point. The next theorem gives the set $T_X(p)$ the structure of n-dimensional $\mathbb{R}$-vector space.

[2] The notation $f^{\cdot}(0)$ means derivative of f at 0.

Theorem 1.6. *For $(X, \mathcal{A})$ an n-dimensional manifold and $p \in X$, $T_X(p)$ is an n-dimensional $\mathbb{R}$-vector space with*

$$[\alpha] + [\beta] = [\gamma] \Leftrightarrow (\varphi \circ \alpha)\dot{}(0) + (\varphi \circ \beta)\dot{}(0) = (\varphi \circ \gamma)\dot{}(0)$$

for all charts $(U, \varphi) \in \mathcal{A}$.

A proof of this is given in [1], § 2.3, and in [2], § 1.8.2.

Examples

1. $GL_n(\mathbb{R})$ is open in the $\mathbb{R}$-vector space $M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$. This shows that tangent vectors of $GL_n(\mathbb{R})$ can be seen as elements of $M_n(\mathbb{R})$, hence $T_{GL_n(\mathbb{R})}(1) = M_n(\mathbb{R})$.

   One can also see this intuitively by looking at curves

   $$K_A : (-\epsilon, \epsilon) \to M_n(\mathbb{R}) : t \mapsto 1 + tA$$

   where $A \in M_n(\mathbb{R})$ and $\epsilon > 0$ is small enough. As ϵ is small enough we have that $[K_A]$ is invertible, with inverse $[K_{-A}]$ because

   $$(1 + tA)(1 - tA) = 1^2 - (tA)^2 = 1 \qquad \text{mod } t^2$$

   so that $K_A \in \mathcal{K}_{GL_n(\mathbb{R})}(1)$. Moreover $K_A(t)\dot{}(0) = A$ hence $K_A \sim K_B \Leftrightarrow A = B$. Therefore we have a bijection from $M_n(\mathbb{R})$ to $\mathcal{K}_{GL_n(\mathbb{R})}(1)/\sim$.

   In general we have that $\text{End}(V)$ is the tangent space at 1 of $\text{Aut}(V) = GL(V)$ for a finite dimensional real or complex vector space V.

   For the tangent spaces of the following submanifolds of $GL_n(K)$ I refer to § 1.2 of [3].

2. The tangent space $T_{SL_n(\mathbb{C})}(1)$ is the subspace of $M_n(\mathbb{C})$ consisting of matrices with trace 0, hence with the sum of their diagonal elements equal to 0.

3. The tangent space $T_{SO(n)}(1)$ is $M_n(\mathbb{R})^-$, consisting of the antisymmetric matrices, that is, the matrices A with $A^t = -A$.

4. The tangent space $T_{SU(2)}(1)$ is the subspace of $M_2(\mathbb{C})$ consisting of matrices A with trace 0 and $A^* = -A$, where A^* is defined as $A^* := \overline{A}^t$. We will see this tangent space many times. It is a 3-dimensional $\mathbb{R}$-vector space spanned by

   $$I = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, K = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

1.4 Derivative

Now that we have introduced the notion of tangent space we can define, in a natural way, what the derivative of a function is. For a function $f : \mathbb{R}^n \to \mathbb{R}^m$ we know such a way: we take as derivative $D(f)$, the jacobi matrix, whose (i, j)-coefficient is $\frac{\partial f_i}{\partial x_j}$. We generalise this principle to the notion of *derivative at a point of a morphism of manifolds.*

Definition 1.7. *Let $f : X \to Y$ be a morphism of manifolds from X to Y, and $x \in X$ a point. The* derivative of f in x *is defined as*

$$D(f)_x : T_X(x) \to T_Y(f(x)) : [\alpha] \mapsto [f \circ \alpha].$$

2 Lie groups and Lie algebras

2.1 Introduction

Lie groups were introduced by the mathematician Sophus Lie in 1870 in order to study symmetries of differential equations, and have been applied in many ways in todays physics. With our knowledge from the preceding chapter we can now define Lie groups and give some important examples. Moreover we will define Lie algebras because they are closely related to Lie groups and they make the determination of the representations of $SU(2)$ in the next chapter easier.

2.2 Lie groups

Now that we have looked at what a manifold is we can proceed to the notion of Lie group.

Definition 2.1. *A* Lie group *G is a topological space with an n-dimensional differentiable atlas and a group law such that the maps $G \times G \to G : (x, y) \mapsto xy$ and $G \to G : x \mapsto x^{-1}$ are differentiable.*

When G is compact as topological space then we call G a compact *Lie group.*

A morphism of Lie groups from G to G' *is a homomorphism $G \to G'$ that is also a morphism of manifolds.*

One can show that the examples from the previous chapter are not only manifolds but actually Lie groups, by showing that the above mentioned maps are differentiable.

In the proof of theorem 3.9 we will use the following lemma.

Lemma 2.2. *Let G be a Lie group and let $g \in G$. Then $l_g : x \mapsto gx$ and $r_g : x \mapsto xg$ are continuous and differentiable.*

Proof. I prove the lemma for l_g, the proof for r_g is analogous. Let G be a Lie group. The the map

$$l : G \to G \times G \ : \ x \mapsto (g, x)$$

is continuous and differentiable. On the first coordinate this map is the constant map $x \mapsto g$ and on the second coordinate it is the identity $x \mapsto x$. Both are continuous and differentiable, hence so is l. As the group law $* : G \times G \to G : (x, y) \mapsto xy$ is continuous and differentiable, so is the composition

$$l_g := * \circ l : G \to G : x \mapsto gx\,.$$

□

Examples

1. $GL_n(\mathbb{R})$ is a Lie group with group law the matrix multiplication "·". In this case it is quite easy to see that both maps that must be differentiable are differentiable, because the multiplication of two matrices boils down, coordinate-wise, to taking sums and products, and those are differentiable operations. When taking inverse one divides by the determinant of the matrix, but indeed for matrices in $GL_n(\mathbb{R})$ this determinant is non-zero and so this operation is differentiable. Also $GL_n(\mathbb{C})$ is a Lie group.

2. All submanifolds of $GL_n(\mathbb{R})$ and $GL_n(\mathbb{C})$ such as $SL_n(\mathbb{C})$ and $SU(n)$ that we gave as examples in the previous chapter are Lie groups because they are as well subgroups as submanifolds of $GL_n(\mathbb{R})$ or of $GL_n(\mathbb{C})$. Therefore the maps $(x, y) \mapsto xy$ and $x \mapsto x^{-1}$ are automatically differentiable.

3. $O(n)$, $SO(n)$ and $SU(n)$ are examples of compact Lie groups because the equations that the matrix coefficients must satisfy are given by $xx^t = x^t x = 1$, respectively $x\overline{x}^t = \overline{x}^t x = 1$ define a closed and bounded subset of $\mathbb{R}^{n^2}$ and by the theorem of Heine Borel such a set is compact.

4. The quotients of a Lie group by a subgroup is a Lie group if the subgroup is normal and closed. This is certainly not trivial, because it is a priori not clear how the quotient is a manifold; see the preceding chapter. In particular $SU(2)/\{1, -1\}$ is a Lie group. We will encounter this one again! For theorems on quotients of Lie groups see also [3], § 1.11.

2.3 Lie algebras

A Lie algebra is a vector space with on it an operation, the Lie bracket, that satisfies the following three conditions.

Definition 2.3. *A* Lie algebra *is a pair* $(L, [\cdot, \cdot])$ *(often denoted just by "L") with* L *an* $\mathbb{R}$*-vector space and a map* $[\cdot, \cdot] : L \times L \to L : (x, y) \mapsto [x, y]$*, the* Lie bracket*, that satisfies the following conditions:*

1. $[\cdot, \cdot]$ *is* $\mathbb{R}$*-bilinear,*

2. $[\cdot, \cdot]$ *is antisymmetric, that is,* $[x, y] = -[y, x]$*,*

3. For all $x, y, z \in L$ *we have* $[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$*, (the* Jacobi identity[3]*).*

A Lie group G gives in a natural way a Lie algebra via the tangent space $\mathfrak{g} := T_G(1)$. We already know that $\mathfrak{g}$ is a vector space, and now we just make a natural choice for the Lie bracket. For that we make use of the so-called adjoint representation, a specific example of a *representation*, a notion that awaits us only in the next chapter.

Definition 2.4. *Let* G *be a Lie group. The map* $\psi_g : G \to G : h \mapsto ghg^{-1}$ *is an automorphism of the Lie group* G*. We have*

$$D(\psi_g)_1 : T_G(1) = \mathfrak{g} \to T_G(\psi_g(1)) = T_G(1) = \mathfrak{g} : [\alpha] \mapsto [\psi_g \circ \alpha]$$

and one can check that this map is $\mathbb{R}$*-linear, with inverse* $D(\psi_{g^{-1}})_1$*, hence* $D(\psi_g)_1 \in Aut(\mathfrak{g})$*. This map is denoted* $Ad : G \to Aut(\mathfrak{g}) : g \mapsto D(\psi_g)_1$ *and is called the* adjoint representation of the Lie group G*.*

We already know that $\text{Aut}(\mathfrak{g}) = GL(\mathfrak{g})$ (automorphisms of $\mathfrak{g}$ as real vector space) and from example 1 of section 1.3 it follows that $T_{GL(\mathfrak{g})}(1) = \text{End}(\mathfrak{g})$, so the adjoint representation of G gives rise to the following map:

$$\text{ad} := D(\text{Ad})_1 : T_G(1) = \mathfrak{g} \to \text{End}(\mathfrak{g})$$

we use this map to define the Lie bracket on $\mathfrak{g}$ and the Lie algebra $\text{Lie}(G)$.

[3]More conceptually, this identity can be understood in 2 ways: (i) the map $[x, \cdot]\colon L \to L$ is a derivation, that is, it satisfies, for all y and z in L, the Leibniz rule for differentiation of a product $[x, [y, z]] = [[x, y], z] + [y, [x, z]]$, (ii) under the map $L \to \text{End}_{\mathbb{R}}(L)$ sending x to $[x, \cdot]$, the Lie bracket is compatible with the commutator in $\text{End}_{\mathbb{R}} L$, that is, for all x and y in L we have $[[x, y], \cdot] = [x, \cdot] \circ [y, \cdot] - [y, \cdot] \circ [x, \cdot]$

Definition 2.5. *Let G be a Lie group. The Lie algebra of G is the vector space $\mathfrak{g}$ together with the map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} : (g,h) \mapsto [g,h] = ad(g)(h)$*

This definition is not complete without a proof that this map $[\cdot,\cdot]$ satisfies the conditions for a Lie bracket and it is also possible to do this (see [3] theorem 1.1.4), but we will look more directly at the specific case of $G = GL_n(K)$ with $K = \mathbb{R}$ or $\mathbb{C}$ because it is then less abstract and more relevant for this thesis.

