

MATHEMATICAL METHODS OF PHYSICS I – 2014

THOMAS CREUTZIG

ABSTRACT. These are lecture notes in progress for Ma Ph 451 – Mathematical Physics I. The lecture starts with a brief discussion of linear algebra, Hilbert spaces and classical orthogonal polynomials. Then as an instructive example the Lie group $SU(2)$ and its Hilbert space of square integrable functions will be discussed in detail. The focus of the second part of the lecture will then be on Lie groups in general, their Lie algebras and its representation theory. Guiding examples are both $sl(2; \mathbb{R})$ and $sl(3; \mathbb{R})$. Finally, we discuss how Lie groups, their harmonic analysis and especially Lie algebras appear in the example of bosonic strings.

CONTENTS

1.	Introduction	4
2.	Linear Algebra	6
	2.1. Vector Spaces	6
	2.2. Linear Transformations and Operators	9
	2.3. Operators	11
	2.4. Eigenvectors and eigenvalues	14
	2.5. Examples	15
	2.6. Exercises	16
	2.7. Solutions	17
3.	Hilbert Spaces	18
	3.1. The definition of a Hilbert space	19
	3.2. Square integrable functions	22
	3.3. Classical orthogonal polynomials	23
	3.4. Gegenbauer polynomials and hypergeometric functions	26
	3.5. Hermite polynomials	28
	3.6. Exercises	29
	3.7. Solutions	30
4.	Harmonic Analysis	32
	4.1. Motivation	32
	4.2. Distributions	33
	4.3. Fourier Analysis	34
	4.4. Harmonic analysis on the Lie group $U(1)$	35
	4.5. Harmonic analysis on S^3	37
	4.6. Summary	48
	4.7. Exercises	49
	4.8. Solutions	49
5.	Lie Groups	50
	5.1. The Haar measure	52
	5.2. Lie subgroups of $GL(n, \mathbb{C})$	54
	5.3. Left-Invariant vector fields	56
	5.4. Example of a Lie supergroup	58
6.	Lie Algebras	61
	6.1. The Casimir element of a representation	65
	6.2. Jordan Decomposition	68
	6.3. Root Space Decomposition	69
	6.4. Finite-dimensional irreducible representations of $sl(2; \mathbb{R})$	72
	6.5. Representation theory	74
	6.6. Highest-weight representations of $sl(3; \mathbb{R})$	75
	6.7. Exercises	76
	6.8. Solutions	77
7.	The bosonic string	79
	7.1. The free boson compactified on a circle	79
	7.2. The Virasoro algebra	82
	7.3. Lattice CFT	83

7.4. Fermionic Ghosts	84
7.5. BRST quantization of the bosonic string	84
8. Possible Exam Questions	85
References	87

1. INTRODUCTION

We will cover subjects of interest in mathematical physics. Lie theory is a fascinating area of mathematical physics with various applications in both areas. So the focus of this lecture will be on Lie theory. Lie groups are smooth manifolds that at the same time are groups. They have been introduced by Sophus Lie. In studying Lie groups it turns out that it is much simpler to consider infinitesimal transformations. These transformations have an algebraic structure called Lie algebra. Lie algebras and Lie groups are important as they appear as the symmetries of physical systems. This leads to a variety of connections of interesting modern topics of mathematical physics. Both Lie groups and Lie algebras are often best visualized using matrices. Examples that are familiar to most physicists are the Heisenberg Lie group and the Lie group of the standard model of particle physics.

The three-dimensional Heisenberg group is

$$H = \left\{ \left(\begin{array}{ccc} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{array} \right) \middle| a, b, c \in \mathbb{R} \right\}.$$

So it is a subgroup of invertible 3×3 matrices. Its Lie algebra \mathfrak{h} on the other hand is spanned by

$$\mathfrak{h} = \text{span} \left\{ p = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}$$

Lie group and Lie algebra are related by the exponential map

$$\exp(X) = \sum_{n=0}^{\infty} \frac{X^n}{n!} \in H \text{ for all } X \in \mathfrak{h}.$$

(You can check this doing explicit matrix multiplication) The commutator of matrices defines an algebra structure on \mathfrak{h} ,

$$[X, Y] = XY - YX.$$

Performing the appropriate matrix multiplication one finds that these commutators or Lie brackets are zero, except for

$$[p, q] = z.$$

But this is exactly the Heisenberg Lie algebra of quantum mechanics under the identification $p = -i\hbar \frac{d}{dx}$, $q = x$, $z = -i\hbar$.

The Lie group of the standard model is $U(1) \times SU(2) \times SU(3)$. This is the product of a one-dimensional Lie group $U(1)$, a three-dimensional Lie group $SU(2)$ and an eight-dimensional one $SU(3)$. The standard model is a gauge theory, and each generator of the Lie group has an associated gauge boson. The electroweak interaction is described by $U(1) \times SU(2)$, and the $U(1)$ -gauge particle is the photon, while the ones for $SU(2)$ are called W^\pm and Z gauge bosons. Quantum chromodynamics is associated to the $SU(3)$ and the eight gauge bosons are called gluons. Quarks and leptons then come in colors, with spins and electric charge. These quantities encode how the particles behave under the action of the gauge group, that is the Lie group $U(1) \times SU(2) \times SU(3)$. A mathematician would call these quantities weights, and the representation theory of the underlying Lie algebra would tell him how the gauge group acts.

The initial purpose of studying Lie theory was the understanding of certain differential equations. If you look back to your quantum mechanics course, the Heisenberg algebra and also the

Lie algebra of infinitesimal rotations describing spin and angular momentum, were used to solve the Schrödinger equation, a differential equation. In general, the representation theory of Lie algebras is a great aide in simplifying certain second order differential equations.

There are more modern connections between Lie theory, physics and mathematics. String theory is a quantum theory of strings and not of point-like particles. It's initial motivation was the search for a quantum theory that both incorporates particle physics and gravity. The framework of string theory uses a variety of mathematics, ranging from geometry, topology, algebra and number theory. But Lie theory is always central, simply since a Lie group or Lie algebra of some kind will always appear as symmetry of the theory in question. Moreover, the world-sheet theory of a string is a two-dimensional conformal field theory. Local conformal transformations in two-dimensions generate an infinite-dimensional algebra, the Witt algebra. Its central extension is called the Virasoro algebra. This is an infinite-dimensional Lie algebra and it is contained in the symmetry algebra of every world-sheet theory of every string. Only due to Lie theory the conformal field theory can be exactly solved. This is extremely special; exactly solvable quantum field theories are very rare. Both conformal field theory and string theory have led to a tremendous progress in pure mathematics. I plan to tell you a little bit about the bosonic string, infinite-dimensional Lie algebras and the monstrous moonshine at the end of the lecture as an interesting application/generalization of what you have learnt. This monstrous moonshine is a surprising connection between modular functions (number theory) and finite groups. But this is by far not the only progress in mathematics due to physics related to Lie theory. Another one is three-dimensional topological quantum field theories and invariants of knots. You surely know what a knot is, and you can probably imagine that it is difficult to describe arbitrary knots on some three-manifold. However, Witten found that there is a topological field theory, called Chern–Simons theory, that can be used to understand the problem of describing knots. If you want to describe a knot in some three-manifold M , you don't really care about the size of the knot. So you need a theory that is metric independent to describe them, the Chern–Simons theory. Let G be a compact Lie group, the gauge group. You should think of this Lie group as a subgroup of invertible $n \times n$ matrices. The Chern-Simons action is then built out of Lie group valued (that is matrix valued) gauge connections A ,

$$S = \frac{k}{4\pi} \int_M \text{tr} \left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right).$$

More correctly, A is a matrix in some representation \mathcal{R} of G . tr denotes the trace, that is the sum over the diagonal entries, of the matrix. To each knot \mathcal{C} , one can associate a Wilson loop

$$W(\mathcal{C}, \mathcal{R}) = \text{tr} \left(P \exp \left(\int_{\mathcal{C}} A \right) \right)$$

as the trace of the path ordered exponential of the integral along the knot of the gauge connection. Expectation values of these Wilson loops are then invariants of knots that are naturally associated to representations of compact Lie groups. This physics description finally allowed to characterize knots in an efficient and treatable manner. Both Edward Witten and Vaughan Jones are the pioneers in this area, and both of them received the fields medal. Chern–Simons theory is a topological quantum field theory. However, if the manifold M has a boundary, then its boundary degrees of freedom are described by a two-dimensional conformal quantum field theory, the Wess–Zumino–Novikov–Witten model of the Lie group G . The action of this theory looks similar as the Chern–Simons action, just that the gauge connection is replaced by the Maurer–Cartan one form. This is the invariant Lie algebra valued one-form of a Lie group.

The conformal field theory is completely described by the representation theory of the affinization of the Lie algebra of G . Turning things around, another huge success of the last decades in mathematical physics was the interaction between infinite-dimensional Lie algebras (Kac-Moody algebras), conformal field theory and modular forms. Note, that Robert Moody has been a professor of the University of Alberta. The success is due to the very surprising fact, that the representation theory of the Lie algebra is guided by modular forms, that is by functions that behave nicely under the action of the Möbius group $SL(2; \mathbb{Z})$ action

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tau = \frac{a\tau + b}{c\tau + d}.$$

However, for a physicists this is less of a surprise. The point being that axioms of two-dimensional conformal field theory tell us, that the representation category of the field theory is a modular one. The really beautiful thing now is, that this modular representation category is essentially a formal way of looking at knots and their invariants. To summarize, physics knows a three-dimensional topological field theories and their two-dimensional conformal field theories at the boundary. These two somehow imply a relation between knots on three-manifolds, infinite-dimensional Lie algebras and modular forms. This is very fascinating and here at the University of Alberta in the mathematical physics group we all perform research in closely related directions.

2. LINEAR ALGEBRA

We start with a review of linear algebra, a course that most of you have taken the first year of your studies. Linear algebra deals with linear transformations on finite dimensional vector spaces. Such linear transformations are best represented using matrices and important examples are translations, rotations and reflections.

We will mostly be concerned with the field of complex numbers

$$\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R}; i^2 = -1\}.$$

This section follows chapter 2 of [H].

2.1. Vector Spaces. A vector space is defined as follows

Definition 1. A *vector space* \mathcal{V} over \mathbb{C} is a set \mathcal{V} whose elements $|a\rangle \in \mathcal{V}$ are called *vectors*. This set is endowed with two operations, addition of vectors

$$+ : \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}, \quad (|a\rangle, |b\rangle) \mapsto |a\rangle + |b\rangle$$

and multiplication by a scalar

$$\cdot : \mathbb{C} \times \mathcal{V} \rightarrow \mathcal{V}, \quad (\lambda, |a\rangle) \mapsto \lambda|a\rangle.$$

Addition satisfies the following list of properties for every three vectors $|a\rangle, |b\rangle, |c\rangle \in \mathcal{V}$.

- (1) *commutativity*: $|a\rangle + |b\rangle = |b\rangle + |a\rangle$;
- (2) *associativity*: $|a\rangle + (|b\rangle + |c\rangle) = (|a\rangle + |b\rangle) + |c\rangle$;
- (3) *additive identity*: There exists a unique vector $|0\rangle$, called the *zero vector*, satisfying $|0\rangle + |a\rangle = |a\rangle$;
- (4) *inverse*: There exists a unique vector $-|a\rangle$ satisfying $|a\rangle + (-|a\rangle) = |0\rangle$;

Multiplication with a scalar satisfies for every $|a\rangle \in \mathcal{V}$ and every two scalars $\lambda, \mu \in \mathbb{C}$.

- (1) *associativity*: $\lambda(\mu|a\rangle) = (\lambda\mu)|a\rangle$;

(2) *multiplicative identity*: $1|a\rangle = |a\rangle$.

Finally, there are two *distributivity* laws combining vector addition and multiplication with a scalar. For every two vectors $|a\rangle, |b\rangle \in \mathcal{V}$ and for every two scalars $\lambda, \mu \in \mathbb{C}$ they are as follows.

$$(1) \lambda(|a\rangle + |b\rangle) = \lambda|a\rangle + \lambda|b\rangle;$$

$$(2) (\lambda + \mu)|a\rangle = \lambda|a\rangle + \mu|a\rangle.$$

We will recall some more well-known linear algebra terms.

(1) A sum of vectors $\lambda_1|a_1\rangle + \dots + \lambda_n|a_n\rangle$ is called a linear combination of the vectors $|a_1\rangle, \dots, |a_n\rangle$. If the relation $\lambda_1|a_1\rangle + \dots + \lambda_n|a_n\rangle = |0\rangle$ implies that $\lambda_1 = \dots = \lambda_n = 0$, then the vectors $|a_1\rangle, \dots, |a_n\rangle$ are said to be linear independent. This means that the vector $|a_i\rangle$ for any $1 \leq i \leq n$ cannot be written as a linear combination of the other $n - 1$ vectors $|a_j\rangle, j \neq i$.

(2) A subset \mathcal{W} of a vector space \mathcal{V} is itself a vector space, that is a sub vector space of \mathcal{V} , if it is closed under addition and multiplication by a scalar. In formulae this means for all $|a\rangle, |b\rangle \in \mathcal{W}$ and any scalar $\lambda \in \mathbb{C}$, also

$$|a\rangle + |b\rangle \quad \text{and} \quad \lambda|a\rangle$$

are in \mathcal{W} .

(3) If \mathcal{S} is a set of vectors in our vector space \mathcal{V} , then the set of all linear combinations of the vectors in \mathcal{S} is a sub vector space of \mathcal{V} . This sub vector space has the name span of \mathcal{S} , and we write $\mathcal{W}_{\mathcal{S}}$.

(4) A basis of a vector space \mathcal{V} is a set \mathcal{B} of linearly independent vectors such that its span is the vector space \mathcal{V} itself, $\mathcal{W}_{\mathcal{B}} = \mathcal{V}$.

(5) All bases of a given vector space have the same cardinality. Especially, if a basis of \mathcal{V} has only finitely many elements, say d , then every basis of \mathcal{V} has d elements. The number d is then called the dimension of \mathcal{V} .

Another useful structure one can introduce on vector spaces is an inner product.

Definition 2. The *inner product* on a vector space \mathcal{V} is a map $\mathcal{V} \times \mathcal{V} \rightarrow \mathbb{C}$ that maps $(|a\rangle, |b\rangle) \in \mathcal{V} \times \mathcal{V}$ to $\langle a|b\rangle \in \mathbb{C}$ satisfying for all $|a\rangle, |b\rangle, |c\rangle \in \mathcal{V}$ and all $\lambda, \mu \in \mathbb{C}$

$$(1) \text{ sesquilinearity or hermiticity: } \langle a|b\rangle = \langle b|a\rangle^*;$$

$$(2) \text{ linearity in one component: } \langle a|(\lambda|b\rangle + \mu|c\rangle) = \lambda\langle a|b\rangle + \mu\langle a|c\rangle;$$

$$(3) \text{ positive definiteness: } \langle a|a\rangle \geq 0, \text{ and } \langle a|a\rangle = 0 \text{ if and only if } |a\rangle = |0\rangle.$$

Here μ^* denotes complex conjugate of the complex number μ .

The Dirac bracket notation has its original use in quantum mechanics. We say that $\langle a|$ is *bra of a* and $|a\rangle$ is *ket of a*. It is nothing but a notation and every other way of denoting vectors is of course allowed, though it is useful to keep some conventions. The inner product singles out a special and very useful kind of basis

Definition 3. Let \mathcal{V} be a vector space with inner product. We call two vectors $|a\rangle, |b\rangle$ *orthogonal* if $\langle a|b\rangle = 0$. A vector $|a\rangle$ is called *normalized* or *unit vector* if $\langle a|a\rangle = 1$. A basis $\mathcal{B} = \{|e_1\rangle, \dots, |e_n\rangle\}$ is called *orthonormal* if

$$\langle e_i|e_j\rangle = \delta_{i,j} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}.$$

Such a basis can always be found, and the algorithm to do so is called the Gram–Schmidt–orthonormalization process.

Theorem 1. Let \mathcal{V} be a finite-dimensional vector space, then there exists an orthonormal basis.

The statement can be best proven using induction. The proof is constructive and provides an algorithm to find an orthonormal basis.

Proof. Let $\{|a_1\rangle, \dots, |a_n\rangle\}$ be a basis of \mathcal{V} . Define the sequence of sub vector spaces

$$\mathcal{W}_1 \subset \dots \subset \mathcal{W}_n$$

where \mathcal{W}_m is the span of $|a_1\rangle, \dots, |a_m\rangle$, so that $\mathcal{W}_n = \mathcal{V}$. By property (3) of the definition of an inner product $\langle a_1 | a_1 \rangle > 0$, hence

$$|e_1\rangle := \frac{|a_1\rangle}{\sqrt{\langle a_1 | a_1 \rangle}}$$

is normal and gives a normal basis of the one-dimensional vector space \mathcal{W}_1 .

Our hypothesis of the induction is now that \mathcal{W}_m has orthonormal basis $|e_1\rangle, \dots, |e_m\rangle$. Define

$$|b_{m+1}\rangle := |a_{m+1}\rangle - \sum_{i=1}^m |e_i\rangle \langle e_i | a_{m+1} \rangle.$$

Then for all $1 \leq j \leq m$, we have

$$\langle e_j | b_{m+1} \rangle = \langle e_j | a_{m+1} \rangle - \sum_{i=1}^m \langle e_j | e_i \rangle \langle e_i | a_{m+1} \rangle = \langle e_j | a_{m+1} \rangle - \sum_{i=1}^m \delta_{i,j} \langle e_i | a_{m+1} \rangle = 0.$$

$|b_{m+1}\rangle \neq |0\rangle$, since $|a_{m+1}\rangle$ is not a vector of \mathcal{W}_m and hence linearly independent of the $|e_1\rangle, \dots, |e_m\rangle$. Hence $\langle b_{m+1} | b_{m+1} \rangle > 0$ and

$$|e_{m+1}\rangle := \frac{|b_{m+1}\rangle}{\sqrt{\langle b_{m+1} | b_{m+1} \rangle}}$$

together with $|e_1\rangle, \dots, |e_m\rangle$ form an orthonormal basis of \mathcal{W}_{m+1} . ■

Using similar ideas one can show the following Schwarz inequality

Theorem 2. Let \mathcal{V} be a vector space with inner product. Then for any two vectors $|a\rangle, |b\rangle$ the inequality

$$\langle a | a \rangle \langle b | b \rangle \geq |\langle a | b \rangle|^2$$

holds. Equality is true if and only if $|a\rangle$ and $|b\rangle$ are proportional to each other.

Proof. Let

$$|c\rangle := |b\rangle - \frac{\langle a | b \rangle}{\langle a | a \rangle} |a\rangle.$$

The geometric interpretation is that $|c\rangle$ is the difference of $|b\rangle$ with its projection onto $|a\rangle$. So that especially $|c\rangle$ is orthogonal to $|a\rangle$. We can use this expression to write the inner product of $|b\rangle$ with itself as

$$\langle b | b \rangle = \frac{|\langle a | b \rangle|^2}{\langle a | a \rangle} + \langle c | c \rangle.$$

Since $\langle c|c\rangle \geq 0$, the Schwarz inequality follows. Equality holds if and only if $\langle c|c\rangle = 0$, that is if and only if $|c\rangle = |0\rangle$. But this is true if and only if $|b\rangle$ coincides with its projection onto $|a\rangle$, that is the two vectors are proportional to each other. ■

In physics, the states of a physical system often form a vector space. In order to measure properties of physical states one then needs an inner product as well as operators that operate on the space of physical states.

2.2. Linear Transformations and Operators. Unless otherwise stated, vector spaces are finite dimensional.

Definition 4. Let \mathcal{V} and \mathcal{W} be two vector spaces over \mathbb{C} . A map

$$T : \mathcal{V} \rightarrow \mathcal{W}$$

is called a *linear map* or a *linear transformation* if for all $|a\rangle, |b\rangle \in \mathcal{V}$ and all $\lambda \in \mathbb{C}$ the following two properties are satisfied.

- (1) $T(|a\rangle + |b\rangle) = T(|a\rangle) + T(|b\rangle)$;
- (2) $T(\lambda|a\rangle) = \lambda T(|a\rangle)$.

The space of linear transformations from \mathcal{V} to \mathcal{W} is denoted by $\mathcal{L}(\mathcal{V}, \mathcal{W})$.

The space of linear transformations has itself an interesting algebraic structure. Let $S, T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ for two complex vector spaces \mathcal{V} and \mathcal{W} . We can define addition by

$$S + T : \mathcal{V} \rightarrow \mathcal{W}, \quad |a\rangle \mapsto S(|a\rangle) + T(|a\rangle)$$

and multiplication by a scalar λ by

$$\lambda T : \mathcal{V} \rightarrow \mathcal{W}, \quad |a\rangle \mapsto \lambda T(|a\rangle).$$

It is a direct computation to verify that these define a vector space structure on $\mathcal{L}(\mathcal{V}, \mathcal{W})$.

We can also compose linear maps. Namely let $\mathcal{V}, \mathcal{W}, \mathcal{X}$ be three vector spaces and let $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, $S \in \mathcal{L}(\mathcal{W}, \mathcal{X})$ be two linear transformations then the composition $S \circ T \in \mathcal{L}(\mathcal{V}, \mathcal{X})$ is defined as

$$S \circ T : \mathcal{V} \rightarrow \mathcal{X}, \quad |a\rangle \mapsto S(T(|a\rangle)).$$

This product allows to define an algebra structure on $\mathcal{L}(\mathcal{V}) := \mathcal{L}(\mathcal{V}, \mathcal{V})$. In the next definition, we use a, b to denote vectors and not the Dirac bracket notation.

Definition 5. Let \mathcal{V} be a vector space over \mathbb{C} , then \mathcal{V} is an \mathbb{C} -algebra if there is a product

$$\cdot : \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}, \quad (a, b) \mapsto ab$$

satisfying

$$a(b + c) = ab + ac, \quad (a + b)c = ac + bc$$

and

$$a(\lambda b) = \lambda(ab), \quad (\lambda a)b = \lambda(ab)$$

for all $a, b, c \in \mathcal{V}$ and all $\lambda \in \mathbb{C}$.

It is again a direct computation to verify that composition defines an algebra structure on $\mathcal{L}(\mathcal{V})$.

Theorem 3. Let \mathcal{V} and \mathcal{W} be two vector spaces over \mathbb{C} . Then $\mathcal{L}(\mathcal{V}, \mathcal{W})$ is a \mathbb{C} -vector space and $\mathcal{L}(\mathcal{V})$ is even an (associative) \mathbb{C} -algebra.

Let now

$$T : \mathcal{V} \rightarrow \mathcal{W}$$

be a linear transformation and $\mathcal{L}(\mathcal{V}, \mathcal{W})$ the \mathbb{C} -vector space of linear transformations. We list some properties and definitions.

- If $\mathcal{W} = \mathbb{C}$, then $\mathcal{L}(\mathcal{V}, \mathcal{W})$ is called the *dual space* of \mathcal{V} and it is denoted by \mathcal{V}^* . The elements of \mathcal{V}^* are called *linear functionals*; they map each vector in \mathcal{V} to a complex number.
- The set of vectors $|a\rangle$ in \mathcal{V} that are mapped by T to the zero vector in \mathcal{W} , that is $T(|a\rangle) = |0\rangle \in \mathcal{W}$, is called the *kernel of T* . Using the axioms of linear transformations one verifies that the kernel of a linear transformation is a sub vector space of \mathcal{V} .
- The *image of T* is the set of vectors $|b\rangle$ in \mathcal{W} , such that there is at least one $|a\rangle$ in \mathcal{V} with $T(|a\rangle) = |b\rangle$. Again using the axioms of linear transformations one can verify that the image of a linear transformations is a sub vector space (of \mathcal{W}).
- The linear transformation T is called *injective* if the kernel of T is just the zero vector $|0\rangle \in \mathcal{V}$. It is called *surjective* if the image of T is the complete vector space \mathcal{W} , and the transformation is called *bijective* if it is both injective and surjective. In that case \mathcal{V} and \mathcal{W} are said to be *isomorphic*. If $\mathcal{V} = \mathcal{W}$, then a bijective linear transformation from \mathcal{V} to \mathcal{V} is said to be an *automorphism*.

Proposition 4. Let $T : \mathcal{V} \rightarrow \mathcal{W}$ be a linear transformation, then

$$\dim(\text{kernel of } T) + \dim(\text{image of } T) = \dim(\mathcal{V}).$$

Proof. This has probably been proven in a linear algebra course. One takes a basis

$$\{|a_1\rangle, \dots, |a_m\rangle\}$$

of the kernel of T and also one

$$\{|b_1\rangle, \dots, |b_n\rangle\}$$

of the image of T . One chooses then vectors $|a_{m+i}\rangle$ in \mathcal{V} with the property that

$$T(|a_{m+i}\rangle) = |b_i\rangle.$$

The $|a_1\rangle, \dots, |a_{m+n}\rangle$ cannot be linearly dependent. Let $|a\rangle$ be an arbitrary vector in \mathcal{V} , then there exists $|c\rangle$ in the span of $|a_{m+1}\rangle, \dots, |a_{m+n}\rangle$ with

$$T(|a\rangle) = T(|c\rangle)$$

and hence $|a\rangle - |c\rangle$ is in the kernel of T . ■

It follows that two vector spaces can only be isomorphic if their dimensions coincide. In fact, we have

Proposition 5. Two vector spaces over \mathbb{C} are isomorphic if and only if they have the same dimension.

Proof. Let \mathcal{V} and \mathcal{W} be two vector spaces of same dimension n and let

$$\{|a_1\rangle, \dots, |a_n\rangle\}$$

be a basis of \mathcal{V} and

$$\{|b_1\rangle, \dots, |b_n\rangle\}$$

a basis of \mathcal{W} . Then

$$T : \mathcal{V} \rightarrow \mathcal{W}, \quad \lambda_1|a_1\rangle + \dots + \lambda_n|a_n\rangle \mapsto \lambda_1|b_1\rangle + \dots + \lambda_n|b_n\rangle$$

defines a linear transformation. As it maps basis vector to basis vector it must be both injective and surjective. ■

Let \mathcal{V} be a vector space with an inner product. We can then think of the vectors in the dual space as the bra-vectors. The reason is as follows. Consider $|a\rangle \in \mathcal{V}$. Then $|a\rangle$ defines a linear functional f_a on \mathcal{V} by

$$f_a : \mathcal{V} \rightarrow \mathbb{C}, \quad |b\rangle \mapsto \langle a|b\rangle.$$

So that it is natural to think of f_a as $\langle a|$. And from now on we will take this intuitive notation $f_a = \langle a|$. The linear functionals span a subvector space of \mathcal{V}^* of same dimension as \mathcal{V} , so that by above proposition they are isomorphic. We even have that all linear functionals are of this form.

Theorem 6. Let \mathcal{V} be a vector space with inner product, then $\mathcal{V} \cong \mathcal{V}^*$, that is the vector space is isomorphic to its own dual.

Proof. We construct an isomorphism. Let $\{|a_1\rangle, \dots, |a_n\rangle\}$ be a basis of \mathcal{V} . Then we define

$$f_i : \mathcal{V} \rightarrow \mathbb{C}, \quad \lambda_1|a_1\rangle + \dots + \lambda_n|a_n\rangle \mapsto \lambda_i.$$

Again we inspect that this map is a linear transformation and clearly the f_1, \dots, f_n are linearly independent. We have to show that every linear functional is in the span of these f_i . For this let f be a linear functional with action on basis vectors $f(|a_i\rangle) = \mu_i$, then

$$f = \mu_1 f_1 + \dots + \mu_n f_n,$$

so that the map is also surjective and we have constructed the desired bijective linear transformation. ■

We can now define two useful maps

Definition 6. Let \mathcal{V} be a vector space over \mathbb{C} with inner product. Then the *dagger* is defined as

$$\dagger : \mathcal{V} \rightarrow \mathcal{V}^*, \quad |a\rangle \mapsto (|a\rangle)^\dagger := \langle a|.$$

Let \mathcal{V} and \mathcal{W} be two vector spaces over \mathbb{C} with inner products, and let $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$. Then the *pullback* $T^* \in \mathcal{L}(\mathcal{W}^*, \mathcal{V}^*)$ of T is the map

$$T^* : \mathcal{W}^* \rightarrow \mathcal{V}^*, \quad f \mapsto f \circ T.$$

2.3. Operators. Linear transformations from a vector space to itself form an algebra. Such linear transformations are called linear operators and they form the algebra of linear operators on the given vector space. Let $\mathcal{L}(\mathcal{V})$ be the algebra of linear operators on the \mathbb{C} -vector space \mathcal{V} with inner product. The identity is a linear transformation, so our algebra has a multiplicative identity. A linear operator that has a non-trivial kernel can not have an inverse, but every bijective linear transformation has, as it maps any basis of \mathcal{V} to another one. We denote the inverse of a linear operator T by T^{-1} and it satisfies

$$T \circ T^{-1} = T^{-1} \circ T = 1.$$

It is the unique operator with this property, since if there is another transformation R with $T \circ R = 1$ then multiplying this equation by the left by T^{-1} and using associativity uniqueness follows. In quantum mechanics, operators represent a physical quantity as energy, some charge or momentum. The measurement of this quantity of a physical state is then given by the

expectation value of this operator in the given state. It is defined as follows,

$$\langle T \rangle_a := \langle a | T | a \rangle$$

for a state $|a\rangle$ and an operator T . Since operators form an algebra, we can use addition and multiplication to define polynomials of operators. Using the inverse of operators we can further define rational functions in the operators. We can even define power series and Laurent series in the operators. These are then formal series, since we don't have a notion of convergence yet. An example is the exponential of an operator T , defined by the series expansion of the standard exponential function on the complex numbers,

$$\exp(T) := \sum_{k=0}^{\infty} \frac{T^k}{k!}.$$

But, operators do not necessarily commute. It is useful to define the commutator of two operators U, T in $\mathcal{L}(\mathcal{V})$ as

$$[T, U] := TU - UT.$$

The commutator is a bilinear map, that is a map from $\mathcal{L}(\mathcal{V}) \times \mathcal{L}(\mathcal{V})$ to $\mathcal{L}(\mathcal{V})$ that is linear in each component. It satisfies the properties

$$[U, T] = -[T, U] \quad (\text{antysymmetry})$$

and

$$[U, [S, T]] + [T, [U, S]] + [S, [T, U]] = 0 \quad \text{Jacobi identity}$$

for all U, S, T in $\mathcal{L}(\mathcal{V})$. A vector space with a bilinear map, that is antisymmetric and that satisfies the Jacobi identity is called a Lie algebra. So that we have

Theorem 7. Let \mathcal{V} be a vector space, then the commutator gives $\mathcal{L}(\mathcal{V})$ the structure of a Lie algebra.

A useful formula for the commutator of two exponentials of operators is the Baker-Campbell-Hausdorff formula:

$$\exp(t(U + S)) = \exp(tU)\exp(tS)\exp\left(-\frac{t^2}{2}[U, S]\right)\exp\left(\frac{t^3}{6}(2[S, [U, S]] + [U, [U, S]])\right) \cdot \dots$$

where the dots indicate exponentials of higher powers of the variable t .

There is a list of important types of operators.

Definition 7. Let \mathcal{V} be a vector space over \mathbb{C} with inner product, and let T in $\mathcal{L}(\mathcal{V})$, then the *adjoint or Hermitian conjugate* of T is denoted by T^\dagger in $\mathcal{L}(\mathcal{V}^*)$ and is defined by

$$\langle a | T | b \rangle^* = \langle b | T^\dagger | a \rangle, \quad \text{for all } |a\rangle, |b\rangle \text{ in } \mathcal{V}.$$

The adjoint is somehow an operator analogue of complex conjugation. The analogue to real numbers are then those operators that do not change under Hermitian conjugation.

Definition 8. A linear operator T in $\mathcal{L}(\mathcal{V})$ is *Hermitian or self-adjoint* if

$$\langle b | T^\dagger | a \rangle = \langle b | T | a \rangle, \quad \text{for all } |a\rangle, |b\rangle \text{ in } \mathcal{V}.$$

One often uses the slightly confusing short-hand notation $T^\dagger = T$. It is called *anti-Hermitian* if

$$\langle b | T^\dagger | a \rangle = -\langle b | T | a \rangle, \quad \text{for all } |a\rangle, |b\rangle \text{ in } \mathcal{V}.$$

And the short-hand notation is $T^\dagger = -T$.

So anti-Hermitian operators are the analogue of purely imaginary numbers. Indeed, by definition of the adjoint, the expectation values of Hermitian operators are real, while those of anti-Hermitian ones are purely imaginary. This statement is actually an if and only if statement.

Definition 9. An operator T in $\mathcal{L}(\mathcal{V})$ is called *positive* if all its expectation values are non-negative integers. We say that it is *positive definite* if all its expectation values are positive integers.

We use the notation $T \geq 0$ for a positive operator, and $T > 0$ for a positive definite one.

Definition 10. An operator T in $\mathcal{L}(\mathcal{V})$ is called *unitary* if

$$\langle a|b \rangle = \langle a|T^\dagger T|b \rangle$$

for all $|a\rangle, |b\rangle$ in \mathcal{V} .

Unitary means that the adjoint of an operator is the same as its inverse.

Let \mathcal{V} be a vector space with inner product, then to every vector $|a\rangle$ in \mathcal{V} we can associate an operator T_a via

$$T_a : \mathcal{V} \rightarrow \mathcal{V}, \quad |b\rangle \mapsto |a\rangle\langle a|b\rangle.$$

In physics one uses the notation

$$T_a := |a\rangle\langle a|,$$

which we will also adapt. What does this operator do? It maps a vector orthogonal to $|a\rangle$ to the zero vector, while every vector parallel to $|a\rangle$ is mapped to itself. If the norm of $|a\rangle$ is one, then it is thus the projection operator onto the subspace spanned by $|a\rangle$. Given a set of orthonormal vectors $\{|a_i\rangle\}$, the operator

$$\sum_{i=1}^m |a_i\rangle\langle a_i|$$

is then the projection operator on the sub vector space spanned by these m orthonormal vectors. Especially if these vectors form an orthonormal basis, this operator is just the identity,

$$\sum_{i=1}^m |a_i\rangle\langle a_i| = 1.$$

This relation is called the completeness relation.

You are surely familiar with matrices, and surely also with the concept that matrices represent linear transformations. So we will only very briefly recall this here. Let \mathcal{V} be a finite-dimensional complex vector space with inner product, and \mathcal{B} a basis, and T a linear operator on \mathcal{V} . Every basis allows a matrix representation of T . Call the basis vectors $|a_1\rangle, \dots, |a_n\rangle$, then the matrix of T with respect to the basis \mathcal{B} has components T_{ij} defined by

$$T(|a_i\rangle) = \sum_{j=1}^n T_{ij}|a_j\rangle.$$

A particular nice basis is an orthonormal one, so let \mathcal{B} be orthonormal, then the matrix entries have the nice form

$$\langle a_i|T|a_j\rangle = \sum_{k=1}^n T_{ik}\langle a_i|a_k\rangle = T_{ij}.$$

Let $T_{\mathcal{B}} = (T_{ij})$ be the matrix for a linear operator T in a given basis \mathcal{B} , then the matrix in the basis \mathcal{B} for the adjoint is the complex conjugate transpose of $T_{\mathcal{B}}$,

$$T_{\mathcal{B}}^\dagger = (T_{\mathcal{B}}^t)^* = (T_{ji}^*).$$

Matrices are then called Hermitian, unitary etc if they are the matrices with respect to some basis of Hermitian, unitary etc operators.

We have mentioned power series and Laurent series in linear operators. Using the matrix form of a linear operator in some basis, the matrix form of the Laurent or power series of the operator is then just the Laurent or power series of its matrix. It then can be said to converge if it converges component wise.

2.4. Eigenvectors and eigenvalues. Linear algebra has taught us, that a matrix is diagonalizable, if there exists a basis such that in this basis the matrix can have non-zero entries only on the diagonal. For a linear operator T on a vector space \mathcal{V} , this translates to the statement that T is diagonalizable if and only if there exists a basis of \mathcal{V} , such that each basis vector $|a_i\rangle$ satisfies $T|a_i\rangle = \lambda_i|a_i\rangle$ for some complex number λ_i called the eigenvalue of the eigenvector $|a_i\rangle$. The set of eigenvectors is usually called the spectrum of the operator. As I have the impression that you are all very familiar with these concepts, we will only repeat few important statements.

Theorem 8. Let \mathcal{V} be a vector space with inner product, and U, T two linear operators on \mathcal{V} . Then U and T are simultaneous diagonalizable, that is they possess a common basis of eigenvectors, if and only if they commute.

Simultaneous diagonalizable operators commute, since diagonal matrices do. The statement is then proven by looking at the eigenspace of U of a given eigenvalue and using commutativity to observe that the eigenspace is T -invariant. Using that distinct eigenspaces are orthogonal one can then show that every invariant subspace of a diagonalizable operator can be decomposed into eigenspaces, so that the theorem follows.

Lemma 9. Let \mathcal{V} be a vector space with inner product and let $\mathcal{V} = M \oplus M^\perp$ be an orthogonal decomposition. This means that every vector in M is orthogonal to every vector in M^\perp . Let T be a linear operator on \mathcal{V} , then M is invariant under T if and only if M^\perp is invariant under T^\dagger .

The proof can be left as an exercise.

Definition 11. An operator T in $\mathcal{L}(\mathcal{V})$ is called *normal* if it commutes with its adjoint.

Both Hermitian and unitary operators are normal.

Theorem 10. (spectral decomposition)

Let T in $\mathcal{L}(\mathcal{V})$ be a normal operator with spectrum $\{\lambda_1, \dots, \lambda_n\}$ and corresponding eigenspaces $\mathcal{V}_1, \dots, \mathcal{V}_n$. Then there exist non-zero mutually orthogonal projection operators $P_1 : \mathcal{V} \rightarrow \mathcal{V}_1, \dots, P_n : \mathcal{V} \rightarrow \mathcal{V}_n$, such that

$$P_1 + \dots + P_n = 1 \quad \text{and} \quad \lambda_1 P_1 + \dots + \lambda_n P_n = T.$$

In other words, every normal operator can be written as a linear combination of projection operators with coefficients the eigenvalues of the normal operator.

Proof. The intersection of two eigenspaces with distinct eigenvalue is trivial, hence the direct sum

$$M := \mathcal{V}_1 \oplus \dots \oplus \mathcal{V}_n$$

exists. We claim that $M = \mathcal{V}$. Let M^\perp be the orthogonal complement of M in \mathcal{V} . T leaves M invariant, since T commutes with its adjoint the same is true for T^\dagger . By above lemma T leaves M^\perp invariant. The eigenvalues of a linear operator are given by the roots of its characteristic

polynomial. The characteristic polynomial is a polynomial of degree the dimension of the vector space. But every polynomial of degree at least one has at least one root over \mathbb{C} . Hence M^\perp must be zero-dimensional. It follows that the sum of projection operator is the identity on $M = \mathcal{V}$. Further the operators $\lambda_1 P_1 + \dots + \lambda_n P_n$ and T coincide on every vector in $M = \mathcal{V}$ and hence must be the same. ■

The eigenvalues of an Hermitian operator are always real and those of unitary ones are always on the unit circle.

One advantage of diagonalizable operators is, that it is now easy to explicitly say what functions in this operator are. So let T in $\mathcal{L}(\mathcal{V})$ be a diagonalizable operator with spectrum $\{\lambda_1, \dots, \lambda_n\}$ and corresponding eigenspaces $\mathcal{V}_1, \dots, \mathcal{V}_n$ and projection operators $P_1 : \mathcal{V} \rightarrow \mathcal{V}_1, \dots, P_n : \mathcal{V} \rightarrow \mathcal{V}_n$. Let f be a function on the space of linear operators, that is for example a polynomial, or a rational function, or a power series or a Laurent series, then

$$f(T) = \sum_{i=1}^n f(\lambda_i) P_i.$$

Now, also the notion of convergence makes sense, and we can say that $f(T)$ converges in a basis consisting of eigenvectors if f converges at all eigenvalues.

2.5. Examples.

Quantum Chromodynamics and $SU(3)$. Those of you who have taken a standard model or quantum field theory course will know that quantum chromodynamics, that is the strong force, is an $SU(3)$ -gauge theory. We will put the objects appearing there in the framework of this section. Let $\mathcal{V} = \mathbb{C}^3$. The Lie group $SL(3; \mathbb{C})$ has \mathcal{V} as defining representation. This means there is a homomorphism

$$\rho : SL(3; \mathbb{C}) \rightarrow \{M \in \text{Mat}_3(\mathbb{C}) \mid \det(M) = 1\}.$$

Warning: In mathematics one distinguishes between the abstract Lie group from the Lie group of matrices in its defining representation. In physics one doesnot, as the two are isomorphic. There is then the risk that one confuses representations of the Lie group with the Lie group itself. In any case you can picture the Lie group $SL(3; \mathbb{C})$ as the group of unit determinant three by three matrices. in the same sense you should picture the unitary real form $SU(3)$ as the subgroup of unitary three by three matrices of determinant one, that is

$$SU(3) \cong \left\{ M \in \text{Mat}_3(\mathbb{C}) \mid \det(M) = 1; M^\dagger = M^{-1} \right\}$$

Let $X = \mathbb{R}^{3|1}$ be our four-dimensional Minkowski space-time. A quantum field is then an operator valued map from X , where the operators themselves act on some infinite-dimensional vector space. We donot want to be concerned with any details of that here. The point of this example is, that the quantum fields of the three quarks, can be organized in a vector

$$\Psi_{\text{quark}}(x) = \begin{pmatrix} \Psi_{\text{red}}(x) \\ \Psi_{\text{green}}(x) \\ \Psi_{\text{blue}}(x) \end{pmatrix}$$

whose components correspond to the three colored quarks. Quantum chromodynamics also has eight gauge fields, the gluons. But $SU(3)$ is an eight-dimensional real Lie group and the eight gluons correspond to a choice of basis of the underlying Lie algebra, called the Gell-Mann matrices $\lambda_1, \dots, \lambda_8$.