Theorem 2.6. *The Lie algebra of $GL_n(K)$ with $K = \mathbb{R}$ or $\mathbb{C}$ is $M_n(K)$ with the map $[\cdot,\cdot] : (X,Y) \mapsto XY - YX$, the commutator of X and Y, and this is a Lie bracket.*

Proof. From example 1 of section 1.3 and from the definition of derivative it follows that

$$D(\psi_g)_1 : T_{GL_n(K)}(1) \to T_{GL_n(K)}(1) : [1+tA] \mapsto [\psi_g \circ (1+tA)] = [g \cdot (1+tA) \cdot g^{-1}] = [1+t(g \cdot A \cdot g^{-1})]$$

where products "$\cdot$" are matrix multiplication. If one identifies $[1+tA]$ with A then one gets the map

$$D(\psi_g)_1 : M_n(K) \to M_n(K) : A \mapsto g \cdot A \cdot g^{-1}\,.$$

Then $\mathrm{Ad} : g \mapsto (A \mapsto gAg^{-1})$, hence $\mathrm{ad} : T_{GL_n(K)}(1) \to \mathrm{End}(M_n(K))$ and

$$\begin{aligned} \mathrm{ad}(X)(Y) &= D(\mathrm{Ad})_1([1+tX])(Y) \\ &= [\mathrm{Ad} \circ (1+tX)](Y) \\ &= [(1+tX)Y(1+tX)^{-1}] \\ &= [(1+tX)Y(1-tX)] \\ &= [Y + tXY - tYX + O(t^2)] \\ &= XY - YX \end{aligned}$$

This map is K-bilinear, antisymmetric and, by a simple computation, satisfies the Jacobi identity. □

With the help of the Lie algebra of a Lie group, one can often prove theorems on Lie groups. In the next chapter we will see an example related to representations. A what simpler example of how Lie algebras and Lie groups are related is the fact that a connected Lie group is commutative if and only if the Lie bracket on the Lie algebra is identically 0. This comes from the fact that ψ_g is the identity if and only if g commutes with all h in the Lie group.

The relation between Lie groups and Lie algebras is such that one can define a functor *Lie* from the category of Lie groups to that of Lie algebras. For this one sends a Lie group to its Lie algebra, and a morphism to the morphism that it induces. One then obtains an equivalence between the category of *connected and simply connected* Lie groups and the category of Lie algebras. For more information see [2].

3 Representations

3.1 Introduction

Representation theory is an important subject of mathematics because it enables mathematicians to transform group-theoretical questions into questions in linear algebra, an easier part of mathematics.

In representation theory one considers the group as a set of transformations of a mathematical object (more precisely: as a homomorphism to the group of automorphisms of the object), in our case this object is a vector space. In this chapter we define, among others, representations, what it means that a representation is irreducible, and we study the representations of $SU(2)$ and of $SO(3)$.

3.2 Representations

A representation of a group over a field $K = \mathbb{R}$ of $\mathbb{C}$ is a, usually finite dimensional, K-vector space with on it an action of the group.

Definition 3.1. *Let G be a group and $K = \mathbb{R}$ of $\mathbb{C}$. A* representation of G over K *is a pair (V, φ) (often denoted just as "φ" or "V") with V a K-vector space and $\varphi : G \to Aut(V) = GL(V)$ a group homomorphism.*

For G a Lie group a representation φ is a representation of the Lie group G *if φ is a morphism of Lie groups.*

A map $f : V \to V'$ with V, V' representations of G over K is a morphism of representations if f is K-linear and, for all $g \in G$ and $v \in V$, we have $f(gv) = gf(v)$.

Representations V and V' are called isomorpic *if there exists an isomorphism between them.*

An example of a representation is the adjoint representation of a Lie group G.

A representation is in fact a map from the group G to the set of linear transformations of a vector space V, that is, after a choice of a basis for V, $\varphi(G)$ is a group of matrices. This has the advantage that one can study properties of Lie groups via linear algebra, one of the work horses of mathematics. This also give insight in what it means that a representation is irreducible.

Definition 3.2. *Let (V, φ) be a representation of G over K. Then (V, φ) is called* irreducible *if V has precisely two subspaces that are invariant under the action of G: $\{0\}$ and V. This means that if for all $g \in G$ one has $\varphi(g)(V') \subset V'$ then geldt $V' = \{0\}$ or V, and $V \neq \{0\}$.*

3.3 Representations of $SU(2)$

Representations of $SU(2)$ play an important role in this thesis because they are relatively simple to construct, and because they are used a lot in physics. Let $\mathbb{C}[x, y]_d$ be the vector space of homogeneous polynomials in two variabels over $\mathbb{C}$ of degree d. An element of $\mathbb{C}[x, y]_d$ is of the the form

$$f(\mathbf{z}) = a_d x^d + a_{d-1} x^{d-1} y + \quad ... \quad + a_0 y^d$$

with $\mathbf{z} = \begin{pmatrix} x \\ y \end{pmatrix}$

Theorem 3.3. *The vector space $(\mathbb{C}[x, y]_d, \varphi)$ is a representation of $SU(2)$ over $\mathbb{C}$ if one defines, for $U \in SU(2)$ and $f(\mathbf{z}) \in \mathbb{C}[x, y]_d$,*

$$\varphi(U)(f)(\mathbf{z}) = f(U^{-1} \cdot \mathbf{z}) .$$

Proof. The map φ is a group homomorphism because $(\varphi(U_1 \cdot U_2)(f)(\mathbf{z}) = f(U_2^{-1} \cdot U_1^{-1} \cdot \mathbf{z}) = \varphi(U_2)(f)(U_1^{-1} \cdot \mathbf{z}) = \varphi(U_1)(\varphi(U_2)(f))(\mathbf{z}) = (\varphi(U_1) \circ \varphi(U_2))(f)(\mathbf{z})$. □

Theorem 3.4. *De representation* $(\mathbb{C}[x,y]_d, \varphi)$ *is irreducible.*[4]

Proof. We will use the Lie algebra $\mathfrak{su}(2)$ of $SU(2)$, as computed in example 4 of section 1.3. Let $V \subset \mathbb{C}[x,y]_d$ non-zero and invariant under the action of $SU(2)$. Now it holds that if V is invariant under the action of $SU(2)$ that it is invariant under the action of $\mathfrak{su}(2)$.

First we observe that $\mathfrak{su}(2)$ is spanned by 3 basis vectors:

$$I = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad K = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

and that $[I,J] = 2K,\ [J,K] = 2I,\ [K,I] = 2J$. Now the action $\varphi : SU(2) \to GL(\mathbb{C}[x,y]_d)$ defines an action $D(\varphi)_1 := \varphi' : \mathfrak{su}(2) \to End(\mathbb{C}[x,y]_d)$ of $\mathfrak{su}(2)$ on $\mathbb{C}[x,y]_d$ by

$$\varphi'(A) = [\varphi(1 + tA)]$$

in other words

$$\varphi(1 + tA)(f(\mathbf{z})) = f(\mathbf{z}) + t\varphi'(A)(f(\mathbf{z})) + O(t^2)$$

For I, J and K we have

$$\begin{aligned}
\varphi(1+tI)(x^a y^b) &= ((1-ti)x)^a((1+ti)y)^b = (x^a - atix^a)(y^b + btiy^b) \\
&= x^a y^b + ti(b-a)x^a y^b \\
\varphi(1+tJ)(x^a y^b) &= (x-ty)^a(tx+y)^b = (x^a - atx^{a-1}y)(y^b + btxy^{b-1}) \\
&= x^a y^b - atx^{a-1}yb + 1 + btx^{a+1}y^{b-1} \\
\varphi(1+tK)(x^a y^b) &= (x-tiy)^a(-tix+y)^b = (x^a - atix^{a-1}y)(y^b - btixy^{b-1}) \\
&= x^a y^b - atix^{a-1}y^{b+1} - btix^{a+1}y^{b-1}
\end{aligned}$$

modulo t^2. In other words:

$$\begin{aligned}
\varphi'(I)(x^a y^b) &= i(b-a)x^a y^b \\
\varphi'(J)(x^a y^b) &= bx^{a+1}y^{b-1} - ax^{a-1}y^{b+1} \\
\varphi'(K)(x^a y^b) &= -bix^{a+1}y^{b-1} - aix^{a-1}y^{b+1}
\end{aligned}$$

As $V \neq \{0\}$ there is a $v \neq 0$ in V; we take one such. Then $v = v_d x^d + v_{d-1}x^{d-1}y + ... + v_0 y^d$ with not all $v_i = 0$. As each term $\neq 0$ of v after applying $\varphi'(I)$ gets another coefficient, one can, by applying $\varphi'(I)$ repeatedly, and by taking suitable linear combinations, and up with a monomial. So, without loss of generality, $x^a y^b \in V$ for some (a,b). From the fact that V is invariant under $\mathfrak{su}(2)$ it then follows that $\varphi'(J)(x^a y^b)$, $\varphi'(K)(x^a y^b)$ and also

$$\frac{1}{2b}(\varphi'(J)(x^a y^b) + i\varphi'(K)(x^a y^b)) = x^{a+1}y^{b-1}$$

and

$$\frac{1}{2a}(i\varphi'(K)(x^a y^b) - \varphi'(J)(x^a y^b)) = x^{a-1}y^{b+1}$$

are elements of V. By applying this repeatedly we get that $x^i y^{d-i}$ is in V, hence that $V = \mathbb{C}[x,y]_d$. Hence V is irreducible. □

[4] Peter-Weyl's theorem 4.4 implies that these are *all* irreducible representations of $SU(2)$. One can also give a direct proof of that; see [2].