In summary, here we see that the quantum fields representing quarks are vector (\mathbb{C}^3) valued objects carrying a representation of the gauge group $SU(3)$. The gauge particles themselves, the gluons, are also vector (or matrix if you wish) ($SU(3)$)-valued objects acting (gauge transformation) on the quarks.

The one-dimensional harmonic oscillator. This is one of the first and instructive problems of a quantum mechanics course. We consider the one-dimensional harmonic oscillator of mass m and frequency ω . The space of physical states are square integrable real-valued functions in one variable. This is an infinite-dimensional vector space, and hence this problem already leads us to what we will deal with the coming lectures. The space and momentum operators x and p act on the space of complex-valued functions in one variable by multiplication with the variable x respectively by $-i\hbar \frac{d}{dx}$. The time-independent Schrödinger equation is

$$H\psi(x) = \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{m\omega^2}{2} x^2 \right) \psi(x) = E\psi(x)$$

It is useful to define the characteristic length

$$x_0 := \sqrt{\frac{\hbar}{m\omega}}.$$

the Schrödinger equation is a differential equation, that might be solved using standard methods. If desired, we can have a look at that later in the course. But there is also an algebraic way to find the eigenfunctions. Let

$$a = \frac{1}{\sqrt{2}} \left(\frac{x}{x_0} + x_0 \frac{d}{dx} \right), \quad a^\dagger = \frac{1}{\sqrt{2}} \left(\frac{x}{x_0} - x_0 \frac{d}{dx} \right).$$

The commutation relations are $[a, a^\dagger] = 1$, and the Hamilton operator becomes

$$H = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right).$$

So that finding the eigenstates of H amounts to finding the eigenstates of $n = a^\dagger a$. Let ψ_ν be an eigenfunction for n with eigenvalue ν and let us denote the inner product on square-integrable functions by (\cdot, \cdot) , then

$$\nu(\psi_\nu, \psi_\nu) = (\psi_\nu, a^\dagger a \psi_\nu) = (a\psi_\nu, a\psi_\nu) \geq 0.$$

the smallest possible eigenvalue is thus $\nu = 0$. Positive definiteness of the inner product implies that

$$0 = a\psi_0(x) = \frac{1}{\sqrt{2}} \left(\frac{x}{x_0} + x_0 \frac{d}{dx} \right) \psi_0(x).$$

The solution of this equation with norm one is

$$\psi_0 = \frac{1}{\sqrt{\sqrt{\pi}x_0}} e^{-\frac{1}{2} \left(\frac{x}{x_0} \right)^2}.$$

Using $[n, a^\dagger] = a^\dagger$ and $[n, a] = -a$, it is straight forward that $a^\dagger \psi_\nu$ is an eigenfunction with eigenvalue $\nu + 1$, while $a\psi_\nu$ is one with eigenfunction $\nu - 1$. The first property allows to construct eigenfunctions with eigenvalues positive integers starting from ψ_0 . The second property together with the positive definiteness property of the inner product of square integrable functions allows to show that these are all.

2.6. Exercises.

- (1) Let \mathcal{V} be a vector space with inner product. Let $|a\rangle$ and $|b\rangle$ be two vectors. $|b\rangle$ can be uniquely written as a sum of a vector parallel to $|a\rangle$ and another one orthogonal to $|a\rangle$. The vector parallel to $|a\rangle$ is called the projection of $|b\rangle$ onto $|a\rangle$. Show that the projection of $|b\rangle$ onto $|a\rangle$ is given by the formula

$$\frac{\langle a|b\rangle}{\langle a|a\rangle}|a\rangle.$$

- (2) Let \mathcal{V} be a finite-dimensional vector space, show that $\mathcal{L}(\mathcal{V})$ is associative, that is for each three elements A, B, C in $\mathcal{L}(\mathcal{V})$,

$$A \circ (B \circ C) = (A \circ B) \circ C$$

holds.

- (3) Give an example of operators on a vector space that do not commute.
 (4) Show that an operator is Hermitian if and only if all its expectation values are real.
 (5) Proof Lemma 9.
 (6) Give an example of a non-diagonalizable operator on a finite-dimensional vector space.

2.7. Solutions.

- (1) We write $|b\rangle = |p(b)_a\rangle + |b\rangle^\perp$. Here $|p(b)_a\rangle$ is the projection of $|b\rangle$ onto $|a\rangle$ and $|b\rangle^\perp$ is orthogonal to $|a\rangle$. Since $|p(b)_a\rangle$ is parallel to $|a\rangle$ there is a λ with $|p(b)_a\rangle = \lambda|a\rangle$. Taking the inner product of $|a\rangle$ with $|b\rangle$, we get

$$\langle a|b\rangle = \lambda\langle a|a\rangle,$$

so that the claim follows.

- (2) \mathcal{V} is a finite-dimensional vector space. Choose a basis \mathcal{B} of \mathcal{V} , and let $A_{\mathcal{B}}, B_{\mathcal{B}}, C_{\mathcal{B}}$ the matrices of A, B, C with respect to this basis. Matrix multiplication is associative (write down the product of three matrices with the indices and use associativity of multiplication in \mathbb{C}).
 (3) Choose two random $n \times n$ matrices, most likely they won't commute. Or the operators x and $\frac{d}{dx}$ acting on a vector space of polynomials in one variable.
 (4) We already noted in the lecture, that the expectation value of a Hermitian operator is real. By the spectral decomposition theorem a Hermitian operator is diagonalizable and eigenspaces are orthogonal and in fact form an orthogonal basis. Say this basis is $\{|a_1\rangle, \dots, |a_n\rangle\}$ with eigenvalues $\{\lambda_1, \dots, \lambda_n\}$. Assume the statement is true for every eigenvector $|a_i\rangle$, that is λ_i is real. Then the expectation value of an arbitrary vector $|a\rangle = \mu_1|a_1\rangle + \dots + \mu_n|a_n\rangle$ is

$$\langle T \rangle_a = |\mu_1|^2 \lambda_1 + \dots + |\mu_n|^2 \lambda_n$$

is also real. It thus remains to show that all λ_i are real. We have

$$\langle T \rangle_{a_i} = \langle a_i|T|a_i\rangle = \lambda_i \langle a_i|a_i\rangle,$$

but also

$$\langle T \rangle_{a_i} = \langle a_i|T|a_i\rangle = \langle a_i|T^\dagger|a_i\rangle^* = \langle a_i|T|a_i\rangle^* = \lambda_i^* \langle a_i|a_i\rangle^* = \lambda_i^* \langle a_i|a_i\rangle$$

so that the claim $\lambda_i = \lambda_i^*$ follows.

- (5) Assume that T leaves M invariant, that is for all $|a\rangle$ in M , we have $T|a\rangle$ in M . This means for all $|b\rangle$ in the orthogonal complement of M , that is in M^\perp , we have

$$\langle b|T|a\rangle = 0,$$

but this is the same as

$$\langle a|T^\dagger|b\rangle = 0$$

and hence for all $|b\rangle$ in M^\perp also $T^\dagger|b\rangle$ is orthogonal to M and hence in M^\perp . The converse direction is exactly the same argument but interchanging M and M^\perp as well as T and T^\dagger .

- (6) Let $\mathcal{V} = \mathbb{C}^2$ with basis v and w . The operator J , that maps v to w and w to zero has Jordan normal form

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

in this basis. It cannot be diagonalized.

3. HILBERT SPACES

Hilbert spaces appear in many problems of physics and mathematics. You have probably first heard about it in quantum mechanics, and we have just seen as an example the one-dimensional quantum mechanical harmonic oscillator.

In mechanics, one is interested in a physical state, an object for example, its properties as energy, momentum, position, etc.; and how it changes in time. We thus have three fundamental properties of a mechanical system: states, observables and time evolution.

In classical mechanics, the states of a physical system are the elements of the phase space; observables are functions on the space of physical states; and time evolution is given by a path or flow in the phase space. In quantum mechanics, the situation is quite different, and the data is summarized as follows

- (1) The space of physical states is associated a separable (that is countable basis), complex Hilbert space \mathcal{H} . The physical states are represented by vectors in the Hilbert space subject to the identification of two vectors that only differ by a phase. These equivalence classes are known as rays in the Hilbert space.
- (2) Each real physical observable is associated with a Hermitian operator T on \mathcal{H} . The only possible measurement is the eigenvalue λ_a of a (normalized) eigenstate $|a\rangle$ of T in \mathcal{H} . For a given arbitrary state $|b\rangle$ in \mathcal{H} , the probability of measuring λ_a is given by $|\langle a|b\rangle|^2$. The expectation value of a measurement, that is the weighted average of possible measurements is $\langle b|T|b\rangle$.
- (3) Symmetries of the physical system are described by unitary operators on \mathcal{H} . Especially the time evolution if the system is given by a one-parameter family of unitary transformations $U(t)$ for t in \mathbb{R} the time-variable. If H is the time-independent Hamiltonian of the physical system, this is a special Hermitian operator whose eigenvalues correspond to the possible energies of states in \mathcal{H} , then

$$U(t) := \exp\left(-\frac{i}{\hbar}tH\right). \quad (3.1)$$

You all have learnt Schrödinger's uncertainty principle, that two observables can only be measured simultaneously if the corresponding operators commute. Given a state $|a\rangle$ in our Hilbert space \mathcal{H} and two operators T, U on \mathcal{H} that do not commute. If we first measure the observable

associated to T , we will get some eigenvalue λ of T . If we repeat the measurement, we will always get the same answer. So the measurement of the observable projects the state onto an eigenstate (of eigenvalue λ) $|b\rangle$ of the associated operator T . But if we now want to measure the observable for U , we will project on an eigenstate of U . Since U and T donot commute such an eigenstate needs not to be one of T and a subsequent measurement of the observable of T will give an answer that needs not to be λ again.

After this short motivation, let's turn to defining a Hilbert space.

3.1. The definition of a Hilbert space. As we like to deal with infinite-dimensional vector spaces, we might like to study infinite sums and thus we need a notion of convergence. So let in this section \mathcal{H} be a possibly infinite-dimensional complex vector space with inner product. But we would like \mathcal{H} to be separable, that is have a countable basis.

Definition 12. A norm on \mathcal{H} is a map $\| \cdot \| : \mathcal{H} \rightarrow \mathbb{R}$ with

- (1) $\| |a\rangle \| \geq 0$ for all $|a\rangle$ in \mathcal{H} ; and $\| |a\rangle \| = 0$ if and only if $|a\rangle = |0\rangle$;
- (2) $\| \lambda |a\rangle \| = |\lambda| \| |a\rangle \|$ for all λ in \mathbb{C} and all $|a\rangle$ in \mathcal{H} ;
- (3) $\| |a\rangle + |b\rangle \| \leq \| |a\rangle \| + \| |b\rangle \|$ for all $|a\rangle, |b\rangle$ in \mathcal{H} .

Proposition 11. Let \mathcal{H} be a possibly infinite-dimensional complex vector space with inner product. Then

$$\| |a\rangle \| := \sqrt{\langle a|a\rangle} \tag{3.2}$$

defines a norm on \mathcal{H} .

Proof. The first two properties follow directly from the inner product properties. For the triangle inequality, recall the Schwarz inequality and we compute

$$\begin{aligned} (\langle a| + \langle b|)(|a\rangle + |b\rangle) &= \langle a|a\rangle + \langle b|b\rangle + \langle a|b\rangle + \langle a|b\rangle^* \\ &\leq \langle a|a\rangle + \langle b|b\rangle + 2|\langle a|b\rangle| \\ &\leq \langle a|a\rangle + \langle b|b\rangle + 2\| |a\rangle \| \| |b\rangle \| = (\| |a\rangle \| + \| |b\rangle \|)^2 \end{aligned}$$

so that the statement follows by taking the square root on both sides of the equation. ■

Having a norm, we can introduce the notion of convergence. Given a sequence of points a_1, a_2, \dots in a set \mathcal{H} with norm d , we say that the sequence converges if there is another point a in \mathcal{H} , such that for every $\varepsilon > 0$ there is an N_0 with for all $n > N_0$ we have $d(a, a_n) < \varepsilon$. We actually need a more restrictive type of convergence, the Cauchy sequence

Definition 13. Let \mathcal{H} be a set with norm d , then a sequence $\{a_n | n \in \mathbb{Z}_{\geq}; a_n \in \mathcal{H}\}$ is called *Cauchy sequence* if for every $\varepsilon > 0$ there exists N_0 such that for all $n, m > N_0$ we have

$$d(a_n, a_m) < \varepsilon.$$

One can show that every convergent sequence is a Cauchy sequence using the triangle inequality. But the converse is noy necessarily true.

Example 1. The reason a Cauchy-sequence might not converge in a set \mathcal{H} is that the limit might not be an element of this set. For example, consider a Cauchy sequence in the rational numbers, that converges to a non-rational one. Such a sequence is thus non-convergent because the set, the rational numbers, is somehow incomplete.

Definition 14. A set \mathcal{H} with norm d is called *complete* if every Cauchy sequence converges in this set. If \mathcal{H} is a vector space with inner product and the norm is given by the inner product as in (3.2), then if \mathcal{H} is complete, then it is called a *Hilbert space*.

We already saw that the rational numbers are not complete, but both real and complex numbers are (with norm the absolute value). This generalizes to finite-dimensional inner product spaces

Proposition 12. Every Cauchy sequence in a finite-dimensional inner product space, with norm defined by (3.2) over the real or complex numbers is convergent. This means, that every finite-dimensional vector space (over the real or complex numbers) with inner product is complete with respect to the norm given by (3.2).

The proof is postponed to the exercises. You can find it in many textbooks as for example [H] page 217. For infinite-dimensional inner product spaces over a complete field (like \mathbb{C} or \mathbb{R}) it is in general difficult to decide whether a space is Hilbert or not. An example is given on page 217 of [H]. Another one is

Example 2. Let X be the space of sequences $\{a_1, \dots | a_i \in \mathbb{R}; \text{ all but finitely many } a_i = 0\}$ with almost all entries zero, and finitely many ones non-zero. Let $a = \{a_1, \dots\}$ and $b = \{b_1, \dots\}$ be such two sequences, then an inner product is defined by

$$(a, b) = \sum_{i=1}^{\infty} a_i b_i.$$

This gives X the structure of an inner product space. A orthonormal basis of this space is given by $\{e(1), e(2), \dots\}$ where $e(n)$ is the sequence that has zeros everywhere, except its n -th component is one. We can then construct a sequence of sequences $\{F_1, \dots\}$ in X via

$$F_1 = \frac{1}{2}e(1), \quad F_n = F_{n-1} + \frac{1}{2^n}e(n).$$

The sequence F_n has first n components non-zero and zeros otherwise. So that its limit (it exists, since the sum $\sum 2^{-n}$ converges) has non-zero entries everywhere, which is a sequence that is not an element of X .

In this example, I have already called the $e(n)$ a basis (even though X is not a Hilbert space). So let us see, what we mean by a basis.

Proposition 13. Let \mathcal{H} be a infinite-dimensional Hilbert space over \mathbb{C} , and let $\{|e_1\rangle, \dots\}$ be an infinite ordered set of orthonormal vectors in \mathcal{H} . Let $|a\rangle$ in \mathcal{H} and define complex numbers $a_i := \langle e_i | a \rangle$. Then the *Bessel inequality*

$$\sum_{i=1}^{\infty} |a_i|^2 \leq \langle a | a \rangle$$

holds.

Proof. Define a sequence of vectors

$$|a_n\rangle := \sum_{i=1}^n a_i |e_i\rangle.$$

The Schwarz inequality proven in last section gives

$$|\langle a | a_n \rangle|^2 \leq \langle a | a \rangle \langle a_n | a_n \rangle = \langle a | a \rangle \sum_{i=1}^n |a_i|^2. \quad (3.3)$$

The inner product of $|a_n\rangle$ with $|a\rangle$ is

$$\langle a|a_n\rangle = \sum_{i=1}^n a_i \langle a|e_i\rangle = \sum_{i=1}^n a_i a_i^* = \sum_{i=1}^n |a_i|^2.$$

Substituting this identity into (3.3) gives

$$\sum_{i=1}^n |a_i|^2 \leq \langle a|a\rangle.$$

That is the sequence $\{A_n = \sum_{i=1}^n |a_i|^2 | n \in \mathbb{Z}_{>0}\}$ of non-negative real numbers is monotonously growing and bounded from above, hence convergent. The limit must satisfy the same inequality. ■

It follows that the vector

$$\sum_{i=1}^{\infty} a_i |e_i\rangle := \lim_{n \rightarrow \infty} \sum_{i=1}^n a_i |e_i\rangle$$

converges in the sense that it has finite norm. The question is whether this vector coincides with $|a\rangle$?

Definition 15. A sequence of orthonormal vectors $\{|e_1\rangle, \dots\}$ in a Hilbert space \mathcal{H} is called *complete* if the only vector that is orthogonal to all $|e_i\rangle$ is the zero vector. In this case $\{|e_1\rangle, \dots\}$ is called a *basis* of \mathcal{H} .

This definition is justified by

Proposition 14. Let $\{|e_1\rangle, \dots\}$ be an orthonormal sequence of a Hilbert space \mathcal{H} . Then the following statements are equivalent

- (1) $\{|e_1\rangle, \dots\}$ is complete.
- (2)

$$|a\rangle = \sum_{i=1}^{\infty} |e_i\rangle \langle e_i|a\rangle, \quad \text{for all } |a\rangle \text{ in } \mathcal{H}.$$

- (3)

$$\sum_{i=1}^{\infty} |e_i\rangle \langle e_i| = 1.$$

- (4)

$$\langle a|b\rangle = \sum_{i=1}^{\infty} \langle a|e_i\rangle \langle e_i|b\rangle, \quad \text{for all } |a\rangle, |b\rangle \text{ in } \mathcal{H}.$$

- (5)

$$\| |a\rangle \|^2 = \sum_{i=1}^{\infty} |\langle e_i|a\rangle|^2, \quad \text{for all } |a\rangle \text{ in } \mathcal{H}.$$

Proof. (1) \rightarrow (2) : Let

$$|c\rangle = |a\rangle - \sum_{i=1}^{\infty} |e_i\rangle \langle e_i|a\rangle,$$

then $|c\rangle$ is orthogonal to all $|e_i\rangle$ and hence must be the zero vector in a complete Hilbert space.

(2) \rightarrow (3) : We have that

$$1|a\rangle = |a\rangle = \sum_{i=1}^{\infty} |e_i\rangle \langle e_i|a\rangle$$

for all $|a\rangle$ in \mathcal{H} . That means

$$\sum_{i=1}^{\infty} |e_i\rangle\langle e_i|$$

acts as the identity on every vector in the Hilbert space and hence is the identity operator.

(3) \rightarrow (4) :

$$\langle a|b\rangle = \langle a|1|b\rangle = \sum_{i=1}^{\infty} \langle a|e_i\rangle\langle e_i|b\rangle$$

for all $|a\rangle, |b\rangle$ in \mathcal{H} .

(4) \rightarrow (5) : This is statement (4) with $|a\rangle = |b\rangle$.

(5) \rightarrow (1) : Let $|a\rangle$ be orthogonal to all $|e_i\rangle$, then by (5) the norm of $|a\rangle$ must be zero. But the only vector of zero norm in an inner product space is the zero vector itself. ■

The equality

$$\| |a\rangle \|^2 = \sum_{i=1}^{\infty} |\langle e_i|a\rangle|^2 = \sum_{i=1}^{\infty} |a_i|^2, \quad a_i := \langle e_i|a\rangle$$

is called the *Parseval equality*, the complex numbers a_i are called the *generalized Fourier coefficients* and the relation

$$\sum_{i=1}^{\infty} |e_i\rangle\langle e_i| = 1$$

is called the *completeness relation*. The definition of a Hilbert space has been rather abstract. Interestingly, there is a very concrete way to picture all separable infinite-dimensional Hilbert spaces over the real or complex numbers.

3.2. Square integrable functions. The set of square integrable functions on a real interval $[a, b]$ is denoted by $\mathcal{L}_{\omega}^2(a, b)$. Here \mathcal{L} is due to Lebesgue, who is responsible for understanding functions that are not continuous. ω is the weight function, this is a strictly positive and hence real valued function on the interval. $\mathcal{L}_{\omega}^2(a, b)$ is then the set of all functions

$$f : [a, b] \rightarrow \mathbb{C}$$

(you can replace \mathbb{C} by \mathbb{R} if you wish) with finite norm, defined by the inner product given by the Lebesgue integral

$$(f, g) := \int_a^b f(x)^* g(x) \omega(x) dx. \quad (3.4)$$

A famous theorem due to Riesz and Fischer is

Theorem 15. The space $\mathcal{L}_{\omega}^2(a, b)$ is a separable Hilbert space and all separable complex Hilbert spaces are isomorphic to a Hilbert space of square integrable functions on some interval with some weight function.

This is really nice, it tells us that studying separable Hilbert spaces, that is Hilbert spaces with countable basis is the same as studying spaces of square integrable functions. Finding a basis is due to a theorem by Stone and Weierstrass

Theorem 16. The sequence of monomials $\{x, x^2, x^3, \dots\}$ forms a basis of any $\mathcal{L}_{\omega}^2(a, b)$.

We are not completely happy yet, as we are usually in physics interested in an orthonormal basis (possibly corresponding to the eigenvectors of some important operator).

3.3. Classical orthogonal polynomials. The following theorem is a generalization of the Gram-Schmid process to a wide class of Hilbert spaces of square integrable functions.

Theorem 17. Define a sequence of functions on the interval $[a, b]$

$$F_n(x) := \frac{1}{\omega(x)} \frac{d^n}{dx^n} (\omega(x)s(x)^n), \quad (3.5)$$

with

- (1) $F_1(x)$ is a polynomial of degree one.
- (2) $s(x)$ is a polynomial of degree at most two with only real roots.
- (3) $\omega(x)$ is a strictly positive function on the interval (a, b) with the boundary conditions $\omega(a)s(a) = \omega(b)s(b) = 0$.

Then $F_n(x)$ is a polynomial of degree n and it is orthogonal to any polynomial $p_k(x)$ of degree $k < n$, that is

$$\int_a^b p_k(x) F_n(x) \omega(x) dx = 0, \quad \text{for } k < n.$$

These polynomials are called *classical orthogonal polynomials*.

This theorem is particularly useful if $s(x)$ is just a polynomial of degree one. In that case define

$$G_n(x) := \omega(x) F_n(x). \quad (3.6)$$

Then there is the obvious recurrence relation

$$G_n(x) = \frac{d}{dx} \frac{d^{n-1}}{dx^{n-1}} (\omega(x)s(x)^{n-1}s(x)) = \left(\frac{d}{dx} G_{n-1}(x) \right) s(x) + n G_{n-1}(x) \frac{d}{dx} s(x). \quad (3.7)$$

Proof. We first claim that the identity

$$\frac{d^m}{dx^m} (\omega(x)s(x)^n P_k(x)) = \omega(x)s(x)^{n-m} P_{k+m}(x) \quad (3.8)$$

holds for all $m \leq n$. Here P_k stands for an arbitrary polynomial of degree at most k . We fix n and proof the statement by induction to m . For $m = 0$ this is an honest identity. Using (3.5) with $n = 1$, we derive

$$s(x) \frac{d}{dx} \omega(x) = \omega(x) F_1(x) - \omega(x) \frac{d}{dx} s(x) = \omega(x) p_1(x) \quad (3.9)$$

since $s(x)$ is a polynomial of degree at most two and $F_1(x)$ one of degree one. We thus get for $m > 0$,

$$\begin{aligned} \frac{d^m}{dx^m} (\omega(x)s(x)^n P_k(x)) &= \frac{d}{dx} \left(\frac{d^{m-1}}{dx^{m-1}} (\omega(x)s(x)^n P_k(x)) \right) \\ &= \frac{d}{dx} (\omega(x)s(x)^{n-m+1} P_{k+m-1}(x)) \\ &= \omega(x)s(x)^{n-m} \left((n-m+1) \frac{d}{dx} s(x) P_{k+m-1}(x) + s(x) \frac{d}{dx} P_{k+m-1}(x) \right) + \\ &\quad + s(x) \frac{d}{dx} \omega(x)s(x)^{n-m} P_{k+m-1}(x) \\ &= \omega(x)s(x)^{n-m} P_{k+m}(x). \end{aligned}$$

So that the claim follows. Setting $k = 0$ and $P_0 = 1$ it follows that

$$\frac{d^m}{dx^m} (\omega(x)s(x)^n) \Big|_{x=a,b} = 0 \quad (3.10)$$

for all $m < n$. Using this equation together with integration by parts, it follows inductively (for $k < n$) that

$$\int_a^b P_k(x)F_n(x)\omega(x)dx = (-1)^k \int_a^b \left(\frac{d^k}{dx^k} P_k(x) \right) \frac{d^{n-k}}{dx^{n-k}} (\omega(x)s(x)^n) dx. \quad (3.11)$$

But the k -th derivative of a polynomial of degree at most k is a constant $(-1)^k C$, so that we get

$$\int_a^b P_k(x)F_n(x)\omega(x)dx = C \int_a^b \frac{d^{n-k}}{dx^{n-k}} (\omega(x)s(x)^n) dx = C \frac{d^{n-k-1}}{dx^{n-k-1}} (\omega(x)s(x)^n) \Big|_a^b = 0. \quad (3.12)$$

So that orthogonality of $F_n(x)$ to all polynomials of degree lower than n follows. It remains to prove that $F_n(x)$ is a polynomial of degree n . The identity (3.8) with $k = 0$ and $P_0 = 1$ and $m = n$ implies that F_n has at most degree n . So we write $F_n(x) = \alpha x^n + p_{n-1}(x)$. Multiplying both sides by $\omega(x)F_n(x)$ and integrating over our interval, we get

$$\int_a^b F_n(x)^2 \omega(x) dx = \int_a^b (\alpha x^n + p_{n-1}(x)) F_n(x) dx = \alpha \int_a^b x^n F_n(x) dx,$$

positive definiteness of the inner product implies that the left-hand side is non-zero if $F_n \neq 0$, and hence in that case α must be non-zero. In other words, in order to finish the proof, we have to show that $F_n \neq 0$. For this, observe that (3.11) also holds by replacing P_k by any square integrable function $f(x)$, so that with $f(x) = x^n$ and $k = n - 1$, we get

$$\int_a^b x^n F_n(x) \omega(x) dx = (-1)^k \int_a^b \left(\frac{d^{n-1}}{dx^{n-1}} x^n \right) \frac{d}{dx} (\omega(x)s(x)^n) dx = (-1)^k n! \int_a^b x \frac{d}{dx} (\omega(x)s(x)^n) dx.$$

Integration by parts gives

$$\int_a^b x \frac{d}{dx} (\omega(x)s(x)^n) dx = x \omega(x)s(x)^n \Big|_a^b - \int_a^b \omega(x)s(x)^n dx = - \int_a^b \omega(x)s(x)^n dx,$$

using (3.10). Positivity of the inner product forces this integral to be non-zero for even n , so that $F_n \neq 0$ for n even. Assume that there is an even n with $F_{n-1} = 0$. Recall the definition of G_n (3.6), then also $G_{n-1} = 0$, and by (3.7) we see that we get a contradiction to $F_n \neq 0$ if $s(x)$ is just a polynomial of degree at most one. Hence let $s(x) = \alpha x^2 + \beta x + \gamma$ with $\alpha \neq 0$. Then

$$\begin{aligned} G_n(x) &= \frac{d}{dx} \frac{d^{n-1}}{dx^{n-1}} (\omega(x)s(x)^{n-1}s(x)) = \frac{n(n-1)}{2} \left(\frac{d^{n-2}}{dx^{n-2}} \omega(x)s(x)^{n-1} \right) \frac{d^2}{dx^2} s(x) \\ &= \alpha n(n-1) \left(\frac{d^{n-2}}{dx^{n-2}} \omega(x)s(x)^{n-1} \right) \end{aligned} \quad (3.13)$$

but the derivative of the right-hand side is proportional to G_{n-1} and hence vanishes, it follows that G_n is constant, and hence F_n is proportional to $1/\omega(x)$. Since we already know that F_n is a polynomial of degree exactly n , the same must be true for $1/\omega(x)$, say $1/\omega(x) = p_n(x)$. Inserting this in the definition of G_{n-1} yields

$$0 = G_{n-1}(x) = \frac{d^{n-1}}{dx^{n-1}} (\omega(x)s(x)^{n-1}) = \frac{d^{n-1}}{dx^{n-1}} \left(\frac{s(x)^{n-1}}{p_n(x)} \right)$$

so that $\frac{s(x)^{n-1}}{p_n(x)}$ must be a polynomial of degree at most $n - 2$, and hence the roots of p_n must be a subset of those of $s(x)^{n-1}$. But s only has two roots (which may coincide) and since p_n is a polynomial of degree n , both roots must be roots of p_n . Positivity of $\omega(x)$ together with the boundary conditions $\omega(a)s(a) = \omega(b)s(b) = 0$ imply that p_n cannot have any roots in $[a, b]$.

This is a contradiction. So that in every case we get an contradiction to our assumption that $F_{n-1} = 0$. This completes the proof. ■

Having this theorem, we would like to know how to use it. There are four cases

- (1) $s(x)$ has no root, that means it is the constant polynomial;
- (2) $s(x)$ has one root with multiplicity one, that means it is a polynomial of degree exactly one;
- (3) $s(x)$ has two distinct roots;
- (4) $s(x)$ has one root with multiplicity two.

The first two cases have been studied by Hassani [H], so let's look at the third one. We would like to determine $\omega(x)$. Let

$$s(x) = (x - \alpha)(x - \beta), \quad (3.14)$$

so α and β are the two distinct real roots of $s(x)$. Recall that

$$F_1(x) = \frac{1}{\omega(x)} \frac{d}{dx} (\omega(x)s(x))$$

is a polynomial of degree one. Dividing both sides of this equation by $s(x)$ is

$$\frac{F_1(x)}{s(x)} = \frac{1}{\omega(x)s(x)} \frac{d}{dx} (\omega(x)s(x)) = \frac{d}{dx} \ln(\omega(x)s(x)).$$

We thus have to integrate the left-hand side in order to determine the weight function $\omega(x)$. For this, note the partial fraction decomposition

$$\frac{F_1(x)}{s(x)} = \frac{A}{x - \alpha} + \frac{B}{x - \beta}.$$

Here $A = F_1(\alpha)/(\alpha - \beta)$ and $B = -F_1(\beta)/(\alpha - \beta)$. (it is a short exercise to verify this identity). This fraction is the derivative of

$$\ln \left((x - \alpha)^A (x - \beta)^B \right),$$

so that up to a constant (which we can set to one), we get

$$\omega(x)s(x) = (x - \alpha)^A (x - \beta)^B$$

and hence

$$\omega(x) = (x - \alpha)^{A-1} (x - \beta)^{B-1}.$$

The boundary conditions $\omega(a)s(a) = \omega(b)s(b) = 0$ imply that the interval of integration is bounded by α and β and that $-1 < A, B$. Also note, that we can always translate and rescale the interval, so that without loss of generality $\alpha = -1$, $\beta = 1$.

The fourth case can be treated analogously. Let

$$s(x) = (x - \alpha)^2, \quad (3.15)$$

then a similar analysis reveals that

$$\omega(x) = e^{-\frac{A}{x-\alpha}} (x - \alpha)^{B-2},$$

with A and B defined by

$$\frac{F_1(x)}{s(x)} = \frac{A}{(x - \alpha)^2} + \frac{B}{x - \alpha}.$$

Translating our integration interval allows to set $\alpha = 0$. It turns out that it is impossible to find a positive weight function with the desired boundary conditions. It does not exist.

So that we can summarize

Theorem 18. The following data inserted in the previous theorem defines up to isomorphism all classical polynomials.

- (1) $s(x) = 1$: the weight function is $\omega(x) = e^{-x^2}$ and the interval is $[-\infty, \infty]$. The resulting polynomials are called *Hermite polynomials* and they are denoted by H_n ;
- (2) $s(x) = x$: the weight function is $\omega(x) = x^\nu e^{-x}$ with $\nu > -1$ and the interval is $[0, \infty]$. The resulting polynomials are called *Laguerre polynomials* and they are denoted by L_n^ν ;
- (3) $s(x) = (1-x)(1+x)$: the weight function is $\omega(x) = (1+x)^\mu(1-x)^\nu$ with $\nu, \mu > -1$ and the interval is $[-1, 1]$. The resulting polynomials are called *Jacobi polynomials* and they are denoted by $P_n^{\mu, \nu}$;
- (4) $s(x) = x^2$: this case does not exist.

There is a Schrödinger equation for these polynomials

Theorem 19. Let k and σ the coefficients $F_1(x) = kx + \dots$ and $s(x) = \sigma x^2 + \dots$. Then the Schrödinger equation

$$s(x) \frac{d^2}{dx^2} F_n(x) + F_1(x) \frac{d}{dx} F_n(x) = (kn + \sigma n(n-1)) F_n(x) \quad (3.16)$$

holds.

Proof. This will be an exercise. ■

One can ask a somehow inverse question, namely under which assumption does a differential equation of type

$$q(x) \frac{d^2}{dx^2} F_n(x) + \ell(x) \frac{d}{dx} F_n(x) = \lambda_n F_n(x) \quad (3.17)$$

admit solutions that are mutually orthogonal for distinct λ_n . A computation reveals that this is true if $\omega(x)q(x)$ vanishes sufficiently fast at the boundary of the interval of integration, and in addition the identity

$$\frac{d}{dx}(\omega(x)q(x)) = \ell(x)\omega(x)$$

is true.

3.4. Gegenbauer polynomials and hypergeometric functions. We now would like to discuss an example. The most prominent examples in physics are the Hermite polynomials in the one-dimensional quantum mechanical harmonic oscillator, and the Gegenbauer polynomials in problems with a spherical symmetric potential, as for example in electro-statics. The Gegenbauer polynomials are the special case $\mu = \nu$ of Jacobi polynomials. Let us define them as

$$C_n^\alpha(x) := P_n^{\alpha-1/2, \alpha-1/2},$$

so that $\alpha < -1/2$. The data for these polynomials is

$$s(x) = 1 - x^2, \quad \omega(x) = (1 - x^2)^{\alpha-1/2}, \quad C_1^\alpha(x) = (2\alpha + 1)x.$$

So that they satisfy the differential equation

$$(1 - x^2) \frac{d^2}{dx^2} C_n^\alpha(x) + (2\alpha + 1) \frac{d}{dx} C_n^\alpha(x) = (2(\alpha + 1)n - n^2) C_n^\alpha(x).$$

The importance of the Gegenbauer polynomials comes from the form of its generating function, that is

$$\frac{1}{(1 - 2xt + t^2)^\alpha} = \sum_{n=0}^{\infty} C_n^\alpha(x) t^n. \tag{3.18}$$

For example in spherical symmetric potentials, we have to deal with expressions of the form

$$\frac{1}{|\vec{r} - \vec{r}'|^{2\alpha}}$$

for two vectors \vec{r}, \vec{r}' . Let r, r' be the length of these vectors, and let ϑ be the angle between these two vectors, then an exercise in linear algebra shows (look at your first year linear algebra textbook) that

$$|\vec{r} - \vec{r}'|^2 = r^2 \left(1 + \frac{r'^2}{r^2} - 2 \frac{r'}{r} \cos \vartheta \right)$$

so that with $t = \frac{r'}{r}$, we get

$$\frac{1}{|\vec{r} - \vec{r}'|^{2\alpha}} = \sum_{n=0}^{\infty} \frac{r'^n}{r^{n+2}} C_n^\alpha(\cos \vartheta).$$

Another important property of these polynomials is their normalization, that is

$$\int_{-1}^1 C_n^\alpha(x) C_n^\alpha(x) (1 - x^2)^{\alpha - \frac{1}{2}} dx = 2^{1-2\alpha} \pi \frac{\Gamma(n + 2\alpha)}{(n + \alpha) \Gamma(\alpha)^2 \Gamma(n + 1)} \tag{3.19}$$

where the Γ function satisfies the relation $\Gamma(x + 1) = x\Gamma(x)$ with $\Gamma(1) = 1$. The Gegenbauer polynomials can be expressed in terms of hypergeometric functions

$$C_n^\alpha(x) = \binom{n + 2\alpha - 1}{n} {}_2F_1 \left(-n, n + 2\alpha, \alpha + \frac{1}{2}; \frac{1}{2}(1 - x) \right) \tag{3.20}$$

where the hypergeometric series is a function that appears frequently in various areas of physics and mathematics. It is

$${}_2F_1(\alpha, \beta, \gamma; x) = \sum_{n=0}^{\infty} \frac{(\alpha)_n (\beta)_n}{n! (\gamma)_n} x^n \tag{3.21}$$

for $|x| < 1$ and $(\alpha)_n$ is a short-hand notation

$$(\alpha)_n := \frac{\Gamma(\alpha + n)}{\Gamma(\alpha)}.$$

The relation to hypergeometric function actually holds to all Jacobi polynomials, it is

$$P_n^{\alpha, \beta}(x) = \frac{(\alpha + 1)_n}{n!} {}_2F_1(-n, \alpha + \beta + 1 + n, \alpha + 1; (1 - 2x)). \tag{3.22}$$

So that this class of classical polynomials is part of hypergeometric functions. The hypergeometric functions satisfy the (Euler hypergeometric) differential equation

$$x(1 - x) \frac{d^2}{dx^2} {}_2F_1(\alpha, \beta, \gamma; x) + (\gamma - (\alpha + \beta + 1)x) \frac{d}{dx} {}_2F_1(\alpha, \beta, \gamma; x) = \alpha\beta {}_2F_1(\alpha, \beta, \gamma; x). \tag{3.23}$$

Hypergeometric functions appear in many areas as in number theory, especially the theory of partitions of integers; in modular and elliptic functions and in conformal field theory and string theory. There a class of theories, called the Virasoro minimal models (but also many others) relates to hypergeometric functions. Especially correlation functions of four fields inserted at $0, 1, x$ and ∞ satisfy the differential equation (3.23). But these correlation functions are usually

computed using the integral identity

$${}_2F_1(\alpha, \beta, \gamma; x) = \frac{\Gamma(\gamma)}{\gamma\beta\Gamma(\gamma-\beta)} \int_0^1 y^{\beta-1} (1-y)^{\gamma-\beta-1} (1-xy)^{-\alpha} dy \quad (3.24)$$

which holds if the real part of γ is larger than the real part of β , which must be positive.

3.5. Hermite polynomials. Every classical orthogonal polynomial satisfies some nice differential and integral equation. We have seen such an equation for Jacobi polynomials in the last section. Here, I would like to present a similar equation for Hermite polynomial as well as some more properties of them. The Hermite polynomial can be represented by the following integral

$$H_n(x) = \frac{2^n (-i)^n e^{x^2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2 + 2itx} t^n dt. \quad (3.25)$$

They also satisfy an integral equation

$$e^{-\frac{x^2}{2}} H_n(x) = \frac{1}{i^n \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ixy} e^{-\frac{y^2}{2}} H_n(y) dy. \quad (3.26)$$

The use of orthogonal polynomials is that they form an orthogonal basis of square integrable functions. This means that every other square integrable function can be expressed in terms of this basis. For Hermite polynomials the more precise statement is

Theorem 20. Let $f(x)$ be in $\mathcal{L}_\omega(-\infty, \infty)$ for $\omega = e^{-x^2}$ and let f be smooth, then

$$f(x) = \sum_{n=0}^{\infty} c_n H_n(x)$$

with

$$c_n = \frac{1}{2^n n! \sqrt{\pi}} \int_{-\infty}^{\infty} n f(t) y \omega(x) H_n(x) f(x) dx.$$

In many cases it is possible to explicitly compute these coefficients.

Example 3. Let $f(x) = x^{2p}$ for $p = 0, 1, \dots$. Then the expansion of this monomial in terms of Hermite polynomials is

$$x^{2p} = \sum_{n=0}^p c_{2n} H_{2n}(x),$$

so that most coefficients in the expansion vanish. In order to compute the c_{2n} , we need the integral representation for the Γ function, that is

$$\Gamma\left(p - n + \frac{1}{2}\right) = \int_{-\infty}^{\infty} e^{-x^2} x^{2p-2n} dx$$

as well as the identity

$$2^{2p-2n} \Gamma\left(p - n + \frac{1}{2}\right) (p-n)! = \sqrt{\pi} (2p-2n)!.$$

Then by partially integrating (the second equation), we get

$$\begin{aligned}
 c_{2n} &= \frac{1}{2^{2n}(2n)!\sqrt{\pi}} \int_{-\infty}^i n f t y e^{-x^2} x^{2p} H_{2n}(x) dx \\
 &= \frac{1}{2^{2n}(2n)!\sqrt{\pi}} \int_{-\infty}^i n f t y e^{-x^2} x^{2p} \frac{d^{2n}}{dx^{2n}} \left(e^{-x^2} \right) dx \\
 &= \frac{1}{2^{2n}(2n)!\sqrt{\pi}} \frac{(2p)!}{(2p-2n)!} \int_{-\infty}^i n f t y e^{-x^2} x^{2p-2n} dx \\
 &= \frac{1}{2^{2n}(2n)!\sqrt{\pi}} \frac{(2p)!}{(2p-2n)!} \Gamma \left(p - n + \frac{1}{2} \right) \\
 &= \frac{(2p)!}{2^{2p}(2n)!(p-n)!}
 \end{aligned} \tag{3.27}$$

Similarly one can show for odd degree monomials that

$$x^{2p+1} = \frac{(2p+1)!}{2^{2p+1}} \sum_{n=0}^p \frac{1}{(2n+1)!(p-n)!} H_{2n+1}(x). \tag{3.28}$$

3.6. Exercises.

- (1) Proof Proposition 12.
- (2) Find the following differential equation for classical orthogonal polynomials. Let k and σ the coefficients $F_1(x) = kx + \dots$ and $s(x) = \sigma x^2 + \dots$. Then

$$\frac{1}{\omega(x)} \frac{d}{dx} \left(\omega(x)s(x) \frac{d}{dx} F_n(x) \right) = (kn + \sigma n(n-1)) F_n(x) \tag{3.29}$$

and

$$s(x) \frac{d^2}{dx^2} F_n(x) + F_1(x) \frac{d}{dx} F_n(x) = (kn + \sigma n(n-1)) F_n(x) \tag{3.30}$$

hold. Proceed in the following steps:

- Use (3.8) to prove

$$\int_a^b F_m(x) \frac{d}{dx} \left(\omega(x)s(x) \frac{d}{dx} F_n(x) \right) dx = 0$$

for $m < n$.