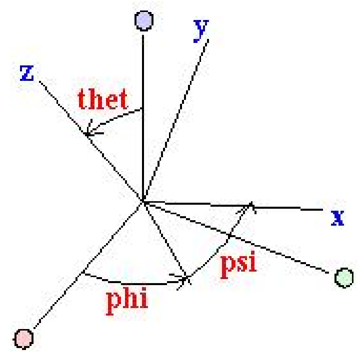

Figure 1: Euler angles φ, θ, ψ.

3.4 $SO(3)$

Before we look at the relation between $SU(2)$ and $SO(3)$ and the representations of $SO(3)$ we first look a bit better at the Lie group $SO(3)$, that plays such an important role in physics.

As a set $SO(3)$ consists of the orthogonal 3×3-matrices with determinant 1. Since the columns of an element A of $SO(3)$ are orthogonal, the image of the standard basis (e_1, e_2, e_3) under A is again an orthogonal basis. Moreover, as $\det A = 1$, it has the same orientation and one can see A as a rotation of $\mathbb{R}^3$.

Following Euler, a rotation of $\mathbb{R}^3$ can be given by 3 *Euler angles* φ, θ and ψ, see section 9.6 of [4], as follows. The rotation with angles φ, θ and ψ, $(0 \le \varphi < 2\pi)$, $(0 \le \theta < \pi)$, $(0 \le \psi < 2\pi)$ is obtained by first rotating about the z-axis over the angle φ, then about the y-axis over the angle θ, and finally again about the z-axis over the angle ψ. This gives the following surjective map from $\mathbb{R}^3$ to $SO(3)$:

$$(\varphi, \theta, \psi) \mapsto \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This map is not injective at all, not only because it is 2π-periodic in the three arguments, but also because for $\theta = 0$ it gives the rotation over the angle $\varphi + \psi$ about the z-axis. It is easy to see that the map is continuous. This easily gives the following lemma.

Lemma 3.5. *The Lie group $SO(3)$ is connected.*

Proof. The map above is continuous and surjective, and its source is connected. □

3.5 Quaternions and representations of $SO(3)$

There is an isomorphism of Lie groups $SU(2)/\{1, -1\} \cong SO(3)$, hence $SU(2)$ is a double cover of $SO(3)$. To see this, it is convenient to introduce quaternions, an $\mathbb{R}$-algebra discovered and applied to mechanics in 3 dimensions by Hamilton.

Definition 3.6. *The* quaternion algebra *is the sub-$\mathbb{R}$-algebra $\mathbb{H}$ of $M_2(\mathbb{C})$ consisting of the matrices* $\begin{bmatrix} a & -\bar{b} \\ b & \bar{a} \end{bmatrix}$ *with $a, b \in \mathbb{C}$. One can also consider $\mathbb{H}$ as the $\mathbb{R}$-vector space $\mathbb{R}1 \oplus \mathbb{R}I \oplus \mathbb{R}J \oplus \mathbb{R}K \subset M_2(\mathbb{C})$ with*

$$1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ I = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix},\ J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},\ K = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

and as multiplication the multiplication of matrices.

It is not hard to prove that $\mathbb{H}$ is indeed a sub-$\mathbb{R}$-algebra and that both definitions are equivalent. The center $Z(\mathbb{H})$ is $\mathbb{R}1$: every element of $M_2(\mathbb{C})$ that commutes with I is diagonal and that it moreover commutes with J means that it is a scalar.

Definition 3.7. *Let, for x in $\mathbb{H}$, $x^* := \overline{x}^t$ in $\mathbb{H}$, and let*

$$\begin{aligned} N(x) &:= x^*x = xx^* = \det(x) \in \mathbb{R}1 \\ tr(x) &:= x + x^* \in \mathbb{R}1 \end{aligned}$$

We define an inner product on $\mathbb{H}$ by

$$\langle x, y \rangle := (x^*y + y^*x)/2 = tr(x^*y)/2\,.$$

The identities on the right of the ":=" must be checked but this is easy. Note that we have

$$\{x \in \mathbb{H} |\ \mathrm{N}(x) = 1\} = SU(2) \subset \mathbb{H}^*$$

and

$$V := \mathbb{R}I \oplus \mathbb{R}J \oplus \mathbb{R}K = \mathfrak{su}(2) \subset \mathbb{H}\,.$$

Consider now the maps $l_x : \mathbb{H} \to \mathbb{H} : y \mapsto xy$ and $r_x : \mathbb{H} \to \mathbb{H} : y \mapsto yx$.

Theorem 3.8. *The maps l_x and r_x zijn orthogonal iff $N(x) = 1$ and $det(l_x) = det(r_x) = (N(x))^2$.*

The proof (left to the reader) of the second part uses the characteristic polynomial of l_x and r_x, see § 5.10 of [2] and for the lineare algebra [6].

Now consider, for a $x \in \mathbb{H}^*$, the function $c_x : \mathbb{H} \to \mathbb{H} : y \mapsto xyx^{-1}$. This function is invertible and on $\mathbb{R}1$ it is the identity, hence it gives, by restriction, a function $c_x : V \to V$ with $V = \mathbb{R}I \oplus \mathbb{R}J \oplus \mathbb{R}K \cong \mathbb{R}^3$. The map $c : x \mapsto c_x$ is a homomorphism from $\mathbb{H}^*$ to $\mathrm{GL}(V)$. For $x \in SU(2)$ this is $c(x)$ and because $c_x = r_{x^{-1}} \circ l_x$ and theorem 3.8 it is an orthogonal linear transformation with determinant 1. Hence by restriction to $SU(2)$ one obtains the map

$$c : SU(2) \to SO(V)\ (\cong SO(3)) : x \mapsto c_x = (y \mapsto xyx^{-1})\,.$$

The kernel of this group homomorphism is the intersection $SU(2) \cap Z(\mathbb{H}) = \{1, -1\}$.

Theorem 3.9. *We have $im(c) = SO(V)$.*

Proof. First we note that c is continuous and differentiable (and so a morphism of Lie groups), because for every choice (v_1, v_2, v_3) of a basis of V the matrix coefficients of

$$c(x) := \begin{bmatrix} \langle xv_1x^{-1}, v_1\rangle & \langle xv_2x^{-1}, v_1\rangle & \langle xv_3x^{-1}, v_1\rangle \\ \langle xv_1x^{-1}, v_2\rangle & \langle xv_2x^{-1}, v_2\rangle & \langle xv_3x^{-1}, v_2\rangle \\ \langle xv_1x^{-1}, v_3\rangle & \langle xv_2x^{-1}, v_3\rangle & \langle xv_3x^{-1}, v_3\rangle \end{bmatrix}$$

are polynomials in $\mathbb{C}[x_{(1,1)}, ..., x_{(2,2)}, \overline{x_{(1,1)}}, ..., \overline{x_{(2,2)}}]$. As $x_{(1,1)} = \overline{x_{(2,2)}}$ and $x_{(1,2)} = -\overline{x_{(2,1)}}$ these are polynomials in $\mathbb{C}[x_{(1,1)}, ..., x_{(2,2)}]$ and so are continuous and differentiable.

If we can show that $\mathrm{im}(c)$ ($\neq \{0\}$) is both open and gesloten then by lemma 3.5 $\mathrm{im}(c) = SO(V)$.

Suppose that the derivative $D(c)_1 : T_{SU(2)}(1) = \mathfrak{su}(2) \to T_{SO(V)}(1)$, a morphism of Lie algebras, is surjective, then the Impliciete Function Theorem on pagina 6 of [1] gives that c is a homeomorphism from a neighborhood of $1 \in SU(2)$ to a neighborhood $1 \in SO(V)$, hence in particular

$$\exists U \subset SU(2) \text{ open, } 1 \in U : c(U) \text{ is open.}$$

Then c is a homomorphism hence $c(gU) = c(g)c(U) \subset \mathrm{im}(c)$, and by lemma 2.2 $l_{c(g)^{-1}}$ is continuous, hence $c(g)c(U)$ is open. As $c(g) \in c(g)c(U)$ is

$$\mathrm{im}(c) = \bigcup_{g \in SU(2)} c(g)c(U)$$

and thus is open.

We know that $\dim(\mathfrak{su}(2)) = \dim(T_{SO(V)}(1)) = 3$, hence $D(c)_1$ is surjective iff $D(c)_1$ is injective.

Let $A \in \mathfrak{su}(2)$ and $v \in V$. Then

$$v \mapsto c(1 + tA)(v) = (1 + tA)v(1 - tA) = v + t(Av - vA) \pmod{t^2}$$

and from this it follows that $D(c)_1(A) = v \mapsto Av - vA$. For $A \in \ker(D(c)_1)$ this means, as we have already seen with the quaternions, that $A \in \mathbb{R}1 \cap \mathfrak{su}(2)$. Then $A = A^* = -A$, that is, $A = 0$, so $D(c)_1$ is injective, surjective and $\mathrm{im}(c)$ is open.

Now $\mathrm{im}(c)$ is a subgroup of $SO(V)$, so $SO(V)$ is the disjoint union, over the $g' \in SO(V)$, of the $g' \cdot \mathrm{im}(c)$, that are all open by lemma 2.2. In particular the complement of $\mathrm{im}(c)$ is and so $\mathrm{im}(c)$ closed.

We have seen that $\mathrm{im}(c)$ is open as well as closed, and so the theorem is proved.

Let us also give a sketch of a **second proof**, that is more direct and moreover gives, for each element in $SO(V)$, its 2 preimages in $SU(2)$. Let g be in $SO(V)$. If $g = \mathrm{id}$, then the 2 preimages are id and $-\mathrm{id}$ in $SU(2)$. So now assume that $g \neq \mathrm{id}$. Then g is a rotation about a unique line in V, and there is $x \in V$, unique un to sign, such that $\|x\| = 1$ and g is a rotation about $\mathbb{R}x$ over a non-zero angle $\varphi \in (-\pi, \pi]$ (here we use the orientation of V for which (I, J, K) is an oriented basis). Then the 2 preimages of g are $\pm(\cos(\varphi/2) + \sin(\varphi/2)x)$.[5] □

So we have a morphism of Lie groups $c : SU(2) \to SO(V)$ such that $SU(2)/\{1, -1\} \cong SO(V)$ as Lie groups. A choice of an orthonormal basis of V, for example $(v_1 = I, v_2 = J, v_3 = K)$ then gives an isomorphism of Lie groups $SO(V) \to SO(3)$. This isomorphism is not unique: another choice (for example $(v_1 = K, v_2 = I, v_3 = J)$) gives another isomorphism. This isomorphism induces an isomorphism of Lie algebras $\mathfrak{so}(V) \to \mathfrak{so}(3)$ and since we already saw that $D(c)_1$ is a bijective morphism of Lie algebras (hence an isomorphism), we conclude that $\mathfrak{su}(2) \cong \mathfrak{so}(3)$. Also this isomorphism depends on the choice of the basis for V.

The irreducible representations of $SO(3)$ can now be seen as the irreducible representations of $SU(2)$ on which the subgroup acts $\{1, -1\}$ trivially. These are precisely the $V_i := \mathbb{C}[x, y]_{2i}$, $(i \geq 0)$ with

$$\rho_i : SO(3) \to GL(V_i) : SO(3) \to SU(2)/\{1, -1\} \to GL(V_i)$$

[5] To prove this, use an orthonormal oriented $\mathbb{R}$-basis (x, y, z) of V, show that $x^2 = y^2 = z^2 = xyz = -1$, giving an automorphism $\mathbb{H}$ sending (i, j, k) to (x, y, z), and do a direct computation of conjugation by $a + bi$.

but this map is not unique, it depends on the choice of basis for V. But if we choose another basis we get isomorphic representations, so in this sense the choice of basis does not matter. The freedom of choice of basis for V is important and we will use it in the next section.

4 The Peter-Weyl theorem

4.1 Introduction

Now we arrive at the central mathematical result of this thesis: the theorem of Peter-Weyl. First it was proved by Hermann Weyl for compact Lie groups, and then for general compact topological groups.

One can see this as a generalisation of Fourier theory of periodic functions to compact Lie groups: as Fourier theory gives a basis for $L^2(S^1)$ (every function in $L^2(S^1)$, that is, every periodic function on $\mathbb{R}$, is written as Fourier series), the Peter-Weyl theorem gives a basis for $L^2(G)$. Here G is (in our case) a compact Lie group contained in some $\mathrm{GL}_n(\mathbb{C})$ and the complexe Hilbertspace $L^2(G)$ is the (completion of the) set of square integrable complex valued functions on G, more about which in a moment. For $f_1, f_2 \in L^2(G)$ one has by definition that $\int_G f_1\overline{f_2}\mu$ is defined. Here μ is a *left-invariant volume form* also called *Haar measure*. More about volume forms can be found in [1] chapters 3 and 5 and in [2] chapter 11. As I onely barely touch on this subject I will not elaborate. As G is compact, one can normalise μ (scale it with a factor in $\mathbb{R}^{\times,>0}$) uniquely, such that $\int_G \mu = 1$. Another consequence of G being compact is that μ is also right-invariant. Here we assume that G is a sub-Lie group of $\mathrm{GL}_n(\mathbb{C})$ because this is all we need and because the proof is easier (the Haar measure is easier to understand). Now first something about Hilbert spaces.

4.2 Hilbert spaces

As already mentioned, $L^2(G)$ is an example of a *Hilbert space*. As such spaces occur in the Peter-Weyl theorem I shall now give a concise definition. For details see [8] chapter 6.

Definition 4.1. *An* inner product space *over* $\mathbb{C}$ *is a* $\mathbb{C}$*-vector space* V *with a positive Hermitian sesquilinear form, the* inner product*. That is, the inner product* $\langle\cdot,\cdot\rangle$ *satisfies:*

1. $\forall x \in V: \ \langle x, x\rangle \geq 0$ *(positive),*
2. $\forall x, y \in V: \ \langle x, y\rangle = \overline{\langle y, x\rangle}$ *(Hermitian),*
3. $\forall x, y, z \in V \ \forall a \in \mathbb{C}: \ \begin{matrix} \langle ax + y, z\rangle = a\langle x, z\rangle + \langle y, z\rangle \\ \langle x, ay + z\rangle = \overline{a}\langle x, y\rangle + \langle x, z\rangle \end{matrix}$ *(sesquilinear).*

A Hilbert space is a special kind of inner product space.

Definition 4.2. *An inner product space* V *is a* Hilbert space *if it is complete for the norm*

$$\|x\| = \langle x, x\rangle^{1/2},$$

that is, every Cauchy sequence in V *sequence converges.*

The normed complex vector space $L^2(G)$ is defined as the completion of the $\mathbb{C}$-vector space $C(G)$ of square integrable continuous functions $G \to \mathbb{C}$. As the inner product on $C(G)$ is defined as $\langle f, g\rangle = \int_G f\overline{g}\mu$ one sees easily that $C(G)$ is an inner product space (not complete unless G is of dimension zero) and that $L^2(G)$ is a Hilbert space.

It will turn out useful (even necessary) to use, for Hilbert spaces, another definition of basis than for vector spaces, see also [8] §6.2.

Definition 4.3. *An orthonormal subset S of a Hilbert space V is called* complete *when $S^\perp := \{x \in V : \forall_{s\in S}\, \langle x, s\rangle = 0\}$ is zero. A complete orthonormal subset of a Hilbert space V is called a* Hilbert basis *of V.*

Loosely speaking this means that, when taking linear combinations of the elements of a Hilbert basis, infinitely many coefficients are allowed to be non-zero, as long as the series (or even integral) converges. In what follows we will mean, whenever necessary, "Hilbert basis" when writing "basis".

The fact that $C(G)$ is dense in $L^2(G)$ means that every $f \in L^2(G)$ is the limit of a sequence $(f_i)_{i\geq 0}$ with $f_i \in C(G)$. This is equivalent with the fact that f can be written as a sum of elementen in $C(G)$: $f = f_0 + \sum_{i=0}^{\infty}(f_{i+1} - f_i)$.

The vector space $C(G)$ is a representation of the Lie group G for the action by right translations $(gf)(x) := f(xg)$ for $f \in C(G)$ and $g \in G$. As elements of $L^2(G)$ cannot be evaluated in elements of g, we cannot just say, for $f \in L^2(G)$, that $(gf)(x) := f(xg)$, but instead we can say that the action of $g \in G$ on $C(G)$ is linear and preserves the inner product, hence extends uniquely to an automorphism of $L^2(G)$, and we say that it is still by right translations.

4.3 The Peter-Weyl theorem

Theorem 4.4. *Let $G \subset GL_n(\mathbb{C})$ for an $n \in \mathbb{N}$ be a compact Lie group, with Haar measure μ and $\{(V_i, \rho_i) | i \in I\}$ a set of representatives of the isomorphism classes of irreducible finite dimensional representations of G. Then the maps $End_{\mathbb{C}}(V_i) \to L^2(G)$ given by $m \mapsto (g \mapsto m(g) := tr(\rho_i(g) \cdot m))$ together give an isomorphism of G-representations*

$$L^2(G) \cong \widehat{\bigoplus}_{i\in I} End_{\mathbb{C}}(V_i)\,.$$

Here $\widehat{\oplus}$ is a Hilbert direct sum*, a sum in which elements are convergent series.*

Let $\langle\cdot,\cdot\rangle$ be a G-invariant inner product on V_i and $v_i := (v_{i,1}, ..., v_{i,dim(V_i)})$ an orthonormal basis of V_i. Then the $\sqrt{dim(V_i)} f_{i,j,k}$ with $f_{i,j,k}$ in $C(G)$ given by

$$f_{i,j,k}(g) := E_{k,j}(g) = tr(\rho_i(g) \cdot E_{k,j}) = \rho_i(g)_{j,k} = \langle \rho_i(g) v_{i,k}, v_{i,j}\rangle\,,$$

where $E_{k,j}$ is the matrix with (k, j)-coefficient 1 and further only zeros and $\rho_i(g)_{j,k}$ the (j, k)-coefficient of the matrix of $\rho_i(g)$ with respect to the basis v_i, form an orthonormal basis of $L^2(G)$.

Proof. First we note that $\mathrm{End}_{\mathbb{C}}(V_i)$ is a representation via left translations: $gm = \rho_i(g) \cdot m$ and $L^2(G)$ via right translations: $(gf)(x) = f(xg)$. The map in the theorem is $\mathbb{C}$-linear and we have

$$(gm)(x) = \mathrm{tr}(\rho_i(x) \cdot (\rho_i(g) \cdot m)) = \mathrm{tr}(\rho_i(xg) \cdot m) = m(xg) = g(m(x))$$

so $\mathrm{End}_{\mathbb{C}}(V_i) \to L^2(G)$ and $\widehat{\bigoplus}_{i\in I}\mathrm{End}_{\mathbb{C}}(V_i) \to L^2(G)$ are morphisms of representations.

We can show that this last map is an isomorphism by showing that it is injective and surjective. Injectivity follows directly from the fact that the $\sqrt{\dim(V_i)}f_{i,j,k}$ are orthonormal by *Schur's orthogonality relations*, to be proved now. Let V and V' be two irreducible representations of G and $u, v \in V$ and $u', v' \in V'$. Then we have

$$\begin{aligned} \int_{g\in G} \langle gu, v\rangle \overline{\langle gu', v'\rangle}\mu &= 0 \text{ if } V \text{ and } V' \text{ are not isomorphic} \\ &= \frac{\langle u, u'\rangle \overline{\langle v, v'\rangle}}{\dim(V)} \text{ if } V = V'. \end{aligned}$$

Proof. Let $f : V \to V'$ be an arbitrary linear map. Then, with

$$F(x) := \int_{g\in G} g(f(g^{-1}x))\mu\,,$$

F is a morphism of representations. Indeed we have $F(hx) = \int_{g\in G} g(f(g^{-1}hx))\mu$ and because for all functions Q we have $\int_{g\in G} Q(g)\mu = \int_{g\in G} Q(hg)\mu$ for $h \in G$ we have

$$F(hx) = \int_{g\in G} (hg)(f((hg)^{-1}hx))\mu = \int_{g\in G} h(g(f(g^{-1}x)))\mu = hF(x)\,.$$