- Use (3.8) to show that

$$\frac{1}{\omega(x)} \frac{d}{dx} \left(\omega(x)s(x) \frac{d}{dx} F_n(x) \right)$$

is a polynomial of at most degree n .

- Use the first step together with the orthogonality of the F_n with respect to the inner product on $\mathcal{L}_{\omega}^2(a, b)$, that is

$$(F_m, F_n) = \int_a^b F_m(x) F_n(x) \omega(x) dx = 0$$

if $n \neq m$, to show that

$$\frac{1}{\omega(x)} \frac{d}{dx} \left(\omega(x)s(x) \frac{d}{dx} F_n(x) \right) = \gamma F_n(x) \tag{3.31}$$

for some proportionality constant γ .

- Use (3.9) to show that

$$\frac{d}{dx} \left(\omega(x)s(x) \frac{d}{dx} F_n(x) \right) = \omega(x)s(x) \frac{d^2}{dx^2} F_n(x) + \omega(x)F_1(x) \frac{d}{dx} F_n(x) \quad (3.32)$$

- We introduce some notation. Let α be the coefficient in front of x^n of F_n , that is $F_n(x) = \alpha x^n + \dots$. Let β be the norm of F_n , that is

$$(F_n, F_n) = \int_a^b F_n(x)F_n(x)\omega(x)dx = \beta.$$

Use this notation together with (3.31) and the previous step to prove (3.29).

- Combine the last two steps to (3.30).

3.7. Solutions.

- (1) We repeat the proof of [H] on page 216/217. Let $|a_1\rangle, \dots$ be a Cauchy sequence of vectors in a finite-dimensional vector space V . Let $|e_1\rangle, \dots, |e_n\rangle$ be an orthonormal basis of V , and define coefficients α_{ik} such that

$$|a_i\rangle = \sum_{k=1}^n \alpha_{ik} |e_k\rangle.$$

Then

$$\begin{aligned} \left\| |a_i\rangle - |a_j\rangle \right\|^2 &= \left\| \sum_{k=1}^n (\alpha_{ik} - \alpha_{jk}) |e_k\rangle \right\|^2 \\ &= \sum_{k=1}^n |\alpha_{ik} - \alpha_{jk}|^2. \end{aligned}$$

The left-hand side tends to zero for large i, j , as the sequence is Cauchy. So the same must be true for the right-hand side. But the right-hand side is a sum of non-negative summands. Hence each term must tend to zero. The complex numbers are complete, hence the limits

$$\lim_{i \rightarrow \infty} \alpha_{ik} := \alpha_k$$

exists. Let

$$|a\rangle = \sum_{k=1}^n \alpha_k |e_k\rangle,$$

then the cauchy sequence converges to this vector since

$$\lim_{i \rightarrow \infty} \left\| |a_i\rangle - |a\rangle \right\|^2 = \lim_{i \rightarrow \infty} \sum_{k=1}^n |\alpha_{ik} - \alpha_k|^2 = \sum_{k=1}^n \lim_{i \rightarrow \infty} |\alpha_{ik} - \alpha_k|^2 = 0.$$

- (2) We will prove this statement following the indicated steps.

- Equation (3.8) with $n = m = 1$ reads

$$\frac{d}{dx} (\omega(x)s(x)P_k(x)) = \omega(x)P_{k+1}(x) \quad (3.33)$$

for every polynomial P_k of degree k . We partially integrate twice and use the boundary behaviour (3.10) to get

$$\begin{aligned} \int_a^b F_m(x) \frac{d}{dx} \left(\omega(x)s(x) \frac{d}{dx} F_n(x) \right) dx &= - \int_a^b \left(\frac{d}{dx} F_m(x) \right) \omega(x)s(x) \frac{d}{dx} F_n(x) dx \\ &= \int_a^b \frac{d}{dx} \left(\left(\frac{d}{dx} F_m(x) \right) \omega(x)s(x) \right) F_n(x) dx. \end{aligned}$$

Using (3.33) with $k = m - 1$ and $P_{m-1} = \frac{d}{dx}F_m(x)$, we get

$$\begin{aligned} \int_a^b \left(\frac{d}{dx} \left(\frac{d}{dx} F_m(x) \right) \omega(x)s(x) \right) F_n(x) dx &= \int_a^b \frac{d}{dx} (P_{m-1}(x) \omega(x)s(x)) F_n(x) dx \\ &= \int_a^b P_m(x) \omega(x) F_n(x) dx = 0 \end{aligned}$$

for all $m < n$ since F_n is orthogonal to any polynomial of degree less than n .

- We again use (3.33), that is the special case of $n = m = 1$ of (3.8). Then for $P_{n-1} = \frac{d}{dx}F_n$, we get

$$\frac{1}{\omega(x)} \frac{d}{dx} \left(\omega(x)s(x) \frac{d}{dx} F_n(x) \right) = \frac{1}{\omega(x)} \omega(x) P_n(x) = P_n(x)$$

that indeed the expression is a polynomial P_n of degree at most n .

- The polynomial P_n of the last step is a polynomial of degree at most n , so we can expand it in terms of those orthogonal polynomials that have degree at most n , i.e.

$$P_n = \sum_{m=1}^n \gamma_m F_m.$$

The first step tells us that $(F_m, P_n) = 0$ for all $m < n$. The orthogonality of the F_m implies that $(F_m, P_n) = \gamma_m (F_m, F_m)$, and hence γ_m must be zero for $m < n$. The claim follows with $\gamma = \gamma_n$.

- We have

$$\frac{d}{dx} \left(\omega(x)s(x) \frac{d}{dx} F_n(x) \right) = \omega(x)s(x) \frac{d^2}{dx^2} F_n(x) + \frac{d}{dx} (\omega(x)s(x)) \frac{d}{dx} F_n(x).$$

Equation (3.9) can be rewritten as

$$\frac{d}{dx} (\omega(x)s(x)) = \omega(x) F_1(x)$$

so that the claim follows.

- We first compute the inner product of F_n with x^n . Since the inner product of F_n with any polynomial of lower degree vanishes, we have $(F_n, F_n - \alpha x^n) = 0$, and hence $(F_n, x^n) = \beta/\alpha$. The task is to compute γ , using the previous two steps we have

$$\begin{aligned} \beta\gamma &= \gamma(F_n, F_n) = (F_n, \gamma F_n) \\ &= \left(F_n, \frac{1}{\omega(x)} \frac{d}{dx} \left(\omega(x)s(x) \frac{d}{dx} F_n(x) \right) \right) \\ &= \left(F_n, s(x) \frac{d^2}{dx^2} F_n(x) + F_1(x) \frac{d}{dx} F_n(x) \right) \\ &= \alpha\sigma n(n-1)(F_n, x^n) + \alpha kn(F_n, x^n) \\ &= \beta(\sigma n(n-1) + kn). \end{aligned}$$

So that the claim follows. Here, we again used the orthogonality of F_n to polynomials of lower degree.

- Inserting now (3.32) in this differential equation proves the final statement.

4. HARMONIC ANALYSIS

So far, we have talked about Hilbert spaces of functions, which have a countable basis consisting of certain polynomials. We now would like to study larger Hilbert spaces and thus have to relax our notion of functions.

4.1. Motivation. Consider a Hilbert space of square integrable functions $\mathcal{H} = \mathcal{L}_\omega(a, b)$. We have learnt that we can expand any function in this space in a given basis, say $\{e_i, \dots\}$,

$$f(x) = \sum_{n=0}^{\infty} f_n e_n(x),$$

where the coefficients are given by

$$f_n = \int_a^b e_n^*(x) f(x) \omega(x) dx$$

if the basis is orthonormal. The expansion of a function is thus given (for an orthonormal basis) by an assignment

$$b : \mathcal{H} \times \mathbb{Z}_{\geq 0} \rightarrow \mathbb{C}, \quad (f, n) \mapsto f_n = \int_a^b e_n^*(x) f(x) \omega(x) dx \quad (4.1)$$

from \mathcal{H} times a set indexing the basis to the complex number. So that the expansion of our function f in the basis becomes

$$f(x) = \sum_{n=0}^{\infty} b(f, n) e_n(x).$$

Now we would like to replace our countable basis by a basis $\{e_y | y \in \mathcal{J}\}$ indexed by an uncountable set \mathcal{J} , then the situation should be analogous. There should be an assignment

$$b : \mathcal{H} \times \mathcal{J} \rightarrow \mathbb{C}, \quad (f, y) \mapsto f_y, \quad (4.2)$$

such that the element f can be expanded in terms of the uncountable basis

$$f = \int_{\mathcal{J}} f_y e_y dy.$$

Since the new basis is uncountable, we here have to replace the sum over the set of basis elements by an integral. There are now some obvious questions:

- What is a natural uncountable set for the space of square integrable functions?
- What is f_y ?
- What are the e_y ?

A candidate for the first question is the interval

$$\mathcal{J} = [a, b]$$

of integration. In order to proceed, we need the notion of orthogonality. We define

Definition 16. The Dirac delta distribution $\delta(y - x)$ is defined by requiring that for any square integrable function on $[a, b]$ the identity

$$\int_a^b f(y) \delta(y - x) dy = f(x)$$

holds for all $x \in [a, b]$.

Then we define a distribution valued inner product on the basis vector e_y by

$$\langle e_z | e_y \rangle = \int_a^b \omega(x) e_z^*(x) e_y(x) dx = \delta(y - z).$$

So that orthonormality should be interpreted in a distributional sense and the role of 1 is replaced by the Dirac delta distribution. With this notion, we can compute the expansion coefficients

$$\langle e_z | f \rangle = \int_a^b e_z^*(x) f(x) \omega(x) dx = \int_a^b \int_a^b e_z^*(x) f_y e_y(x) \omega(x) dy dx = \int_a^b \delta(z - y) f_y dy = f_z.$$

Here we interchanged the order of integrations. We assume this to be allowed. We observe, that the expansion coefficients are as in the seperable case determined by the inner product with the (in some distributional sense orthonormal) basis vectors. Now, let

$$g(x) = \int_a^b g_y e_y(x) dy$$

be another element of our Hilbert space. The inner product is then coefficient wise.

$$\langle f | g \rangle = \int_a^b \int_a^b \int_a^b f_y^* e_y^*(x) g_z e_z(x) \omega(x) dx dy dz = \int_a^b \int_a^b f_y^* g_z \delta(y - z) dy dz = \int_a^b f_y^* g_y dy$$

There is also a completeness relation. Consider

$$f(x) = f = \int_a^b f_y e_y(x) dy = \int_a^b e_y^*(z) \omega(z) e_y(x) f(z) dz dy$$

so that the identity opertor is the integral over

$$\delta(x - z) = \int_a^b e_y^*(z) \omega(z) e_y(x) dy. \tag{4.3}$$

4.2. Distributions. We have seen, that we need to learn about distributions in order to study bases of uncountable cardinality. Distributions act on suitable spaces of test functions. We choose our space of test functions to be $C^\infty(\mathbb{C})$. That is the space of smooth (infinitely differentiable functions) on \mathbb{R} with values in \mathbb{C} .

Definition 17. A *distribution* is a linear functional on $C^\infty(\mathbb{C})$, that means it maps every smooth function to a complex number.

Example 4. Let f be a sufficiently integrable function on the real line, then the distribution T_f is defined by

$$T_f(g) := \int_{\mathbb{R}} f(x) g(x) dx.$$

Usually one abuses notation and identifies the distribution corresponding to a function by the same name, that is

$$T_f = f.$$

We will adopt this notation.

Example 5. In the same sense, the Dirac delta distribution is defined by

$$\delta(f) = f(0)$$

and one abuses the notation as in the definition of last section, that is

$$\delta(f) = f(0) = \int_{\mathbb{R}} \delta(x) f(x) dx.$$

The derivative of a distribution ϕ is defined by partial integration

$$\int_{\mathbb{R}} \frac{d}{dx} \phi(x) f(x) dx = - \int_{\mathbb{R}} \phi(x) \frac{d}{dx} f(x) dx,$$

that is we assume that test functions vanish sufficiently fast at $\pm\infty$. This definition then implies that

$$\delta(x-y) = \frac{d}{dx} \theta(x-y) \tag{4.4}$$

where the Heavyside step function is defined as

$$\theta(x) := \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases}.$$

A useful way to represent the delta distribution is as a limit of a series of distributions associated to some functions.

Definition 18. Let $\{\phi_1(x), \dots\}$ be a series of functions such that

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \phi_n(x) f(x) dx$$

exists for all smooth functions $f(x)$. Then we say that the series of functions $\phi_1(x), \dots$ converges to the distribution $\phi(x)$ defined by this limit.

Example 6. One can show that the two sequences

$$\left\{ \frac{n}{\sqrt{\pi}} e^{-n^2 x^2} \right\}, \quad \left\{ \frac{\sin(nx)}{\pi x} \right\}$$

both converge to the Dirac delta distribution. A good idea to proof the statement is to proof it for a basis of the Hilbert space. But we will not go into the details here. We just note, that the second sequence can be used to derive the following integral representation of the Dirac delta distribution

$$\delta(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ixt} dt.$$

This follows, since the second sequence has the integral representation

$$\frac{\sin(nx)}{\pi x} = \frac{1}{2\pi} \int_{-n}^n e^{ixt} dt.$$

We will see another example soon in the context of Fourier series.

4.3. Fourier Analysis. We turn to functions on the circle S^1 of radius one. The circle is

$$S^1 = \{e^{i\alpha} \mid -\pi \leq \alpha < \pi\}.$$

So that we can identify functions on the circle with functions on the interval $[-\pi, \pi]$. We know that a good basis for the space of some square integrable functions are monomials. But there are also other possibilities.

Theorem 21. The Hilbert space of functions on the circle, $\mathcal{H} = \mathcal{L}(-\pi, \pi)$, has orthonormal basis $\{e_1, \dots\}$ given by

$$e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}.$$

This means that for every square integrable function $f \in \mathcal{H}$, there is an expansion

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{n \in \mathbb{Z}} f_n e^{inx}$$

with Fourier coefficients

$$f_n = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x)e^{-inx} dx.$$

We donot proof this theorem. Orthonormality of the e_n is a direct computation, while completeness requires a Stone-Weierstrass theorem in two variables, that then can be specialized to functions on the circle. Recall that in a finite-dimensional vector space, the identity operator is the sum of projection operators associated to a basis. In our case, the operator

$$P_n(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)e^{-iny} dy e^{inx} = f_n e_n(x)$$

projects f onto the subspace spanned by $e_n(x)$. Further, the completeness relation becomes

$$\sum_{n \in \mathbb{Z}} P_n = 1,$$

since

$$\sum_{n \in \mathbb{Z}} P_n(f) = \sum_{n \in \mathbb{Z}} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)e^{-iny} dy e^{inx} = \sum_{n \in \mathbb{Z}} f_n e_n(x) = f(x)$$

so that we get the representation of the delta distribution on the circle

$$\delta(x - y) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} dy e^{in(x-y)}.$$

Next, we ask what does it mean that a function converges to its Fourier series. Let $f \in \mathcal{H}$, then we say that its Fourier series

$$\frac{1}{\sqrt{2\pi}} \sum_{n \in \mathbb{Z}} f_n e^{inx}$$

converges in $\mathcal{H} = \mathcal{L}(-\pi, \pi)$ if the partial sums

$$f_N = \frac{1}{\sqrt{2\pi}} \sum_{n=-N}^N f_n e^{inx}$$

converge in the sense that

$$\lim_{N \rightarrow \infty} \|f_N - f\| = 0.$$

This means the difference of f_N and f becomes a function of measure zero. If we allow for piecewise continuous functions, then at the points of discontinuity the difference between original function and its Fourier series is given by the following theorem.

Theorem 22. The Fourier series of a piecewise continuous function $f(x)$ on $[-\pi, \pi]$ converges pointwise to

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{2} (f(x + \varepsilon) + f(x - \varepsilon))$$

at $x \in (-\pi, \pi)$ and to

$$\frac{1}{2} (f(\pi) + f(-\pi))$$

at $x = \pm\pi$.

Fourier analysis on the circle is the simplest example of harmonic analysis on a Lie group. So let us express everything we know in an unusual language.

4.4. Harmonic analysis on the Lie group $U(1)$. The real unitary Lie group $U(1)$ is the one-dimensional real manifold S^1 . We can parameterize elements in S^1 by

$$e^{i\alpha}, \quad -\pi \leq \alpha < \pi.$$

Then the product

$$e^{i\alpha} e^{i\beta} = e^{i(\alpha+\beta)}$$

defines a group structure on S^1 , the Lie group $U(1)$. The tangent space at a point is nothing but a line. An infinitesimal translation on the line $\ell = a + bx$ is given by the derivative $\frac{d}{dx}$. The operator of infinitesimal translations on the tangent space of a Lie group is called invariant vector field. $U(1)$ is a commutative Lie group, in which case there is only one type of invariant vector field. In the general non-commutative case there are both left- and right-invariant vector fields, corresponding to infinitesimal left and right translations. The invariant vector fields satisfy the commutation relations of the Lie algebra of the Lie group. The Lie algebra of $U(1)$ is called $u(1)$ and it is the one-dimensional abelian Lie algebra. Indeed, the vector field $\frac{d}{dx}$ commutes with itself. Harmonic analysis is then the study of square integrable functions on the Lie group. These functions respect the Lie group symmetry in the sense, that they are representations of the Lie algebra of vector fields. In our case, the square integrable functions on the Lie group are best described as periodic functions on the real line. We then saw that a basis is given by the e^{inx} , and indeed each basis vector carries a one-dimensional representation of $u(1)$ given by

$$\frac{d}{dx} e^{inx} = in e^{inx}.$$

In the non-commutative setting, good basis elements won't be invariant under the Lie algebra of vector fields, but they will carry a nice, that means irreducible, representation of the Lie algebra. But there is another operator, called the Casimir or Laplace operator. It is a second-order differential operator, that is invariant under conjugation by a Lie group element. In our case, this is simply

$$\Delta = \frac{d}{dx} \frac{d}{dx}.$$

Conjugating by a Lie group element is rather trivial

$$e^{i\alpha} \Delta e^{-i\alpha} = \Delta.$$

Of course our good basis elements are eigenfunctions of the Laplace operator,

$$\Delta e^{inx} = -n^2 e^{inx}.$$

This will still be true in the non-commutative setting. One essential task of harmonic analysis is then to find these eigenfunctions. Finally, we need a measure on our Lie group. This measure must respect the Lie group symmetry, and it is called the invariant measure or Haar measure. In our case, it is

$$d\mu(U(1)) = dx.$$

Invariance has the following meaning. Let $g = e^{ix}$ our group-valued (that is S^1 -valued) function, then a constant group element $e^{i\alpha}$ acts by multiplication

$$e^{ix} e^{i\alpha} = e^{i(x+\alpha)}.$$

It thus translates our variable $x \mapsto x + \alpha$. Now, consider an integral of some function $f(x)$ on S^1 , that is $f(x)$ is periodic with periodicity 2π :

$$\int_{-\pi}^{\pi} f(x + \alpha) dx = \int_{-\pi+\alpha}^{\pi+\alpha} f(x) dx = \int_{-\pi}^{\pi} f(x) dx.$$

In the last equality we used the periodicity of $f(x)$. We thus see, that the measure is translation invariant, that is it respects the symmetry of the Lie group of the circle.

The quantum mechanics interpretation of harmonic analysis on the circle is a free particle with momentum. The momentum operator is

$$p = -i\hbar \frac{d}{dx},$$

and the Hamiltonian describing the kinetic energy of the particle is

$$H = -\frac{\hbar^2}{2m} \Delta,$$

so that the particle (of mass m) given by the wave-function

$$\psi_n(x) = e^{inx}$$

has momentum $\hbar n$ and kinetic energy $\frac{\hbar^2 n^2}{2m}$.