Suppose now that V and V' are not isomorphic, then $F = 0$ because V and V' are irreducible. Let $f : x \mapsto \langle x, u\rangle u'$. Then we have:

$$\begin{aligned} 0 &= \langle v', F(v)\rangle = \langle v', \int_{g\in G} g(f(g^{-1}v))\mu\rangle \\ &= \int_{g\in G} \langle v', g(f(g^{-1}v))\rangle\mu = \int_{g\in G} \langle g^{-1}v', f(g^{-1}v)\rangle\mu \\ &= \int_{g\in G} \langle g^{-1}v', \langle g^{-1}v, u\rangle u'\rangle\mu = \int_{g\in G} \overline{\langle g^{-1}v, u\rangle}\langle g^{-1}v', u'\rangle\mu \\ &= \int_{g\in G} \langle gu, v\rangle \overline{\langle gu', v'\rangle}\mu\,. \end{aligned}$$

Suppose now that $V = V'$, then $F = \lambda 1$ for a $\lambda \in \mathbb{C}$. W take f as above. Then we have:

$$\begin{aligned} \lambda \cdot \dim(V) &= \operatorname{tr}(F) = \sum_{i=1}^{\dim(V)} \langle F(v_i), v_i\rangle \\ &= \int_{g\in G} \sum_{i=1}^{\dim(V)} \langle g(f(g^{-1}v_i), v_i\rangle\mu = \int_{g\in G} \operatorname{tr}(g \circ f \circ g^{-1})\mu \\ &= \int_{g\in G} \operatorname{tr}(f)\mu = \operatorname{tr}(f) = \sum_{i=1}^{\dim(V)} \langle f(v_i), v_i\rangle \\ &= \sum_{i=1}^{\dim(V)} \langle\langle v_i, u\rangle u', v_i\rangle = \sum_{i=1}^{\dim(V)} \langle u', v_i\rangle\langle v_i, u\rangle = \langle u', u\rangle \end{aligned}$$

Hence:

$$\frac{\langle u, u'\rangle\overline{\langle v, v'\rangle}}{\dim(V)} = \overline{\frac{\langle u', u\rangle}{\dim(V)}}\langle v, v'\rangle = \overline{\lambda\langle v, v'\rangle} = \langle v', F(v)\rangle$$
$$= \ldots = \int_{g\in G}\langle gu, v\rangle\overline{\langle gu', v'\rangle}\mu$$

This proves Schur's orthogonality relations, and the fact that the are $\sqrt{\dim(V_i)}f_{i,j,k}$ orthonormal.

The proof of the surjectivity is more difficult and will not be given completely. An important step is to show that vector space $E \subset L^2(G)$ spanned by the $f_{i,j,k}$ is closed under multiplication and under complex conjugation, for which properties of dual vector spaces and tensor products of vector spaces are used. For the proof see [2] section 12.4 and propositie 12.5 and for dual vector spaces and for tensor products see [6] chapter 2 and appendix 4.

So E is a sub-$\mathbb{C}$-algebra closed under complex conjugation. As $G \subset \mathrm{GL}_n(\mathbb{C})$, the elements of G are matrices. Consider now the functions $x_{p,q} \in L^2(G)$ with $1 \leq p, q \leq n$ such that $x_{p,q} : G \to \mathbb{C} : g \mapsto (g)_{p,q}$, that is, $x_{p,q}$ is the (p, q)-coordinate function. If we can prove that $x_{p,q} \in E$ for all (p, q) then it follows from the *Stone-Weierstrass theorem* that $\overline{E} = L^2(G)$, with $\overline{E}$ the closure of E. Indeed, the Stone Weierstrass theorem says that a continuous complex function f on a compact subset $C \subset \mathbb{C}^{n^2}$ is the limit, for the sup norm, of sequence of polynomials in the n^2 coordinates and their conjugates, see also [8], §1.6. For this it follows that E is dense in $L^2(G)$, because the subspace of continuous functions in $L^2(G)$ is dense in $L^2(G)$, and, on the space of continuous functions on the compact space G, the sup norm is stronger than the L^2-norm.

We show that $x_{p,q} \in E$ by noting that $\mathbb{C}^n$ together with the embedding $\rho : G \to GL_n(\mathbb{C})$ is a representation of G and hence is a direct sum of irreducible representations. Hence there is an isomorphism of representations of G:

$$f\colon \bigoplus_{i\in I} V_i^{m_i} \xrightarrow{\sim} \mathbb{C}^n\,,$$

with $m_i \in \mathbb{N}$ all but finitely many equal to 0 (recall that the $(V_i)_{i\in I}$ are the irreducible representations of G). On both sides of this isomorphism we have a basis: the standard basis of $\mathbb{C}^n$, and, for each $i \in I$, the chosen basis v_i. The matrix of f with respect to these two bases expresses the $x_{p,q}$ as linear combinations of the $f_{i,j,k}$, and so the $x_{p,q}$ are in E. □

4.4 Right or left?

In the previous section, we have viewed $C(G)$ as a representation of G via the action of G on itself by *right*-translations:

$$(g{\cdot}f)(x) = f(xg)\,.$$

But suppose that we want to use left-translations:

$$(g\bullet f)(x) = f(g^{-1}x)\,,$$

does that change the result of theorem 4.4? The answer is "no", because the map

$$\iota\colon G \to G\,, \quad x \mapsto x^{-1}\,,$$

is an isomorphism from G, with G acting by right-translations, to G with G action by left-translations:

$$\iota(xg) = (xg)^{-1} = g^{-1}x^{-1} = g^{-1}\iota(x)\,.$$

Then the map

$$\iota^*\colon C(G) \to C(G)\,, \quad f \mapsto \iota^*(f)\,, \quad \text{with } (\iota^* f)(x) = f(\iota(x)) = f(x^{-1})$$

is an isomorphism from $C(G)$, with G acting via left-translations, to $C(G)$ with G acting by right-translations:

$$(\iota^*(g\bullet f))(x) = (g\bullet f))(x^{-1}) = f(g^{-1}x^{-1}) = f((xg)^{-1}) = f(\iota(xg)) = (\iota^* f)(xg) = (g\cdot(\iota^* f))(x)\,.$$

4.5 Inner Product on $\mathbb{C}[x, y]_d$

In the next part we will apply the Peter-Weyl theorem to the irreducible representations of $G = SU(2)$, and the $V_d = \mathbb{C}[x, y]_d$. The Peter-Weyl theorem requires that we choose a G-invariant inner product on V_d. We note that $SU(2)$ leaves the standard inner product on $\mathbb{C}^2$ invariant. We consider the elements of V_d as functions $\mathbb{C}^2 \to \mathbb{C}$ and choose as inner product:

$$\langle f, g\rangle := \int_{\substack{v\in\mathbb{C}^2 \\ \|v\|=1}} f\overline{g}\,\mu_{SU(2)}$$

with $\mu_{SU(2)}$ the usual volume form on $S^3 \subset \mathbb{R}^4 = \mathbb{C}^2$ (, see further down, it is $SO(4)$-invariant, hence $SU(2)$-invariant) and $f, g \in V_d$. This inner product is automatically $SU(2)$-invariant.

Theorem 4.5. *Let $d \in \mathbb{N}$. Then $(x^d, x^{d-1}y, \ldots, y^d)$ is an orthogonal basis of V_d.*

Proof. We must show that for $x^a y^b$ and $x^{a'} y^{b'}$ in V_d with $a \neq a'$ (and therefore $b \neq b'$), we have $\langle x^a y^b, x^{a'} y^{b'}\rangle = 0$. We do this by using the action of the diagonal subgroup of G. For $t \in \mathbb{R}$ let $\lambda(t) := e^{it}$, and let

$$\Lambda(t) := \begin{bmatrix} \lambda(t) & 0 \\ 0 & \lambda(t)^{-1} \end{bmatrix} \quad \text{in } SU(2).$$

Then we have, for all $t \in \mathbb{R}$:

$$\begin{aligned} \Lambda(t)x^a y^b &= (\lambda(t)x)^a(\lambda(t)^{-1}y)^b = \lambda(t)^{a-b}x^a y^b \\ \Lambda(t)x^{a'} y^{b'} &= (\lambda(t)x)^{a'}(\lambda(t)^{-1}y)^{b'} = \lambda(t)^{a'-b'}x^{a'} y^{b'} \end{aligned}$$

As $a - b \neq a' - b'$ it holds for sufficiently general t that $\lambda(t)^{a-b} \neq \lambda(t)^{a'-b'}$, hence $x^a y^b$ and $x^{a'} y^{b'}$ are eigenfunctions of $\Lambda(t)$ with distinct eigenvalues. The we have:

$$\langle x^a y^b, x^{a'} y^{b'}\rangle = \langle \Lambda(t)x^a y^b, \Lambda(t)x^{a'} y^{b'}\rangle = \lambda(t)^{a-b}\lambda(t)^{-1(a'-b')}\langle x^a y^b, x^{a'} y^{b'}\rangle$$

and because $\lambda(t)^{a-b}\lambda(t)^{-1(a'-b')} \neq 1$ we get that $\langle x^a y^b, x^{a'} y^{b'}\rangle = 0$. □

Now that we know that $(x^d, ..., y^d)$ is orthogonal, all that is left to do is to scale them, with factors say $\lambda_{d,j}$, such that $(\lambda_{d,d}x^d, \lambda_{d,d-1}x^{d-1}y, ..., \lambda_{d,0}y^d)$ is orthonormal. For this, we must compute the $\langle x^j y^{d-j}, x^j y^{d-j}\rangle$.

For working with $S^3 = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^2 + |z_2|^2 = 1\}$ we use the *Hopf coordinates*:

$$\begin{aligned} z_1 &= e^{i\xi_1} \sin \eta \\ z_2 &= e^{i\xi_2} \cos \eta \end{aligned}$$

where ξ_1 and ξ_2 range over $[0, 2\pi]$ and η over $[0, \pi/2]$. The *volume form* is then

$$dV = \sin \eta \cos \eta \, |d\eta \wedge d\xi_1 \wedge d\xi_2|$$

Then, as Maple tells us,

$$\begin{aligned} \langle x^j y^{d-j}, x^j y^{d-j} \rangle &= \int_{\xi_2=0}^{2\pi} \int_{\xi_1=0}^{2\pi} \int_{\eta=0}^{1/2\pi} \sin \eta^{2j+1} \cos \eta^{2(d-j)+1} \, d\eta \, d\xi_1 \, d\xi_2 \\ &= 4\pi^2 \frac{\Gamma(d-j+1)\Gamma(j+1)}{2\Gamma(d+2)} \\ &= 4\pi^2 \frac{1}{2(d+1)} \frac{j!(d-j)!}{d!} \\ &= \frac{2\pi^2}{d+1} \left(\begin{array}{c} d \\ j \end{array} \right)^{-1} \end{aligned}$$

So we scale $x^j y^{d-j}$ with a factor

$$\lambda_{d,j} := \frac{1}{\pi} \sqrt{\left(\begin{array}{c} d \\ j \end{array} \right) \frac{d+1}{2}},$$

and then $(\lambda_{d,d} x^d, \ldots, \lambda_{d,0} y^d)$ is an orthonormal basis for $\mathbb{C}[x, y]_d$.

5 Spherical harmonic functions

5.1 Introduction

In the previous sections we have treated, in as compact a way as possible, the representation theory of compact Lie groups, culminating in the Peter-Weyl theorem. In addition, we have determined the irreducible representations of $SU(2)$ and of $SO(3)$, using the double cover of $SO(3)$ by $SU(2)$. Now we apply all this theory to realise the goal of this thesis: determining the usual spherical-harmonic functions that form a basis of the Hilbert space $L^2(S^2)$, the space of square integrable functions on the sphere S^2.

Recall from definition 4.3 that to give a Hilbert basis for $L(S^2)$ it suffices to give an orthonormal collection of functions in $C(S^2)$ that is complete. We will do this by relating S^2 with $SO(3)$, more precisely, we will view S^2 as a quotient of $SO(3)$.

5.2 Determination of a Hilbert basis for $L^2(S^2)$

From section 3.4 it follows that $G = SO(3)$ acts in a natural way on S^2 via matrix multiplication, and that an element $g \in G$ can be seen as a rotation. As a base point in S^2 we take $N = (0, 0, 1)$

(N for "north pole"). The stabiliser of N is the group of rotations about the z-axis, and is isomorphic with $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ via the map:

$$\varphi \mapsto \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This gives us a surjective differentiable map with surjective derivative

$$F : G \to S^2\,, \quad g \mapsto g{\cdot}N\,,$$

with $F^{-1}(g{\cdot}N) = gS^1$; it is a quotient for the action of S^1 on G by right-translations. The map F is equivariant for the action by G on G by left-translations and on S^2 by matrix multiplication: $F(g_1g_2) = (g_1g_2){\cdot}N = g_1{\cdot}(g_2{\cdot}N) = g_1{\cdot}F(g_2)$. As the Haar measure μ_G with volume 1 on G is invariant under the right-translations by S^1, it induces, via F, the G-invariant measure μ_{S^2} on S^2. This means that for f_1 and f_2 in $C(S^2)$ we have

$$\int_{S^2} f_1\overline{f_2}\mu_{S^2} = \int_G (f_1 \circ F)\overline{(f_2 \circ F)}\mu_G\,.$$

Hence, with

$$C(G)^{S^1} := \{f \in C(G) \mid \forall g \in G\,,\ \forall \varphi \in S^1,\ f(g\varphi) = f(g)\}\,,$$

the map

$$F^*\colon C(S^2) \to C(G) : f \mapsto f \circ F\,.$$

preserves inner products, has image $C(G)^{S^1}$, and gives an isomorphism of inner product spaces from $C(S^2)$ to $C(G)^{S^1}$. So, we get a Hilbert basis for $L^2(S^2)$ by applying the inverse of this isomorphism to an orthonormal complete subset of $C(G)^{S^1}$. And to obtain such a subset, we will apply the standard orthogonal projection $C(G) \to C(G)^{S^1}$, to be defined in a moment, to a suitably chosen complete orthonormal subset of $C(G)$ (all elements in it are eigenvectors for S^1).

The *averaging map* that sends f in $C(G)$ to the function $F_*(f)\colon G \to \mathbb{C}$, $g \mapsto \int_{\varphi\in S^1} f(g\varphi)\mu_{S^1}$. One checks, using that integration is continuous for the sup-norm, that $F_*(f)$ is in $C(G)$, and in fact in $C(G)^{S^1}$. Moreover we have that, for $f \in C(G)^{S^1}$, $F_*(f) = f$, hence $F_*^2 = F_*$, that is, F_* is an idempotent. It is a nice exercise to show that F_* is self-adjoint: for all f_1 and f_2 in $C(G)$ we have $\langle F_*(f_1), f_2\rangle = \langle f_1, F_*(f_2)\rangle$. From this it follows that $C(G) = \ker(F_*) \oplus C(G)^{S^1}$, and that the 2 summands are orthogonal to each other. So F_* is the orthogonal projection from $C(G)$. Anyway, if now S is a complete subset of $C(G)$, then we claim that $F_*(S)$ is a complete subset of $C(G)^{S^1}$. Here is a proof. Suppose that f in $C(G)^{S^1}$ is orthogonal to $F_*(S)$. Then for all s in S we have $0 = \langle f, F_*s\rangle = \langle F_*f, s\rangle = \langle f, s\rangle$, hence $f = 0$ by the completeness of S.

Now is the moment that we produce our complete orthonormal set in $C(G)$ that has the desired propery: every element is an eigenvector for S^1. The irreducible representations (up to isomorphism) of G are the representations $V_l = \mathbb{C}[x,y]_{2l}$, for $l \geq 0$, of $SU(2)$, viewed as representations of G via the morphism $SU(2) \to SO(3)$ obtained from the action by conjugation of $SU(2)$ on V (the quaternions of trace zero) via the basis (J, K, I) of V. The reason for this choice of basis is that the $I = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$, as element of $SU(2)$, is diagonal, so that our orthonormal basis $v_{l,j} = \lambda_{2l,j}x^jy^{2l-j}$, with $0 \leq j \leq 2l$, of V_l from section 4.5, consists of eigenvectors for the diagonal subgroup of $SU(2)$. The conjugation by this subgroup fixes the element I of V, hence we want this subgroup to be sent to the stabiliser of $N=(0,0,1)$ in $S^2\subset\mathbb{R}^3$.

The theorem 4.4 of Peter-Weyl says that the $\sqrt{2l+1}f_{l,j,k}$ with $l\geq 0$, $0\leq j,k\leq 2l$, form a complete orthonormal set of $C(G)$, where, for all $g\in G$, $f_{l,j,k}(g)=\langle\rho_l(g)v_{l,k},v_{l,j}\rangle$.

We compute how $\varphi\in S^1\subset G$ acts on the $f_{l,j,k}$. The 2 preimages in $SU(2)$ of φ are these:

$$\pm\begin{pmatrix} e^{i\varphi/2} & 0\\ 0 & e^{-i\varphi/2}\end{pmatrix}\mapsto\begin{bmatrix}\cos\varphi & -\sin\varphi & 0\\ \sin\varphi & \cos\varphi & 0\\ 0 & 0 & 1\end{bmatrix}$$

So we have:

$$\rho_l(\varphi)v_{l,k}=\begin{pmatrix} e^{i\varphi/2} & 0\\ 0 & e^{-i\varphi/2}\end{pmatrix}x^ky^{2l-k}=e^{(2k-2l)i\varphi/2}x^ky^{2l-k}=e^{(2k-2l)i\varphi/2}v_{l,k}\,.$$

For $g\in G$, we have

$$\begin{aligned} f_{l,j,k}(g\varphi)&=\langle\rho_l(g\varphi)v_{l,k},v_{l,j}\rangle=\langle\rho_l(g)\rho_l(\varphi)v_{l,k},v_{l,j}\rangle=\langle\rho_l(g)e^{(2k-2l)i\varphi/2}v_{l,k},v_{l,j}\rangle\\ &=e^{(2k-2l)i\varphi/2}f_{l,j,k}(g)\,.\end{aligned}$$

We conclude that $F_*(f_{l,j,k})=0$ if $k\neq l$ and that $F_*(f_{l,j,l})=f_{l,j,l}$. Hence the $f_{l,j,l}$, with $0\leq l$ and $0\leq j\leq 2l$ form a complete orthonormal set in $C(G)^{S^1}$.

Now it remains to compute the images $Y_{l,j}$ of the $\sqrt{2l+1}f_{l,j,l}$ under the inverse of our isomorphism $F^*\colon C(S^2)\to C(G)^{S^1}$. This means that for a point P in S^2, we must find an h in $SU(2)$ such that $h{\cdot}N=P$ and then take the value

$$Y_{l,j}(P)=\sqrt{2l+1}f_{l,j,l}(h)=\sqrt{2l+1}\langle h{\cdot}v_{l,l},v_{l,j}\rangle\,.$$

To do so, we write points in S^2 in the usual spherical coordinates:

$$P(\theta,\varphi)=(\sin(\theta)\cos(\varphi),\sin(\theta)\sin(\varphi),\cos(\theta))\,,\quad\text{with } 0\leq\theta\leq\pi,\ -\pi\leq\varphi\leq\pi.$$

We observe that (draw the usual picture for spherical coordinates, and note the rotation over θ about the y-axis followed by the rotation over φ about the z-axis):

$$\begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix}=\begin{bmatrix}\cos\varphi & -\sin\varphi & 0\\ \sin\varphi & \cos\varphi & 0\\ 0 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix}\cos\theta & 0 & \sin\theta\\ 0 & 1 & 0\\ -\sin\theta & 0 & \cos\theta\end{bmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

We need the inverse images in $SU(2)$ of the 2 matrices just above. These are $\pm e^{(\varphi/2)I}$ and $\pm e^{(\theta/2)K}$ (exponentials in $\mathbb{H}$) because, with our choices, the z-axis corresponds to I and the y-axis to K. Writing these exponentials out as complex 2 by 2 matrices we get:

$$h=h_2h_1\,,\quad h_2=\begin{pmatrix} e^{i\varphi/2} & 0\\ 0 & e^{-i\varphi/2}\end{pmatrix},\quad h_1=\begin{pmatrix}\cos(\theta/2) & i\sin(\theta/2)\\ i\sin(\theta/2) & \cos(\theta/2)\end{pmatrix}.$$