4.5. Harmonic analysis on S^3 . we now turn to a much more complicated example of harmonic analysis, functions on the three-sphere S^3 . The three-sphere is embedded in \mathbb{R}^4 as

$$S^3 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1\}.$$

Equivalently, we can write $a = x_1 + ix_2$ and $b = x_3 + ix_4$, then we can view the three-sphere as being embedded in \mathbb{C}^2 ,

$$S^3 = \{(a, b) \in \mathbb{C}^2 \mid |a|^2 + |b|^2 = 1\}.$$

This second way gives us a good matrix representation of S^3 . Consider a 2×2 matrix

$$M = \begin{pmatrix} a & b \\ c & c \end{pmatrix}.$$

We would like this matrix to have unit determinant, that is $ad - bc = 1$ and to be unitary, that is its inverse coincides with its adjoint:

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = M^{-1} = M^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}.$$

So that $d = a^*$ and $c = -b^*$. The determinant one condition then translates to $|a|^2 + |b|^2 = 1$, so that we see that

$$S^3 = SU(2) = \left\{ M \in \text{Mat}_2(\mathbb{C}) \mid \det M = 1, M^{-1} = M^\dagger \right\}.$$

We see that the set of points on the three-sphere can be identified with unitary two by two matrices of determinant one. Given two such matrices A, B , we can take its matrix product. The inverse of the product is

$$(AB)^{-1} = B^{-1}A^{-1} = B^\dagger A^\dagger = (AB)^\dagger.$$

So that we observe that the product of two unitary matrices is still unitary. Its determinant is still one, since

$$\det(AB) = \det(A) \det(B) = 1 \cdot 1 = 1.$$

All unitary matrices are by definition invertible and hence matrix multiplication defines a group structure. This group is isomorphic to the unitary real form $SU(2)$ of the Lie group $SL(2, \mathbb{C})$. It is a real Lie group. Let us learn more about this Lie group. In order to study Lie groups, one usually first considers its infinitesimal analogue the underlying Lie algebra. Its Lie algebra is

called $su(2)$. It consists of the following matrices

$$\begin{aligned} su(2) &= \left\{ M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{Mat}_2(\mathbb{C}) \mid \text{tr}(M) = a - d = 0, M^\dagger = -M \right\} \\ &= \left\{ M = \begin{pmatrix} ia & b + ic \\ -b + ic & -ia \end{pmatrix} \in \text{Mat}_2(\mathbb{C}) \mid a, b, c \in \mathbb{R} \right\} \end{aligned} \quad (4.5)$$

We thus see, that $su(2)$ is a three-dimensional Lie algebra (over \mathbb{R}) generated by

$$\sigma_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

These are the famous Pauli matrices of physics. Their products are easily computed, they are

$$\sigma_i \sigma_j = \varepsilon_{ijk} \sigma_k - \delta_{i,j} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Here $\delta_{i,j}$ is the Kronecker delta

$$\delta_{i,j} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

and ε_{ijk} is completely anti-symmetric in all three indices

$$\varepsilon_{ijk} = -\varepsilon_{jik} = -\varepsilon_{ikj}.$$

Its normalization is $\varepsilon_{123} = 1$. It follows that the commutator of the Pauli matrices is

$$[\sigma_i, \sigma_j] = 2\varepsilon_{ijk} \sigma_k.$$

For computational purposes, the Pauli matrices are not perfect. It is convenient to pass to the complexification $sl(2; \mathbb{C}) = su(2) \otimes_{\mathbb{R}} \mathbb{C}$. Define

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

These matrices form a basis of the complexification. Their commutation relations are

$$[h, e] = 2e, \quad [h, f] = -2f, \quad [e, f] = h.$$

You might have seen this algebra in discussing the angular momentum and spin in quantum mechanics. e and f are the ladder operators, while the h -eigenvalue is the spin or angular momentum in a distinguished direction.

The Lie algebra can be obtained as the algebra of left or right-invariant vector fields of the Lie group. The invariant vector fields are infinitesimal translation operators, they are the differential operators acting on functions on the Lie group. So really, we are now computing the Lie group analogue of $\frac{d}{dx}$. For this let

$$g(\phi, \theta_1, \theta_2) = \begin{pmatrix} e^{i\theta_1} \cos \phi & e^{i\theta_2} \sin \phi \\ -e^{-i\theta_2} \sin \phi & e^{-i\theta_1} \cos \phi \end{pmatrix}.$$

g is a map from $[0, 2\pi]^3$ to $SU(2)$. The interval is chosen such that the map is injective. Surjectivity is verified by writing $a = e^{i\theta_1} r$ and $b = e^{-i\theta_2} s$ for two real numbers r and s . The unit determinant condition then implies $r^2 + s^2 = 1$, so that $r = \cos \phi$ and $s = \sin \phi$ is a good choice of parameterization.

Even though we haven't yet said yet what a Lie group is (we will do that later), let's define invariant vector fields. Just think of the Lie group as $SU(2)$.

Definition 19. Let g be a parameterization of a Lie group G , and let \mathfrak{g} be the Lie algebra of G . Then the *left-invariant vector field* for $X \in \mathfrak{g}$ is the differential operator satisfying

$$L_X g = -Xg.$$

The *right-invariant vector field* is defined analogously (up to a minus sign)

$$R_X g = gX.$$

You should think of the invariant vector fields as two commuting copies of the Lie algebra.

Theorem 23. The map

$$(X, Y) \mapsto (L_X, R_Y)$$

is a homomorphism from two commuting copies of the Lie algebra to the vector space of differential operators on the Lie group.

Proof. Let X, Y be two arbitrary elements of \mathfrak{g} . We have to show that $L_X L_Y - L_Y L_X = L_{[X, Y]}$, but

$$(L_X L_Y - L_Y L_X)g = -L_X Yg + L_Y Xg = -Y L_X g + X L_Y g = (YX - XY)g = -[X, Y]g = L_{[X, Y]}g.$$

So that the action of $L_X L_Y - L_Y L_X$ and $L_{[X, Y]}$ coincides on G and hence they are the same operators. Similarly for the right invariant vector fields

$$(R_X R_Y - R_Y R_X)g = R_X gY - R_Y gX = g(XY - YX) = g[X, Y] = R_{[X, Y]}g.$$

The commutativity of the two types of vector fields follows similarly

$$R_X L_Y g = -R_X Yg = -Y R_X g = -Y gX = L_Y gX = L_Y R_X g.$$

■

This is pretty nice, it means that functions on the Lie group carry an action of its Lie algebra. In other words analyzing the space of functions directly leads us to the representation theory of the Lie algebra. After this general interlude, let us turn to the technical question of computing the invariant vector fields for $SU(2)$. We choose to do computation in the e, f, h basis (it is simpler). The computation goes in a few steps.

(1) Let α, β, γ be functions on $[0, 2\pi)^3$, then the differential operator

$$D_{\alpha, \beta, \gamma} = \alpha \frac{d}{d\theta_1} + \beta \frac{d}{d\theta_2} + \gamma \frac{d}{d\phi}$$

acts on g as follows

$$D_{\alpha, \beta, \gamma} g = \begin{pmatrix} e^{i\theta_1}(i\alpha \cos \phi - \gamma \sin \phi) & e^{i\theta_2}(i\beta \sin \phi + \gamma \cos \phi) \\ e^{-i\theta_2}(i\beta \sin \phi - \gamma \cos \phi) & e^{-i\theta_1}(-i\alpha \cos \phi - \gamma \sin \phi) \end{pmatrix}$$

The proof is a direct computation, take the derivatives of the components of the matrix g .

(2) The elements e, f, h act as follows on g

$$\begin{aligned} -eg &= \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} g = \begin{pmatrix} e^{i\theta_2} \sin \phi & -e^{-i\theta_1} \cos \phi \\ 0 & 0 \end{pmatrix}, \\ -hg &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} g = \begin{pmatrix} -e^{i\theta_1} \cos \phi & -e^{i\theta_2} \sin \phi \\ -e^{-i\theta_2} \sin \phi & e^{-i\theta_1} \cos \phi \end{pmatrix}, \\ -fg &= \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} g = \begin{pmatrix} 0 & 0 \\ -e^{i\theta_1} \cos \phi & -e^{i\theta_2} \sin \phi \end{pmatrix}. \end{aligned}$$

These identities are verified by performing the appropriate matrix multiplications.

(3) The equation

$$-eg = D_{\alpha, \beta, \gamma} g \tag{4.6}$$

is true for

$$\alpha = -\frac{i \sin \phi}{2 \cos \phi} e^{-i(\theta_1 + \theta_2)}, \quad \beta = \frac{i \cos \phi}{2 \sin \phi} e^{-i(\theta_1 + \theta_2)}, \quad \gamma = -\frac{1}{2} e^{-i(\theta_1 + \theta_2)}.$$

You can now directly verify this identity. But the way to derive it is to consider (4.6) and to solve this matrix equation for its components. You thus get four equations for the three unknown functions α, β, γ . It turns out that this system has the unique solution given above.

(4) In analogy to step three one shows that the equation

$$-hg = D_{\alpha, \beta, \gamma} g$$

is true for

$$\alpha = \beta = i, \quad \gamma = 0.$$

(5) In analogy to step three one shows that the equation

$$-fg = D_{\alpha, \beta, \gamma} g$$

is true for

$$\alpha = -\frac{i \sin \phi}{2 \cos \phi} e^{i(\theta_1 + \theta_2)}, \quad \beta = \frac{i \cos \phi}{2 \sin \phi} e^{i(\theta_1 + \theta_2)}, \quad \gamma = \frac{1}{2} e^{i(\theta_1 + \theta_2)}.$$

We summarize

Theorem 24. The left invariant vector fields of $SU(2)$ are

$$\begin{aligned} L_e &= -\frac{i \sin \phi}{2 \cos \phi} e^{-i(\theta_1 + \theta_2)} \frac{d}{d\theta_1} + \frac{i \cos \phi}{2 \sin \phi} e^{-i(\theta_1 + \theta_2)} \frac{d}{d\theta_2} - \frac{1}{2} e^{-i(\theta_1 + \theta_2)} \frac{d}{d\phi}, \\ L_h &= i \left(\frac{d}{d\theta_1} + \frac{d}{d\theta_2} \right), \\ L_f &= -\frac{i \sin \phi}{2 \cos \phi} e^{i(\theta_1 + \theta_2)} \frac{d}{d\theta_1} + \frac{i \cos \phi}{2 \sin \phi} e^{i(\theta_1 + \theta_2)} \frac{d}{d\theta_2} + \frac{1}{2} e^{i(\theta_1 + \theta_2)} \frac{d}{d\phi}. \end{aligned} \tag{4.7}$$

It is instructive to verify that the commutation relations of these vector fields are indeed those of $su(2)$. This was the first step of harmonic analysis on the Lie group $SU(2)$, finding infinitesimal translation operators. The second step is to find a Laplace operator. For this we define

Definition 20. The *universal enveloping algebra* of $sl(2, \mathbb{C})$ is the ring of functions

$$U(sl(2, \mathbb{C})) = \mathbb{C}[e, h, f]/I.$$

Where the ideal I is generated by the polynomials

$$[x, y] - (xy - yx)$$

for all x, y in $sl(2, \mathbb{C})$. In words, the universal enveloping algebra is the polynomial ring in the generators of the Lie algebra, where the Lie bracket $[x, y]$ coincides with the commutator of elements, that is $(xy - yx)$. In the universal enveloping algebra you can compute Lie brackets as you are used to in matrix algebras.

Definition 21. A very important element of the universal enveloping algebra is the Casimir element. For $sl(2, \mathbb{C})$ it is

$$\Omega = hh - 2h + 4ef.$$

Let us compute the commutator of Ω with the generators of the Lie algebra.

$$\begin{aligned} \Omega e - e\Omega &= hhe - 2he + 4efe - ehh + 2eh - 4eef \\ &= 4eh + 4efe - 4eef \quad (\text{since } ehh = heh - 2eh = hhe - 2he - 2eh) \\ &= 4he - 4he = 0 \\ \Omega h - h\Omega &= hhh - 2hh + 4hef - hhh + 2hh + 4efh = -8ef + 8ef = 0 \\ \Omega f - f\Omega &= hhf - 2hf + 4eff - fhh + 2fh - 4fef \\ &= -4hf + 4eff - 4fef \quad (\text{since } fhh = hfh + 2fh = hhf + 2hf + 2fh) \\ &= -4hf + 4hf = 0 \end{aligned} \tag{4.8}$$

So its importance is that the Casimir commutes with the Lie algebra. For $sl(2, \mathbb{C})$ a stronger statement is even true.

Theorem 25. The center of $sl(2, \mathbb{C})$ inside its universal enveloping algebra is the polynomial ring $\mathbb{C}[\Omega]$.

The Laplace operator is then defined to be¹

$$-\Delta = L_h L_h - 2L_h + 4L_e L_f. \tag{4.9}$$

By construction it commutes with the left-invariant vector fields. The theory of compact Lie groups implies that this operator also commutes with the right-invariant vector fields. We know from linear algebra, that two operators that commute can be simultaneously diagonalized. One of our main goals of this course will be to learn about representations of Lie algebras. So how does this statement help us? In harmonic analysis we are interested in the question: What are the representations of the invariant vector fields which the functions on the Lie group carry? The Laplacian commuting with the Lie algebra of vector field means that two functions transforming in the same irreducible representation (we will learn what this word means later) of the Lie algebra have the same eigenvalue of the Laplacian. This means that finding the eigenvalue of the Laplacian is the first step in analyzing functions on a Lie group. Computing the Laplacian

¹The minus sign is just a choice of normalization, but it ensures that we get the same result as deducing the Laplace operator from the Riemannian metric on S^3 .

explicitly is now straight-forward though tedious. We compute

$$\begin{aligned} 4L_e L_f &= -\frac{\sin^2 \phi}{\cos^2 \phi} \frac{d^2}{d\theta_1^2} - \frac{\cos^2 \phi}{\sin^2 \phi} \frac{d^2}{d\theta_2^2} + 2 \frac{d}{d\theta_1} \frac{d}{d\theta_2} - \frac{d^2}{d\phi^2} + \\ &\quad + 2i \left(\frac{d}{d\theta_1} + \frac{d}{d\theta_2} \right) - \left(\frac{\cos \phi}{\sin \phi} - \frac{\sin \phi}{\cos \phi} \right) \frac{d}{d\phi} \\ L_h L_h &= -\frac{d^2}{d\theta_1^2} - \frac{d^2}{d\theta_2^2} - 2 \frac{d}{d\theta_1} \frac{d}{d\theta_2}, \end{aligned}$$

so that the explicit form of the Laplace operator is

$$\Delta = \frac{1}{\cos^2 \phi} \frac{d^2}{d\theta_1^2} + \frac{1}{\sin^2 \phi} \frac{d^2}{d\theta_2^2} + \frac{d^2}{d\phi^2} + \left(\frac{\cos \phi}{\sin \phi} - \frac{\sin \phi}{\cos \phi} \right) \frac{d}{d\phi}. \quad (4.10)$$

You might wish to rewrite this expression using the trigonometric identity

$$\left(\frac{\cos \phi}{\sin \phi} - \frac{\sin \phi}{\cos \phi} \right) = 2 \frac{\cos 2\phi}{\sin 2\phi}.$$

We have differential operators and a Laplace operator for our Lie group. Next, we need an invariant measure. Computing this is a little but subtle. I will give a method that works for any Lie group.

Definition 22. Let g be a Lie group-valued element, then its *Maurer-Cartan* one-form is

$$\omega(g) = g^{-1} dg.$$

This one-form is defined such that it is left-invariant under the translation (from the left) of a constant Lie group element h ,

$$\omega(hg) = (hg)^{-1} d(hg) = g^{-1} h^{-1} h dg = g^{-1} dg = \omega(g)$$

The Maurer Cartan form is a one-form with values in the Lie algebra of G . We can thus write

$$\omega(g) = \omega(e)'e + \omega(f)'f + \omega(h)'h.$$

The $\omega(x)'$ for x in $su(2)$ are also one-forms, and they are called the dual one-forms to the left-invariant vector fields L_x . The Haar measure is then defined to be

$$d\mu(g) = \omega(e)' \wedge \omega(f)' \wedge \omega(h)'.$$

It is by definition left-invariant, that is for any function $f(g)$ on our Lie group G , we have

$$\int_G f(hg) d\mu(g) = \int_G f(hg) d\mu(hg) = \int_G f(g) d\mu(g).$$

How do we compute this quantity in an efficient way? The Haar measure of a compact Lie group has the nice property that

$$d\mu(g) = d\mu(g^{-1}).$$

We have

$$\omega(g^{-1}) = -(dg)g^{-1} = \left(\frac{d}{d\theta_1} g \right) g^{-1} d\theta_1 + \left(\frac{d}{d\theta_2} g \right) g^{-1} d\theta_2 + \left(\frac{d}{d\phi} g \right) g^{-1} d\phi.$$

Our task is to expand this expression in the Lie algebra basis. For this, it is useful to express the derivatives in terms of left-invariant vector fields:

$$\begin{aligned}\frac{d}{d\phi}g &= \left(e^{-i(\theta_1+\theta_2)}L_f - e^{i(\theta_1+\theta_2)}L_e \right)g \\ \frac{d}{d\theta_1}g &= \left(-i\cos^2\phi L_h + i\sin\phi\cos\phi \left(e^{-i(\theta_1+\theta_2)}L_f + e^{i(\theta_1+\theta_2)}L_e \right) \right)g \\ \frac{d}{d\theta_2}g &= \left(-i\sin^2\phi L_h - i\sin\phi\cos\phi \left(e^{-i(\theta_1+\theta_2)}L_f + e^{i(\theta_1+\theta_2)}L_e \right) \right)g.\end{aligned}\tag{4.11}$$

Here we did nothing but inverted the 3×3 matrix which expresses the left-invariant vector fields in terms of the derivatives of our angles θ_1, θ_2, ϕ . Having this expression, we insert it in the definition of the Maurer Cartan form of g^{-1} , to get

$$\begin{aligned}\omega(g^{-1}) &= \omega(e)e + \omega(f) + \omega(h) \\ \omega(e) &= e^{i(\theta_1+\theta_2)}(d\phi + i\sin\phi\cos\phi(d\theta_1 - d\theta_2)) \\ \omega(f) &= e^{-i(\theta_1+\theta_2)}(-d\phi + i\sin\phi\cos\phi(d\theta_1 - d\theta_2)) \\ \omega(h) &= i\sin^2\phi d\theta_2 + i\cos^2\phi d\theta_1.\end{aligned}\tag{4.12}$$

Using the anti-symmetry of the wedge product, the Haar measure takes the explicit form

$$\begin{aligned}d\mu(g^{-1}) &= \omega(e) \wedge \omega(f) \wedge \omega(h) \\ &= 2i\sin\phi\cos\phi(d\phi \wedge (d\theta_2 - d\theta_1) \wedge (i\sin^2\phi d\theta_2 + i\cos^2\phi d\theta_1)) \\ &= -2\sin\phi\cos\phi(d\phi \wedge d\theta_1 \wedge d\theta_2)(-\sin^2\phi - \cos^2\phi) \\ &= \sin 2\phi(d\phi \wedge d\theta_1 \wedge d\theta_2).\end{aligned}\tag{4.13}$$

In the case of S^3 , there is a more intuitive way to derive this measure. Let me present this second derivation. However note, that this second derivation does not nicely generalize to any compact Lie group. Consider the threeball

$$B_r^3 = \{(a, b) \in \mathbb{C}^2 \mid |a|^2 + |b|^2 \leq r\}$$

of radius r . Then the boundary of the unit three-ball is the unit-three sphere. On \mathbb{C}^2 , we have the natural measure $da \wedge da^* \wedge db \wedge db^*$. Let us parameterize elements in B_r^3 by

$$a = r\cos\phi e^{i\theta_1}, \quad b = r\sin\phi e^{i\theta_2},$$

so that at $r = 1$ we recover our previous parameterization of S^3 . Let

$$da \wedge da^* \wedge db \wedge db^* = \mu(r)dr \wedge \mu(\phi, \theta_1, \theta_2)d\phi \wedge d\theta_1 \wedge d\theta_2$$

be the measure in the new coordinates. Let f be a function on the unit three-ball, then its integral is

$$\int_0^1 \mu(r)dr \wedge \int_{S^3} f(r, \phi, \theta_1, \theta_2)\mu(\phi, \theta_1, \theta_2)d\phi \wedge d\theta_1 \wedge d\theta_2.$$

And this integral is invariant under rotating the angles by means of an $SU(2)$ transformation, but rotations only act on the latter part, hence for every r

$$\int_{S^3} f(r, \phi, \theta_1, \theta_2)\mu(\phi, \theta_1, \theta_2)d\phi \wedge d\theta_1 \wedge d\theta_2$$

must already be $SU(2)$ invariant. So $\mu(\phi, \theta_1, \theta_2)d\phi \wedge d\theta_1 \wedge d\theta_2$ must be an invariant measure. It is computed from the Jacobian of the change of coordinates

$$\begin{aligned} \mu(r)\mu(\phi, \theta_1, \theta_2) &= \det \begin{pmatrix} \frac{d}{dr}a & \frac{d}{d\phi}a & \frac{d}{d\theta_1}a & \frac{d}{d\theta_2}a \\ \frac{d}{dr}a^* & \frac{d}{d\phi}a^* & \frac{d}{d\theta_1}a^* & \frac{d}{d\theta_2}a^* \\ \frac{d}{dr}b & \frac{d}{d\phi}b & \frac{d}{d\theta_1}b & \frac{d}{d\theta_2}b \\ \frac{d}{dr}b^* & \frac{d}{d\phi}b^* & \frac{d}{d\theta_1}b^* & \frac{d}{d\theta_2}b^* \end{pmatrix} \\ &= r^3 \sin 2\phi \end{aligned} \quad (4.14)$$

so that

$$\mu(r) = r^3 \quad \text{and} \quad \mu(\phi, \theta_1, \theta_2) = \sin 2\phi.$$

We thus obtain the exact same Haar measure as before.