Then

$$\begin{aligned} Y_{l,j}(P(\theta,\varphi))&=\sqrt{2l+1}\langle h_2h_1{\cdot}v_{l,l},v_{l,j}\rangle=\sqrt{2l+1}\langle h_1{\cdot}v_{l,l},h_2^{-1}{\cdot}v_{l,j}\rangle\\ &=\sqrt{2l+1}\,\lambda_{2l,l}\,\lambda_{2l,j}\,e^{i\varphi(j-l)}\langle h_1{\cdot}x^ly^l,x^jy^{2l-j}\rangle\,.\end{aligned}$$

We compute $h_1 \cdot x^l y^l$. Note that h_1 gives the automorphism of the $\mathbb{C}$-algebra $\mathbb{C}[x,y]$ with

$$h_1\colon x \mapsto \cos(\theta/2)x + i\sin(\theta/2)y\,, \quad h_1\colon y \mapsto i\sin(\theta/2)x + \cos(\theta/2)y\,.$$

Hence

$$h_1 \cdot x^l y^l = \big(\cos(\theta/2)x + i\sin(\theta/2)y\big)^l \big(i\sin(\theta/2)x + \cos(\theta/2)y\big)^l\,.$$

Using Newton's binomial identity, we get

$$h_1 \cdot x^l y^l = \sum_{j=0}^{2l} c_{l,j} x^j y^{2l-j}\,, \quad c_{l,j} = \sum_{\substack{0\le n,m\le l\\ n+m=j}} \binom{l}{n}\binom{l}{m} \cos(\theta/2)^{n+l-m}\big(i\sin(\theta/2)\big)^{m+l-n}\,.$$

It follows that

$$\langle h_1 \cdot x^l y^l, x^j y^{2l-j}\rangle = \lambda_{2l,j}^{-2} c_{l,j}\,.$$

So our final result is that the $Y_{l,j}$ with $l \in \mathbb{N}$ and $0 \le j \le 2l$, given by:

$$\begin{aligned} Y_{l,j}(P(\theta,\varphi)) &= \sqrt{2l+1}\,\lambda_{2l,l}\,\lambda_{2l,j}\,e^{i\varphi(j-l)}\,c_{l,j} = \\ &= \sqrt{2l+1}\binom{2l}{l}^{1/2}\binom{2l}{j}^{-1/2} e^{i\varphi(j-l)} \sum_{\substack{0\le n,m\le l\\ n+m=j}} \binom{l}{n}\binom{l}{m}\cos(\theta/2)^{n+l-m}\big(i\sin(\theta/2)\big)^{m+l-n}\,, \end{aligned}$$

form an orthonormal Hilbert-basis of $L^2(S^2)$.

5.3 Examples

We make a few cases more explicit. For $l = 0$ we have $j = 0$ and $n = m = 0$, and that gives:

$$Y_{0,0}(P(\theta,\varphi)) = 1\,.$$

It is nice to note that indeed $\int_{S^2} Y_{0,0}\mu_{S^2} = 1$, as μ_{S^2} has been normalised for this to hold, it is the invariant probability measure. Physicists often use the invariant volume form on S^2 that comes from the euclidean metric on $\mathbb{R}^3$ and then the area of S^2 is 4π, which means that their constant spherical harmonic has value $1/\sqrt{4\pi}$.

Let us now take $l = 1$. Then $j \in \{0,1,2\}$. For $j = 0$ we have $n = m = 0$, for $j = 1$ there are 2 cases, $n = 0$ and $m = 1$, $n = 1$ and $m = 0$. For $j = 2$ we have $n = m = 1$. This gives us:

$$\begin{aligned} Y_{1,0}(P(\theta,\varphi)) &= \sqrt{3/2}\,ie^{-i\varphi}\sin(\theta)\,, \\ Y_{1,1}(P(\theta,\varphi)) &= \sqrt{3}\,\cos(\theta)\,, \\ Y_{1,2}(P(\theta,\varphi)) &= \sqrt{3/2}\,ie^{i\varphi}\sin(\theta)\,. \end{aligned}$$

For $l = 2$ we have $j \in \{0,1,2,3,4\}$. We find:

$$\begin{aligned} Y_{2,0}(P(\theta,\varphi)) &= 4^{-1}\sqrt{30}\,e^{-2i\varphi}(\cos(\theta)^2 - 1)\,, \\ Y_{2,1}(P(\theta,\varphi)) &= 2^{-1}\sqrt{30}\,e^{-i\varphi}i\,\cos(\theta)\sin(\theta)\,, \\ Y_{2,2}(P(\theta,\varphi)) &= 2^{-1}\sqrt{5}\big(3\cos(\theta)^2 - 1\big)\,, \\ Y_{2,3}(P(\theta,\varphi)) &= 2^{-1}\sqrt{30}\,e^{i\varphi}i\,\cos(\theta)\sin(\theta)\,, \\ Y_{2,4}(P(\theta,\varphi)) &= 4^{-1}\sqrt{30}\,e^{2i\varphi}(\cos(\theta)^2 - 1)\,. \end{aligned}$$

All these functions are, up to the factor $1/\sqrt{4\pi}$ that we already explained, equal to the spherical-harmonic functions as given in [5], table 4.3, with the dictionary that our $Y_{l,j}$ corresponds to Y_l^{j-l} there ($j-l$ is the number that gives the action of S^1 acting on S^2, we will come back to this).

In figures 2, 3 and 4 the $Y_{l,j}$ are visualised as follows: the set $\{|Y_{l,j}(P)|\cdot P : P \in S^2\}$ is plotted for $l = 0, 1, 2$, but then without the normalising constants. For example, for $l = 1$ and $j = 0$ it is the parametrised plots of the function

$$[0,\pi] \times [-\pi,\pi] \to \mathbb{R}^3\,, \quad (\theta,\varphi) \mapsto |ie^{-i\varphi}\sin(\theta)|\cdot(\sin(\theta)\cos(\varphi), \sin(\theta)\sin(\varphi), \cos(\theta))\,.$$

It is a bit unfortunate that, in these plots, the argument of the $Y_{l,j}$ is not visible. w would need a second series of plots for that, or some informative colour coding in the given plots.

We note that the collection of Y_l^m is not the only orthonormal basis of $L^2(S^2)$! A priori many more choices are possible. We see this in the choices we had for the basis of V and of the orthonormal basis for $\mathbb{C}[x,y]_d$. So it appears more or less as a coincidence that the we find the orthonormal basis that is used in quantum mechanics! That this is not a coincidence at all is made clear in the next chapter, about the relation between the mathematics of the preceding sections and physics.

5.4 Quantum mechanics and the hydrogen atom

The foundations quantum mechanics have been laid in the beginning of the 20th century by many famous scientists such as Heisenberg, Planck, Schrödinger, Pauli and many others. Thus it is a relatively modern theory and still now it is, together with the theory of relativity, one of the most fundamental theories in physics.

Quantum mechanics meant a radical revolution in physics: while before it the state of an object was described by a point in phase space, and observables by functions on this space — think of a particle with a certain speed and position: the state of the particle is described by the vector $(x_1, x_2, x_3, v_1, v_2, v_3) \in \mathbb{R}^6$, and its kinetic energy by the function $(\ldots) \mapsto \frac{1}{2}m(v_1^2 + v_2^2 + v_3^2)$ — in quantum mechanics the state of the object is described by a *(wave) function* in a complex Hilbert space and an observable by a self-adjoint linear *operator* on the same space. These operators typically do not commute with each other, so one can view these operators as elements of a Lie algebra with a non-trivial Lie bracket.

A beautiful example of this are the operators L_x, L_y and L_z that correspond, respectively, to the x, y, and z-components of the angular momentum of a particle. These operators are given by the equation for the angular momentum (from classical mechanics):

$$\overrightarrow{L} = \overrightarrow{r} \times \overrightarrow{p}$$

with $\overrightarrow{r}$ the position operator and $\overrightarrow{p}$ the momentum operator given by $\overrightarrow{p} = \frac{\hbar}{i}\overrightarrow{\nabla}$. When one choses the units such that $\hbar = 1$, then one has

$$[L_x, L_y] = iL_z, \quad [L_y, L_z] = iL_x, \quad [L_z, L_x] = iL_y$$

In this case we see, after some thought, that if one makes the following identifications:

$$\begin{aligned} L_x &\mapsto \frac{i}{2}I \\ L_y &\mapsto \frac{i}{2}J \\ L_z &\mapsto \frac{i}{2}K\,, \end{aligned}$$

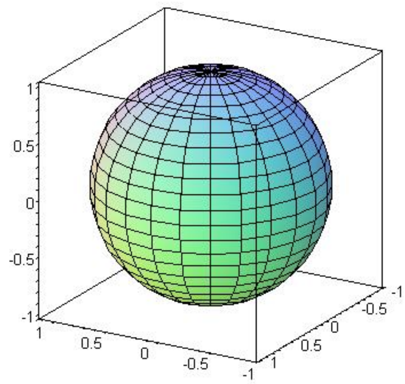

Figure 2: $l = 0$, the function "$r = 1$".

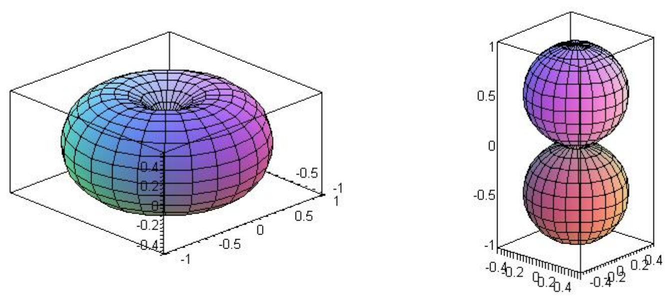

Figure 3: $l = 1$, the functions "$r = \sin\theta$" (left), "$r = |\cos\theta|$" (right).

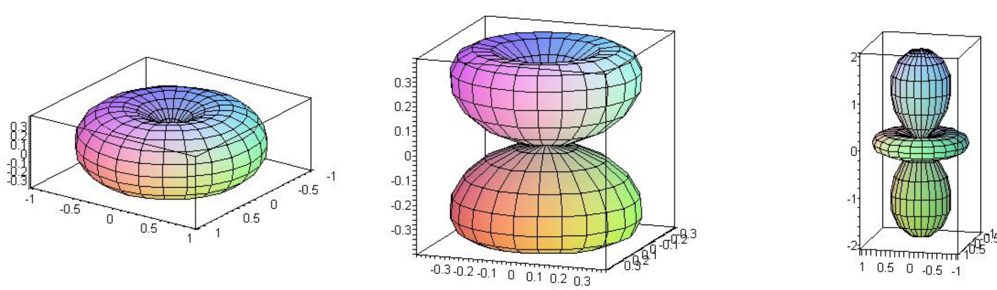

Figure 4: $l = 2$, the functions "$r = (\sin\theta)^2$" (left), "$r = |\cos\theta|\sin\theta$" (middle) and "$r = |3\cos\theta - 1|$" (right).

the real vector space spanned by L_x, L_y, L_z is, as Lie algebra, isomorphic with $\mathfrak{su}(2)$. The action of the operators L_x, L_y and L_z then gives an action of $\mathfrak{su}(2)$ on the Hilbert space of wave functions. It appears, see [7], chapter 8, section 3, that this action of $\mathfrak{su}(2)$ matches the rotations of the space, as follows. Let $R \in SO(3)$. Then R acts on the wave function $\Psi(\overrightarrow{x})$ as

$$R\Psi(\overrightarrow{x}) = \Psi(\overrightarrow{x} \cdot R)$$

The induced action of the Lie algebra $\mathfrak{so}(3)$ ($\cong \mathfrak{su}(2)$) is then "the same".

The time evolution of the function $\Psi(\mathbf{x}, t)$ that describes the state of the object is given by the *Schrödinger equation*:

$$i\hbar\frac{\partial \Psi}{\partial t} = H\Psi\,.$$

Here $H = \frac{p^2}{2m} + V$ with V the potential. In the case of the hydrogen atom the "object" that we study is the electron subject to the electric potential of a proton sitting at the origin. In this case the potential V does not depend on time, and therefore there is a set of *stationairy states* Ψ_i, $(i > 0)$, spanning the Hilbert space and where every Ψ_i has a time independent factor ψ_i that satisfies the *time independent* Schrödinger equation:

$$H\psi_i = E_i\psi_i$$

with $E_i \in \mathbb{R}$ the energy of ψ_i. As the potential V depends only on the distance to the origin, the hamiltonian H is invariant is under rotations, that is, under the action of $SO(3)$ defined above. It follows that the operators L_x, L_y and L_z each commute with H and that each has a common set of eigenfunctions with H, see [5] or [7]. In the quantum mechanics one mostly uses the eigenfunctions of L_z and H. In spherical coordinates an eigenfunction ψ_i depends on (r, θ, φ). For a fixed $r = r_0$, $\psi_i(r_0, \theta, \varphi)$ is in $L^2(S^2)$, and hence can be written as

$$\psi_i(r_0, \theta, \varphi) = \sum_j R_j(r_0)Y_j(\theta, \varphi)$$

for a basis $\{Y_j \mid j \in J\}$ of $L^2(S^2)$. As H and L_z commute, one can ask that the Y_j are normalised orthogonal eigenfunctions of L_z. These Y_j arise in quantum mechanics as the spherical harmonic functions and form an orthonormal basis of eigenfunctions of L_z in $L^2(S^2)$.

From the above we see how important the theory of Lie groups and their representations is for moderne physics, and in particular for quantum mechanics. Operators form a Lie algebra and the space of wave functions with the action of operators form a representation. The theory of spin is a beautiful example of this. Where angular momentum has a classical analog, spin is a purely mathematical construction. The Lie algebra $\mathfrak{su}(2)$ acts on an irreducible representation $\mathbb{C}[x, y]_d$ of $SU(2)$. In contrast with the more "physical" group $SO(3)$, representations can now be also of odd degree d zijn. Analogous to the definition of l in section 5.3, now we have $s = \frac{1}{2}d$. This explains the curious fact that the spin can be half integer! See also [5], chapter 4. As stationairy states of the spin one searches for an orthonormal basis of eigenvectors of $\frac{i}{2}K$ for $\mathbb{C}[x, y]_{2s}$. Take for example the electron, it has $s = \frac{1}{2}$. Then $\{\frac{x}{\pi}, \frac{y}{\pi}\}$ form an orthonormal basis for $\mathbb{C}[x, y]_1$. An eigenvector of $\frac{i}{2}K$ is a eigenvector of K, so we search eigenvectors of the matrix

$$K = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

We have

$$\begin{array}{rcl} K(x+y) & = & i(x+y) \\ K(x-y) & = & -i(x-y) \end{array}$$

hence

$$(\frac{\sqrt{2}}{\pi}(x+y), \frac{\sqrt{2}}{\pi}(x-y))$$

form an orthonormal basis of eigenvectors of K. This gives eigenvalues of $-\frac{1}{2}$ and $\frac{1}{2}$ for $\frac{i}{2}K$, respectively, corresponding to the familiar spin up and down states of $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$ for $\hbar = 1$.

In the preceding chapters we have noticed the amount of freedom in choosing a basis: we could freely choose the basis for V, and also the orthonormal basis for $\mathbb{C}[x,y]_d$. So it was a "coincidence" that we got the right orthonormal basis of $L^2(S^2)$ via our choices. That this is not a complete coincidence can be understood as follows: the basis that we seek is one of eigenfunctions of

$$L_z = xp_y - yp_x = i\hbar y\frac{\partial}{\partial x} - i\hbar x\frac{\partial}{\partial y}.$$

As mentioned above angular momentum operators correspond with elements of $\mathfrak{so}(3)$ as follows: take the infinitesimal rotation about the z-as, $R_z(\epsilon) \in SO(3)$. It acts on a function $\psi(x,y,z)$ as

$$\begin{array}{rcl} R_z(\epsilon)(\psi(x,y,z)) & = & \psi\left((x,y,z)\circ\begin{bmatrix} \cos(\epsilon) & -\sin(\epsilon) & 0 \\ \sin(\epsilon) & \cos(\epsilon) & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \\ & \approx & \left((x,y,z)\circ\begin{bmatrix} 1 & -\epsilon & 0 \\ \epsilon & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \\ & = & \psi(x+\epsilon y, y-\epsilon x, z) \\ & \approx & \psi(x,y,z) + \epsilon\left(y\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial y}\right) \\ & = & (1 + \frac{\epsilon}{i\hbar}L_z)\psi(x,y,z) \end{array}$$

As $R_z(\epsilon) = 1 + \epsilon\sigma$ with $\sigma = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{so}(3)$, ψ is an eigenfunction of L_z iff it is an eigenfunction of $\sigma \in \mathfrak{so}(3)$. In section 5.2 we have chosen the map $\mathfrak{su}(2) \to \mathfrak{so}(V) \to \mathfrak{so}(3)$ so that σ corresponded with the matrix $I \in \mathfrak{su}(2)$ up to a scalar factor.

From the same section it follows that the functions $f_{i,j,k}$ could be written as $f_{i,j,k}(g) = \langle\rho_i(g)v_{ij}, v_{ik}\rangle$. The action of $R_z(\epsilon)$ then gives

$$R_z(\epsilon)(f_{i,j,k}(g)) = f_{i,j,k}(g\cdot R_z(\epsilon)) = \langle\rho_i(g)\cdot\rho_i(R_z(\epsilon))v_{ij}, v_{ik}\rangle$$

and this gives the action of σ:

$$\sigma(f_{i,j,k}(g)) = \langle\rho_i(g)\frac{1}{2}Iv_{ij}, vik\rangle$$

But as we chose the basis $\{v_{i1}, ..., v_{id}\}$ of V_i so that it consisted of eigenvectoren of I, see section 4.5, we have

$$\sigma(f_{i,j,k}(g)) = \lambda\langle\rho_i(g)v_{ij}, v_{ik}\rangle = \lambda f_{i,j,k}(g)$$

for a certain λ. That's why the functions $f_{i,j,k}(g)$ are eigenfunctions of σ and L_z. We could of course have chosen the map $\mathfrak{so}(V) \to \mathfrak{so}(3)$ so that σ would correspond with K, or an arbitrary other element in $\mathfrak{su}(2)$, but then we should have chosen our orthonormal basis of V_i so that it consisted of eigenvectors of this arbitrary element. The choices I have made in this thesis are in my eyes the choices that make the computations easiest.

6 Finally . . .

By computing the the spherical harmonic functions via representation theory of compact Lie groups I have reached the goal of my thesis. On my way to this goal I have treated, briefly, differentiable manifolds, Lie groups, Lie algebras and representations. I have delved deeper into the irreducible representations of $SU(2)$, their bases, the covering $SU(2) \to SO(3)$ and the Peter Weyl theorem. All this has enabled me to compute the spherical-harmonic functions. The result agrees with the literature, a fine reward for all this theory. Finally I have briefly touched upon applications of the theory to quantum mechanics, with an explanation of how spin relates to $\mathfrak{su}(2)$ and some explanation of how angular momentum operators relate to $\mathfrak{so}(3)$. In my research I have amply used the sources below and I have tried to refer, at important points, to the relevant literature. But what does not show in the bibliography is the help that I got from my three enthusiastic supervisors, Theo van den Bogaart, Gerard Nienhuis and Bas Edixhoven. Many thanks!

References

[1] Klaus Jänich: *Vector Analysis*, Springer-Verlag New York (2001)

[2] Bas Edixhoven: *Lie groups and Lie algebras*, Université de Rennes (2001),

`http://www.math.leidenuniv.nl/~edix/public_html_rennes/cours/dea0001.pdf`

[3] J.J. Duistermaat, J.A.C. Kolk: *Lie Groups*, Springer-Verlag Berlin (2000)

[4] Grant R. Fowles, George L. Cassiday: *Analytical Mechanics*, Thomson Learning, Inc. (6th Ethision)

[5] David J. Griffiths: *Introduction to Quantum Mechanics*, Pearson Education (2nd Edition).

[6] Peter D. Lax: *Linear Algebra*, John Wiley & Sons (1997)

[7] Albert Messiah: *Quantum Mechanics, Volume II*, North-Holland Publishing Company Amsterdam (1973)

[8] A. Mukherjea, K. Pothoven: *Real and Functional Analysis*, Plenum Press New York (1978)

[9] Ernest B. Vinberg: *Linear representations of groups*, Translated from the Russian by A. Iacob. Basel Textbooks, 2. Birkhäuser Verlag, Basel, 1989.