Let us summarize what we have done so far. We have computed a measure, differential operators and a Laplace operator on the Lie group. So that we can finally turn to the most important question. What are functions on the Lie group? And what is a good basis of functions? There is a very powerful theorem due to Peter and Weyl.

Theorem 26. Let G be a compact Lie group and \mathcal{H} a Hilbert space, such that \mathcal{H} is a unitary representation of G , then \mathcal{H} is a direct sum of irreducible finite-dimensional representations of the underlying Lie algebra.

I still haven't told you what a Lie group is, and we will do that in the next chapter. Here, we have introduced another word, *representation* of a Lie algebra. What is that?

Definition 23. Let V be a vector space and \mathfrak{g} a Lie algebra, then V is called a *representation* ρ of \mathfrak{g} , if there is a linear map

$$\rho : \mathfrak{g} \rightarrow \text{End}(V)$$

from \mathfrak{g} to the ring of linear operators on V (this is denoted by $\text{End}(V)$), satisfying

$$\rho([x, y]) = \rho(x)\rho(y) - \rho(y)\rho(x)$$

for all x, y in \mathfrak{g} .

Such a representation is called finite-dimensional if V is finite-dimensional and it is called irreducible if there is no subrepresentation than V itself. This means, there is no subvector space W of V that is invariant under the action of \mathfrak{g} .

Let us turn to our example of $sl(2; \mathbb{C})$. \mathbb{C}^2 is a representation of it, and the action of the Lie algebra is given by the matrices we used to define the Lie algebra. This is the reason that this representation carries the special names fundamental representation, standard representation and also defining representation. Clearly a basis of this representation is given by

$$v = \begin{pmatrix} 1 & 0 \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} 0 & 1 \end{pmatrix}.$$

Using the explicit form, we can compute

$$hv = v, \quad ev = 0, \quad fv = w.$$

We see that v is an eigenvector of eigenvalue one of h . Such an eigenvalue is denoted as weight in the theory of Lie algebras. We also see that v is annihilated by e . A vector with such a property is called a highest-weight vector (of weight one). If we look at our second vector w , we find

$$hw = -w, \quad ew = v, \quad fw = 0.$$

So indeed the weight of w is lower (minus one) than the weight of v . Since w has the lowest weight in this representation, it is called a lowest-weight vector. How do we get more interesting representations of this Lie algebra? Consider the symmetric product of \mathbb{C}^2 , $\text{Sym}^2(\mathbb{C}^2)$. This is the vector space with basis $\{vv, vw, ww\}$. We can define an action of $sl(2; \mathbb{C})$ by

$$x(ab) = (xa)b + a(xb)$$

for all x in $sl(2; \mathbb{C})$ and all a, b in \mathbb{C}^2 . What is the action on our basis then?

$$\begin{aligned} hvv &= 2vv, & evv &= 0, & fvv &= 2vw, \\ hvw &= 0, & evw &= ww, & fvw &= vv, \\ hww &= -2ww, & eww &= 2vw, & fww &= 0. \end{aligned}$$

Is this representation irreducible? It is, and in order to see this, we assume that it is not irreducible. In that case, there must be a vector $z = a_1ww + a_2vw + a_3vv$ for some complex numbers a_1, a_2, a_3 such that this vector is a vector of an $sl(2; \mathbb{C})$ -invariant subspace of $\text{Sym}^2(\mathbb{C}^2)$. For this let $i \in \{1, 2, 3\}$ be such that $a_i \neq 0$ but $a_j = 0$ for $j < i$. Then

$$e^{3-i}z = a_i vv,$$

so that vv is in this sub-module, but then applying f to vv gives vw and applying it twice ww , so that the invariant subspace is already the complete vector space $\text{Sym}^2(\mathbb{C}^2)$ and we have a contradiction to our assumption.

In other words, we obtain the three-dimensional irreducible highest-weight representation of $sl(2; \mathbb{C})$, the highest-weight is 2, the highest-weight-vector is vv . In this case, ww is the lowest-weight vector of lowest-weight -2 .

Exercise 1. Show that the n -fold symmetric product of \mathbb{C}^2 carries the $n + 1$ -dimensional irreducible representation of $sl(2; \mathbb{C})$ (you have to show that the representation is irreducible). Find the highest-weight vector and the highest-weight.

The n -fold symmetric product has basis $\{v^k w^{n-k}\}$. This vector space has a unique $su(2)$ -invariant inner product. We fix normalization to be $(v^n, v^n) = 1$. The adjoint of e is f . We compute

$$f(v^k w^{n-k}) = kv^{k-1} w^{n-k+1}, \quad e(v^k w^{n-k}) = (n-k)v^{k+1} w^{n-k-1},$$

so that

$$\begin{aligned} (n-k)^2 \left((v^{k+1} w^{n-k-1}, v^{k+1} w^{n-k-1}) \right) &= \left((e(v^k w^{n-k}), e(v^k w^{n-k})) \right) \\ &= \left((fe(v^k w^{n-k}), v^k w^{n-k}) \right) \\ &= (n-k)(k+1) \left(v^k w^{n-k}, v^k w^{n-k} \right). \end{aligned}$$

Iterating this procedure, we find

$$\left(v^k w^{n-k}, v^k w^{n-k} \right) = \frac{(n-k)!k!}{n!} (v^n, v^n)$$

so that we can conclude that there is a positive inner product with orthonormal basis

$$e_{n,k} = \binom{n}{k}^{\frac{1}{2}} v^k w^{n-k}.$$

Now, we want to find highest-weight vectors and thus highest-weight representations that are functions on our Lie group. Let

$$\psi_{a,b}(\phi, \theta_1, \theta_2)$$

be a (periodic) function in our three angles ϕ, θ_1, θ_2 and let a, b be non-negative integers. We want it to be a highest-weight vector of highest-weight $n = a + b$ for the action of $sl(2; \mathbb{C})$ given by the left-invariant vector fields. In other words, we want $\psi_{a,b}$ to satisfy

$$L_h \psi_{a,b}(\phi, \theta_1, \theta_2) = n \psi_{a,b}(\phi, \theta_1, \theta_2)$$

and

$$L_e \psi_{a,b}(\phi, \theta_1, \theta_2) = 0.$$

The first condition can be satisfied if we write

$$\psi_{a,b}(\phi, \theta_1, \theta_2) = f_{a,b}(\phi) e^{-ia\theta_1 - ib\theta_2}.$$

It thus remains to find $f_{a,b}(\phi)$. The condition $L_e \psi_{a,b} = 0$ translates into

$$-a \frac{\sin \phi}{\cos \phi} + b \frac{\cos \phi}{\sin \phi} - \frac{1}{f_{a,b}(\phi)} \frac{d}{d\phi} f_{a,b}(\phi) = 0$$

This equation can be rewritten using

$$\frac{d}{d\phi} \ln \cos \phi = -\frac{\sin \phi}{\cos \phi}, \quad \frac{d}{d\phi} \ln \sin \phi = \frac{\cos \phi}{\sin \phi}, \quad \frac{1}{f_{a,b}(\phi)} \frac{d}{d\phi} f_{a,b}(\phi) = \frac{d}{d\phi} \ln f_{a,b}(\phi).$$

Namely we get

$$\frac{d}{d\phi} \left(\ln \left(f_{a,b}(\phi) (\cos(\phi))^{-a} (\sin(\phi))^{-b} \right) \right) = 0$$

and hence up to a normalization

$$f_{a,b}(\phi) = (\cos(\phi))^a (\sin(\phi))^b$$

and the highest-weight vector becomes

$$\psi_{a,b}(\phi, \theta_1, \theta_2) = (\cos(\phi))^a (\sin(\phi))^b e^{-ia\theta_1 - ib\theta_2}.$$

The theorem of Peter and Weyl tells us, that this function must transform in an irreducible-highest weight representation of weight $n = a + b$. But such a representation is the unique $n + 1$ -dimensional irreducible representation of highest-weight n , call it ρ_n . It is spanned by

$$\psi_{a,b}, L_f \psi_{a,b}, L_f^2 \psi_{a,b}, \dots, L_f^n \psi_{a,b}.$$

Note, that a non-trivial implication of the Peter Weyl theorem is that

$$L_f^{n+1} \psi_{a,b} = 0.$$

Computing the basis vectors is both straight-forward and tedious. For example one gets

$$\begin{aligned} L_f \psi_{a,b} &= e^{-i(a-1)\theta_1 - i(b-1)\theta_2} \left(-a(\cos \phi)^{a-1} (\sin \phi)^{b+1} + b(\cos \phi)^{a+1} (\sin \phi)^{b-1} \right) \\ L_f^2 \psi_{a,b} &= e^{-i(a-2)\theta_1 - i(b-2)\theta_2} \left(a(a-1)(\cos \phi)^{a-2} (\sin \phi)^{b+2} + 2ab(\cos \phi)^a (\sin \phi)^b + \right. \\ &\quad \left. + b(b-1)(\cos \phi)^{a+2} (\sin \phi)^{b-2} \right). \end{aligned}$$

The lowest-weight vector can be computed exactly in the same way as in the highest-weight case.

Exercise 2. Let a, b be non-negative integers with $a + b = n$. Make the Ansatz

$$\phi_{a,b}(\theta_1, \theta_2, \phi) = g_{a,b}(\phi) e^{ia\theta_1 + ib\theta_2}.$$

Show that $\phi_{a,b}$ has weight $-n$, i.e.

$$L_h \phi_{a,b} = -n \phi_{a,b}.$$

We require that $\phi_{a,b}$ is a lowest-weight vector, that is

$$L_f \phi_{a,b} = 0.$$

Show that this implies that $g_{a,b}$ satisfies exactly the same differential equation as $f_{a,b}$ before, that is

$$\frac{d}{d\phi} \left(\ln \left(g_{a,b}(\phi) (\cos(\phi))^{-a} (\sin(\phi))^{-b} \right) \right) = 0.$$

Conclude that the lowest-weight vector is

$$\phi_{a,b}(\phi, \theta_1, \theta_2) = (\cos(\phi))^a (\sin(\phi))^b e^{ia\theta_1 + ib\theta_2}.$$

In other words, the lowest-weight vector is the complex conjugate of the highest-weight vector.

We thus get for the norm of the highest-weight vector

$$\begin{aligned} (\Psi_{a,b}, \Psi_{a,b}) &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} (\cos(\phi))^{2a} (\sin(\phi))^{2b} \sin 2\phi d\theta_1 d\theta_2 d\phi \\ &= 8\pi^2 \int_{-\pi}^{\pi} (\cos(\phi))^{2a+1} (\sin(\phi))^{2b+1} d\phi \quad (2 \cos \phi \sin \phi = \sin 2\phi) \\ &= 8\pi^2 \int_{-1}^1 (1-x^2)^a x^{2b+1} dx \quad (x = \sin \phi, dx = \cos \phi d\phi, (\cos \phi)^2 = 1-x^2) \\ &= 8\pi^2 \frac{a}{b+1} \int_{-1}^1 (1-x^2)^{a-1} x^{2(b+1)+1} dx \tag{4.15} \\ &\vdots \\ &= 8\pi^2 \frac{a!b!}{(b+a)!} \int_{-1}^1 x^{2(b+a)+1} dx \\ &= 8\pi^2 \frac{a!b!}{(b+a+1)!}. \end{aligned}$$

We have learnt before, that there is up to normalization of the norm of the highest-weight vector a unique invariant inner product on a representation. So that we have determined all norms of all vectors in the highest-weight representation.

There is one more nice observation we can make. The Casimir (the Laplacian in terms of the vector fields) commutes with all elements of the Lie algebra. Hence all elements of ρ_n have the same eigenvalue of the Laplacian. It is most easy to compute this number acting on the lowest-weight state. However, since we already have an explicit expression for the highest-weight state, let us rewrite the Laplacian as

$$\Delta = -L_h L_h + 2L_h - 4L_e L_f = -L_h L_h - 2L_h - 4L_f L_e.$$

Here we used the relation $L_f L_e - L_e L_f = -h$. Now, we can compute

$$\Delta \Psi_{a,b} = (-L_h L_h - 2L_h - 4L_f L_e) \Psi_{a,b} = (-L_h L_h - 2L_h) \Psi_{a,b} = (-n^2 - 2n) \Psi_{a,b}.$$

In other words, every element in the representation ρ_n has Casimir eigenvalue $-n^2 - 2n$. We summarize our findings

Theorem 27. The Hilbert space of square integrable functions on the Lie group $SU(2)$ has basis

$$\left\{ L_f^m \psi_{a,b}(\phi, \theta_1, \theta_2) \mid a, b \in \mathbb{Z}_{\geq 0}, 0 \leq m \leq a + b; \psi_{a,b}(\phi, \theta_1, \theta_2) = (\cos(\phi))^a (\sin(\phi))^b e^{-ia\theta_1 - ib\theta_2} \right\}.$$

The highest-weight vectors of highest weight n are the $\psi_{a,b}(\phi, \theta_1, \theta_2)$ with $a + b = n$ and the functions

$$\psi_{a,b}, L_f \psi_{a,b}, L_f^2 \psi_{a,b}, \dots, L_f^n \psi_{a,b}$$

span the irreducible $n + 1$ dimensional highest-weight representation ρ_n of $sl(2; \mathbb{C})$ of highest-weight n . The Laplacian has eigenvalue $-n^2 - 2n$ on each function in this representation. As a representation of the left-invariant vector fields, the space of square-integrable functions on $SU(2)$ decomposes as

$$\mathcal{L}_\mu^2(SU(2)) = \bigoplus_{n \in \mathbb{Z}_{\geq 0}} \dim(\rho_n) \rho_n, \quad (4.16)$$

since there are $\dim(\rho_n) = n + 1$ distinct highest-weight vectors of highest-weight n .

If one also considers the right-invariant vector fields, one can do better. Recall that right-invariant vector fields commute with the left-invariant ones. A representation of two commuting copies of a Lie algebra \mathfrak{g} decomposes as

$$\rho_1 \otimes \rho_2,$$

where ρ_1 is a representation of the first copy of the algebra and ρ_2 of the second one. The action of $\mathfrak{g} \oplus \mathfrak{g}$ on a vector $v \otimes w$ is then defined as

$$(x \oplus y)(v \otimes w) = (xv \otimes w) \oplus (v \otimes yw).$$

We can rewrite our $SU(2)$ decomposition (4.16) as

$$\mathcal{L}_\mu^2(SU(2)) = \bigoplus_{n \in \mathbb{Z}_{\geq 0}} \rho_n^L \otimes \mathbb{C}^{n+1}.$$

Here we put an upper index L on the representation to indicate that it is a representation of the left-invariant vector fields. It is now suggestive that the right-invariant ones act on the multiplicity vector spaces \mathbb{C}^{n+1} , and indeed the following theorem is true.

Theorem 28. Under the action of both left- and right-invariant vector fields, the space of square integrable functions on $SU(2)$ decomposes as

$$\mathcal{L}_\mu^2(SU(2)) = \bigoplus_{n \in \mathbb{Z}_{\geq 0}} \rho_n^L \otimes \rho_n^R. \quad (4.17)$$

In order to prove this statement, you have to show that the $n + 1$ functions $\psi_{a,b}$ (with $a + b = n$) carry the $n + 1$ -dimensional representation of the $sl(2; \mathbb{C})$ -action given by the right-invariant vector fields. By the theorem of Peter and Weyl, this amounts to finding a highest-weight vector of highest-weight n .

Question: Can you formulate an analogous theorem for the circle $S^1 = U(1)$?

4.6. Summary. This section has been quite some work, so let us summarize what we have done. We have started by learning that the three-sphere carries a group structure, namely of the Lie group $SU(2)$. We then have decided that harmonic analysis on the three-sphere is harmonic analysis on the group. For the latter, we had to compute differential operators. These operators

are called invariant vector fields. There are left-invariant ones and right-invariant ones. They make functions on the Lie group into a representation of two commuting copies of the underlying Lie algebra, $(\mathfrak{su}(2))$. We then decided that it is easier to work with the complexification of this real Lie algebra. The complexification is called $\mathfrak{sl}(2; \mathbb{C})$. Once having obtained these differential operators, we immediately obtained the Laplacian from the Casimir element of the Lie algebra. The next step was then to find an invariant measure, which has been computed from a left-invariant one-form, the Maurer-Cartan form. After all this preparation, we have learnt that $\mathfrak{sl}(2; \mathbb{C})$ has special representations, called irreducible finite-dimensional highest-weight representations. There is a very important theorem due to Peter and Weyl, that tells us that exactly these representations appear in the harmonic analysis of $SU(2)$. Indeed, we were then able to explicitly compute highest- and lowest-weight vectors as well as their norm. The norm of all other elements in the representation is then determined by the one of the highest-weight vector. The final theorem is then your homework problem, to find the decomposition of square integrable functions on $SU(2)$ under both left- and right-action of the underlying Lie algebra of vector fields.

4.7. Exercises.

- (1) Proof Theorem 28.

In other words, compute the right-invariant vector fields in a similar manner as we have done it for the left-invariant ones. The first step for that is the same as for the left-invariant ones. The second step is also very analogous to the case of the left-invariant vector fields, namely it is to compute the right-action of the matrices corresponding to the Lie algebra on the Lie group valued function g . Then you have to solve equations of type (4.6), that is for example $ge = D_{\alpha, \beta, \gamma} g$ gives you R_e .

Secondly, you need to use these vector fields to find highest-weight vectors, that is functions ψ that satisfy the differential equations

$$R_e \psi = 0, \quad R_h \psi = n \psi.$$

4.8. Solutions.

- (1) The first task is to compute the right-invariant vector fields. We use the same parameterization of a Lie group valued matrix g as in the lecture. We then compute

$$\begin{aligned} ge &= \begin{pmatrix} 0 & e^{i\theta_1} \cos \phi \\ 0 & -e^{-i\theta_2} \sin \phi \end{pmatrix} \\ gh &= \begin{pmatrix} e^{i\theta_1} \cos \phi & -e^{i\theta_2} \sin \phi \\ -e^{-i\theta_2} \sin \phi & -e^{-i\theta_1} \cos \phi \end{pmatrix} \\ gf &= \begin{pmatrix} e^{i\theta_2} \sin \phi & 0 \\ e^{-i\theta_1} \cos \phi & 0 \end{pmatrix} \end{aligned}$$

The next step is to solve the differential matrix equations

$$D_{\alpha, \beta, \gamma} g = gX$$

for $X = e, h, f$. This gives four equations for the three functions α, β, γ . There must be a unique solution in each case and indeed the answer to these equations turns out to be

$$R_e = -\frac{i \sin \phi}{2 \cos \phi} e^{i(\theta_1 - \theta_2)} \frac{d}{d\theta_1} - \frac{i \cos \phi}{2 \sin \phi} e^{i(\theta_1 - \theta_2)} \frac{d}{d\theta_2} + \frac{1}{2} e^{i(\theta_1 - \theta_2)} \frac{d}{d\phi},$$

$$R_h = -i \left(\frac{d}{d\theta_1} - \frac{d}{d\theta_2} \right),$$

$$R_f = -\frac{i \sin \phi}{2 \cos \phi} e^{-i(\theta_1 - \theta_2)} \frac{d}{d\theta_1} - \frac{i \cos \phi}{2 \sin \phi} e^{-i(\theta_1 - \theta_2)} \frac{d}{d\theta_2} - \frac{1}{2} e^{-i(\theta_1 - \theta_2)} \frac{d}{d\phi}.$$

we are now looking for functions $\phi_{a,b}(\theta_1, \theta_2, \phi)$ with $a + b = n$ satisfying the system of differential equations

$$R_h \phi_{a,b} = n \phi_{a,b}, \quad R_e \phi_{a,b} = 0.$$

For this we make the separation of variables Ansatz

$$\phi_{a,b}(\theta_1, \theta_2, \phi) = g_{a,b}(\phi) e^{i(a\theta_1 - b\theta_2)}$$

so that the equation $R_h \phi_{a,b} = n \phi_{a,b}$ is satisfied. Plugging the Ansatz in the second differential equation, we can solve it exactly with the same technique as in the lecture to get that

$$g_{a,b}(\phi) = (\cos \phi)^a (\sin \phi)^b.$$

So that we get $n + 1$ solutions for our differential equation. These are all highest-weight vectors for the $n + 1$ dimensional irreducible representation of $sl(2; \mathbb{R})$. But especially, we see that $\phi_{0,n} = \psi_{0,n}$ is both a highest-weight vector for the left and for the right regular action. Since these two actions commute this is a highest-weight vector for the tensor product and the theorem follows.

5. LIE GROUPS

In this section, we will repeat part of the analysis of last section in a much more general setting. We will learn about important concepts of Lie groups. I use the book by Daniel Bump [Bu] as reference.

Definition 24. Let G be a n -dimensional real manifold, and let $\{U_\alpha\}$ be a set of open subsets of G that cover G , such that each open subset looks like \mathbb{R}^n in the sense that there is a bijective map

$$\phi_\alpha : U_\alpha \rightarrow \mathbb{R}^n$$

that is bi-continuous. In other words both ϕ_α and its inverse are continuous maps. A *transition map* is then a map

$$\phi_{\alpha\beta} = \phi_\beta \circ \phi_\alpha^{-1} \Big|_{\phi_\alpha(U_\alpha \cap U_\beta)}$$

from $\phi_\alpha(U_\alpha \cap U_\beta)$ to $\phi_\beta(U_\alpha \cap U_\beta)$. A manifold is called *smooth* if all transition maps are smooth, that is infinitely differentiable. A Lie group G (over \mathbb{R}) is a smooth manifold and also a group, such that group multiplication and inversion are smooth maps.

Example 7. A list of examples are

- (1) \mathbb{R}^n with group operation addition of vectors.
- (2) $\mathbb{R} \setminus \{0\}$ with group operation multiplication.
- (3) positive real numbers with group operation multiplication

- (4) the unit circle S^1 . We have called this the Lie group $U(1)$.
- (5) real invertible $n \times n$ matrices. This is called the Lie group $GL(n, \mathbb{R})$. Most Lie groups of interests will turn out to be subgroups of the complex Lie group $GL(n, \mathbb{C})$.
- (6) unitary $n \times n$ matrices of determinant one, the special unitary Lie group $SU(n)$.

As a physicist, we are interested in problems, where the Lie group is the continuous group of symmetries of the problem. A physical observable is then an objective that carries an action of the symmetry, that is an action of the Lie group. There are two important notions of action: representation and action.

Definition 25. An *action* of a Lie group G on a manifold M is an assignment ρ , that assigns to each g in G a diffeomorphism $\rho(g)$ on M , with the properties that this assignment is compatible with the group structure, that is

$$\rho(1) = Id, \quad \rho(gh) = \rho(g)\rho(h)$$

for all h, g in the Lie group. In addition one wants $\rho(g)$ to be a smooth map on M .

A *representation* of a Lie group G is a vector space V together with a group morphism $\rho : G \rightarrow \text{End}(V)$. We will be interested in finite-dimensional representations, that is $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$. In that case $\text{End}(V)$ is the Lie group of invertible (real or complex) $n \times n$ matrices. The map ρ is then a morphism of Lie groups, which means it is a smooth map that respects the Lie group structure in the sense, that $\rho(gh) = \rho(g)\rho(h)$ and the image of the identity is the identity matrix.

Example 8. A very important example of a group action is the group action on a coset space. Let G be a Lie group and H be a sub Lie group of G , that is a subset of G that itself is a Lie group, then the space of cosets

$$G/H := \{gh | g \in G, h \in H, gh = g'h' \text{ if } g^{-1}g' \in H\}$$

is called the coset G/H of G . It carries an action of G by multiplication from the left. Similarly one can also define the coset

$$H \backslash G := \{hg | g \in G, h \in H, hg = h'g' \text{ if } g'g^{-1} \in H\}$$

it then carries an action of the Lie group by multiplication from the right. Important interesting manifolds like spheres can be constructed as such coset manifolds.

A very important example of a representation of a Lie group is the Hilbert space of square integrable function on the Lie group, denoted by

$$\mathcal{L}^2_{\mu}(G),$$

where $\mu = \mu(G)$ refers to the Haar measure, that is a measure that respects the group action. This measure is the weight function of the Hilbert space, and part of our coming analysis will also be concerned with this measure. Also the Hilbert space of square integrable function on a coset space is a representation of a Lie group. The analysis of this Hilbert space uses the resulting analysis of the parent Lie group Hilbert space. We will however not look at such examples, as it is getting more and more complicated.

We will now describe harmonic analysis on a Lie group in more detail. Harmonic analysis is nothing but quantum mechanics on the group manifold. One might also like to view it as a somehow semi-classical limit of quantum field theory on the Lie group, so in any case it is very important for a theoretical quantum physicist. We will see, that we essentially need to

understand a nice class of representations of Lie algebras. The most important theorem is the famous theorem of Peter and Weyl that we have already mentioned in the previous chapter.

Definition 26. Let \mathcal{H} be a Hilbert space, a representation $\rho : G \rightarrow \text{End}(\mathcal{H})$ is called *unitary* if the inner product respects the Lie group in the sense that

$$(\rho(g)v, \rho(g)w) = (v, w)$$

for all g in G and all v, w in \mathcal{H} .

If \mathcal{H} is the Hilbert space of square integrable functions on G , then every function f carries two-commuting actions of G , the left-regular action defined by

$$L_h : \mathcal{H} \rightarrow \mathcal{H}, \quad f(g) \mapsto f(gh).$$

and the right-regular one

$$R_h : \mathcal{H} \rightarrow \mathcal{H}, \quad f(g) \mapsto f(h^{-1}g)$$

The Peter-Weyl theorem then has two versions. Both of them you have already seen in the example of $SU(2)$ in the last chapter.

Theorem 29. (Peter and Weyl)

- (1) Let \mathcal{H} be a Hilbert space and $\rho : G \rightarrow \text{End}(\mathcal{H})$ a unitary representation of a compact group G . Then \mathcal{H} is a direct sum of finite-dimensional irreducible representations.
- (2) Let $\mathcal{H} = \mathcal{L}^2_\mu(G)$ be the Hilbert space of square integrable functions of a compact Lie group G . Then \mathcal{H} is a unitary representation of G . Moreover, let \mathcal{R} be the set of all finite-dimensional unitary representations of G , then under the left-right action of G , the Hilbert space decomposes as

$$\mathcal{L}^2_\mu(G) = \bigoplus_{\rho \in \mathcal{R}} \rho^L \otimes \bar{\rho}^R.$$

Here, by $\bar{\rho}$ we denote the representation conjugate to ρ , meaning that the one-dimensional representation is contained in the tensor product of the two representations and hence there exists a linear map on this product.

Compare this result with the last theorem of last chapter. We will now turn on the relation between Lie group and Lie algebra, and then rephrase this theorem in terms of representations of the Lie algebra of invariant vector fields. We start with the invariant measure.

5.1. The Haar measure. We require that G is a locally compact Lie group. This means that every point x in G has a compact neighbourhood. This is not a severe restriction, as every compact Lie group is especially locally compact, but also \mathbb{R}^n is locally compact and more generally every Lie group that looks locally like Euclidean space is locally compact. Looking like Euclidean space means that every point has an open neighbourhood that is homeomorphic (there is a bi-continuous map) to Euclidean space.

Definition 27. A measure μ on a locally compact group G is called a *left Haar* measure if it is regular, that is

$$\mu(X) = \inf\{\mu(U) | U \supset X, U \text{ open}\} = \sup\{\mu(K) | K \subset X, K \text{ compact}\}$$

and if it is left-invariant, that is $\mu(X) = \mu(gX)$

Such a measure has the property, that any compact set has finite measure and any nonempty open set has measure > 0 .

Theorem 30. If G is a locally compact group, then there is a unique left Haar measure.

Left-invariance of the measure amounts to left-invariance of the integral of an integrable function f

$$\int_G f(\gamma g) d\mu(g) = \int_G f(g) d\mu(g).$$

There is also a unique right Haar measure. Left and right measure do not necessarily coincide. For example, let

$$G = \left\{ \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \mid x, y \in \mathbb{R}, y > 0 \right\}$$

then one can show that the left measure is

$$d\mu_L = y^{-2} dy dx,$$

but the right measure is

$$d\mu_R = y^{-1} dy dx.$$

Definition 28. A Lie group G is called *unimodular* if left and right Haar measure coincide.

So unimodular Lie groups are in a sense very symmetric. These are the type of Lie groups appearing frequently in physical problems. Let g in G , then g acts on G via conjugation, that is $h \mapsto ghg^{-1}$. Every conjugation is an automorphism of G , since

$$g1g^{-1} = 1, \quad ghg^{-1}gh'g^{-1} = gh'h'g^{-1}.$$

Every automorphism takes a measure to another measure, so conjugation maps the left Haar measure to another left Haar measure. Uniqueness implies that this must be a constant multiple of the original measure. Thus for every g in G , there is $\delta(g) > 0$, such that

$$\int_G f(g^{-1}hg) d\mu_L(h) = \int_G f(h) d\mu_L(ghg^{-1}) = \delta(g) \int_G f(h) d\mu_L(h).$$

Definition 29. A *quasicharacter* is a continuous homomorphism

$$\chi : G \rightarrow \mathbb{C} \setminus \{0\}.$$

If $|\chi(g)| = 1$ for all g in G , then χ is called a unitary quasicharacter.

Proposition 31. The function

$$\delta : G \rightarrow \mathbb{R}_{>0}$$

is a quasicharacter and the measure $\delta(h)\mu_L(h)$ is right invariant.

Proof. Conjugation by first g_1 and then by g_2 is the same as conjugating by g_1g_2 . Thus $\delta(g_1)\delta(g_2) = \delta(g_1g_2)$ and hence the map is a homomorphism, it is also continuous and hence a quasicharacter. Using the left invariance of the Haar measure, we get

$$\delta(g) \int_G f(h) d\mu_L(h) = \int_G f(g^{-1}hg) d\mu_L(h) = \int_G f(gg^{-1}hg) d\mu_L(h) = \int_G f(hg) d\mu_L(h).$$

Replace f by δf , so that

$$\delta(g) \int_G f(h) \delta(h) d\mu_L(h) = \int_G f(hg) \delta(hg) d\mu_L(h).$$

Now using that δ is a homomorphism and dividing by $\delta(g)$ gives the result. ■

Proposition 32. If G is compact, then G is unimodular and μ is finite.

Proof. The map δ is a homomorphism, so that its image must be a subgroup of $\mathbb{R}_{>0}$. Since G is compact and δ is continuous its image must also be compact. The only compact subgroup of the positive real numbers is $\{1\}$. Thus $\delta(g) = 1$ for all g in G . Hence the left Haar measure coincides with the right Haar measure. The volume of a compact subset of a locally compact group is finite, so the volume of G must already be finite. ■

Finally, another useful result is

Proposition 33. If G is unimodular, then $g \mapsto g^{-1}$ is an isometry.

Proof. The map $g \mapsto g^{-1}$ turns a left-invariant measure into a right-invariant measure. If both measures agree, then the map $g \mapsto g^{-1}$ multiplies the left Haar measure by a positive constant. Since the map is of order two, this constant must be one. ■

We now turn to the essential example of Lie groups, these are subgroups of $GL(n, \mathbb{C})$.

5.2. Lie subgroups of $GL(n, \mathbb{C})$. Every classical Lie group is such a subgroup, so actually we are studying a quite generic situation in this section. To get an idea we list the standard examples

Example 9. Interesting subgroups of $GL(n, \mathbb{C})$ are

- (1) The orthogonal (compact) group $O(n)$,

$$O(n) = \{g \in GL(n, \mathbb{R}) \mid gg^t = I\}$$

- (2) The unitary (compact) group $U(n)$,

$$U(n) = \{g \in GL(n, \mathbb{C}) \mid g(g^*)^t = I\}$$

- (3) The special (compact) unitary group $SU(n)$,

$$SU(n) = \{g \in U(n) \mid \det g = 1\}$$

- (4) The symplectic group (compact) $Sp(2n)$

$$Sp(2n) = \{g \in GL(2n, \mathbb{R}) \mid gJg^t = J\}$$

for

$$J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$$

Exercise 3. Show, that the sets of the example are groups.

We want to find the Lie algebra of these groups. So let us recall the definition.

Definition 30. A Lie algebra \mathfrak{g} is a vector space \mathfrak{g} together with a bilinear operation

$$[\ , \]: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$$

satisfying

$$[U, T] = -[T, U] \quad (\text{antisymmetry})$$

and

$$[U, [S, T]] + [T, [U, S]] + [S, [T, U]] = 0 \quad \text{Jacobi identity}$$

for all U, S, T in \mathfrak{g} .

Let $\text{Mat}_n(\mathbb{C})$ be the vector space of $n \times n$ matrices over the complex numbers. The Lie group $GL(n, \mathbb{C})$ is the group of all invertible such matrices. The important map is the exponential map:

$$\exp: \text{Mat}_n(\mathbb{C}) \rightarrow GL(n, \mathbb{C}), \quad \exp(X) = I + X + \frac{1}{2}X^2 + \frac{1}{6}X^3 + \dots \quad (5.1)$$

This series converges for every X . You prove this by observing this statement for matrices in Jordan normal form and then use that every matrix can be brought into such a form. Let G be a Lie sub group of $GL(n, \mathbb{C})$, we want to somehow think of the Lie algebra of G , called $Lie(G)$ as all matrices X , such that $\exp(X)$ is in G . We will now explain this intuitive picture.

Proposition 34. Let U be an open subset of \mathbb{R}^n , and let x in U . Then there is a smooth function f with compact support contained in U that does not vanish at x .

Proof. Assume that $x = (x_1, \dots, x_n)$ is the origin (otherwise translate the point x into the origin). Define

$$f(x_1, \dots, x_n) = \begin{cases} \exp\left(-\left(1 - \frac{|x|^2}{r^2}\right)^{-1}\right) & \text{if } |x| \leq r \\ 0 & \text{otherwise} \end{cases}.$$

The function f is smooth, and has support inside the ball $|x| \leq r$. For sufficiently small r it vanishes outside of U . ■

Definition 31. Let G be a Lie group, and let g be a point in G . Let $\gamma_i : (-1, 1) \rightarrow G$ be a curves with $\gamma_i(0) = g$ for $i = 1, 2$. Then choose an open set $U \subset G$ containing g and a chart $\phi : U \rightarrow \mathbb{R}^n$. γ_1 and γ_2 are equivalent if $\phi \circ \gamma_1 = \phi \circ \gamma_2$. The set of such equivalence classes is called the *tangent space* of G at the point g .

Proposition 35. Let G be a subgroup of $GL(n, \mathbb{C})$ and let X be an $n \times n$ matrix over \mathbb{C} . Then the path $t \rightarrow \exp(tX)$ is tangent to G at the identity if and only if it is contained in G for all t .

Proof. If $\exp(tX)$ is contained in G , then it is tangent to G at the identity. Suppose that there is $t_0 > 0$ with $\exp(t_0X)$ not an element of G , but the path is still tangent to G . We will derive a contradiction to this assumption. With the previous proposition, we know that there is a smooth compactly supported function ϕ_0 on $GL(n, \mathbb{C})$ with $\phi_0(g) = 0$ for all g in G and $\phi_0(\exp(t_0X)) \neq 0$. Let

$$f(t) = \phi(\exp(tX)), \quad \phi(h) = \int_G \phi_0(hg) d\mu$$

with the left Haar measure $d\mu$ on G . So that ϕ is constant on the left cosets hG of G and especially vanishes on G , but is non-zero at $\exp(t_0X)$. For any t , we can write the derivative as

$$f'(t) = \frac{d}{du} \phi(\exp(tX)\exp(uX)) \Big|_{u=0} = 0$$

since the path $\exp(tX)\exp(uX)$ is tangent to the coset $\exp(tX)G$ and ϕ is constant on such cosets. Moreover $f(0) = 0$ and hence $f(t) = 0$ for all t , but this is a contradiction to $f(t_0) \neq 0$. ■

Corollary 36. Let G be a subgroup of $GL(n, \mathbb{C})$. The set $Lie(G)$ of all $X \in Mat_n(\mathbb{C})$ such that $\exp(tX) \in G$ is a vector space whose dimension is equal to the dimension of G as a manifold.

Proposition 37. Let G be a subgroup of $GL(n, \mathbb{C})$. The map

$$X \rightarrow \exp(X)$$

gives a diffeomorphism of a neighborhood of the identity in $Lie(G)$ onto a neighborhood of the identity in G .

Proof. Recall the expansion of $\exp(X) = I + X + X^2/2 + \dots$ so that the Jacobian of \exp at the identity is one. Hence \exp induces a diffeomorphism of an open neighborhood of the identity in $Mat_n(\mathbb{C})$ onto a neighborhood of the identity in $GL(n, \mathbb{C})$. Since $Lie(G)$ is a vector space of

the same dimension as G the Inverse Function Theorem (you might have had that in analysis) implies that $\text{Lie}(G) \cap U$ must be mapped onto an open neighborhood of the identity in G . ■

Proposition 38. Let G be a subgroup of $GL(n, \mathbb{C})$, then for X, Y in $\text{Lie}(G)$, also $[X, Y]$ is in $\text{Lie}(G)$.

Proof. Using the expansion of \exp , we see that

$$\exp(e^{tX}Ye^{-tX}) = e^{tX}\exp(Y)e^{-tX},$$

so that with Y also $e^{tX}Ye^{-tX}$ is an element of $\text{Lie}(G)$. Thus $\text{Lie}(G)$ contains

$$\frac{1}{t}(e^{tX}Ye^{-tX} - Y) = XY - YX = \frac{t}{2}(X^2Y - 2XYX + YX^2) + \dots$$

this is true for all t . Taking the limit $t \rightarrow 0$ shows that $[X, Y] = XY - YX$ is also in $\text{Lie}(G)$. ■

In the algebra chapter, we have seen, that the commutator gives an associative algebra the structure of a Lie algebra.

Theorem 39. $\text{Lie}(G)$ is a Lie subalgebra of $gl(n, \mathbb{C})$ of same dimension as G .

Example 10. The important examples are

- (1) The Lie algebra of $O(n, \mathbb{R})$ is $o(n, \mathbb{R})$. It consists of all matrices X satisfying $X + X^t = 0$.
- (2) The unitary Lie algebra is denoted by $u(n, \mathbb{R})$ and consists of all complex matrices X satisfying $X^* + X^t = 0$.
- (3) The special unitary Lie algebra is denoted by $su(n, \mathbb{R})$ and consists of all traceless unitary matrices.
- (4) The symplectic Lie algebra is denoted by $sp(2n, \mathbb{R})$ and consists of all matrices X satisfying $XJ + JX^t = 0$ with $J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$.

In order to see that these examples are indeed the Lie algebras of the Lie groups presented at the beginning of this section, one has to exponentiate the relations. For example, let X in $o(n, \mathbb{R})$, then $tX^t = -tX$ for all t and hence

$$\exp(tX)^{-1} = \exp(tX)^t$$

so that $\exp(tX)$ is in $O(n, \mathbb{R})$ for all t . Thus $o(n, \mathbb{R})$ is a subgroup of $\text{Lie}(O(n, \mathbb{R}))$. For the converse direction, suppose that X in $\text{Lie}(O(n, \mathbb{R}))$, then for all t

$$\begin{aligned} I &= \exp(tX)\exp(tX)^t \\ &= \left(I + tX + \frac{1}{2}t^2X^2 + \dots \right) \left(I + tX^t + \frac{1}{2}t^2(X^t)^2 + \dots \right) \\ &= I + t(X + X^t) + \frac{1}{2}t^2(X^2 + 2XX^t + (X^t)^2) + \dots \end{aligned}$$

This is true for all t , hence $X + X^t = 0$ and hence $\text{Lie}(O(n, \mathbb{R}))$ is $o(n, \mathbb{R})$.

5.3. Left-Invariant vector fields. Let G be a Lie group. G acts on itself by multiplication from the left. Let us call this action L_g ,

$$L_g : G \rightarrow G, \quad h \mapsto gh.$$

Consider the Tangent space at a point h , that is $T_h(G)$. Recall that this space was the space of equivalence classes of curves $\gamma : (-1, 1) \rightarrow G$ with $\gamma(0) = h$. If we compose such a curve with

L_g , we get a tangent vector at the point gh . We thus have defined a map

$$L_{g,*} : T_h \rightarrow T_{gh}, \quad \gamma \mapsto L_g \circ \gamma.$$

Definition 32. A vector field X on a Lie group G is a collection of assignments, that assigns to each $g \in G$ an element X_g in $T_g(G)$. A vector field X on G is called *left-invariant* if

$$L_{g,*}(X_h) = X_{gh}.$$

Proposition 40. The vector space of left-invariant vector fields is closed under the commutator $[\ , \]$ and is a Lie algebra of dimension $\dim(G)$. If e is the identity of G and if X_e in $T_e(G)$, then there is a unique left-invariant vector field X on G with the prescribed tangent vector at the identity.

Proof. Given a tangent vector at the identity X_e , a left-invariant vector field is defined by $X_g = L_{g,*}(X_e)$. Conversely every left-invariant vector field must satisfy this identity. Hence the space of left-invariant vector fields is isomorphic to the tangent space of G at the identity. Therefore its vector space dimension equals the dimension of G . The Lie algebra structure is given by the tangent vectors at the identity. ■

The Lie algebra of a Lie group, is the Lie algebra of the left-invariant vector fields. However, we have in the previous section already defined the Lie algebra of $GL(n, \mathbb{C})$ and its subgroups in terms of $n \times n$ matrices. We thus need to see that the two definitions define isomorphic Lie algebras. let $G = GL(n, \mathbb{C})$. Then both the tangent space and the Lie algebra of $n \times n$ matrices are of dimension n^2 . Now, let X be an $n \times n$ matrix. The tangent space at the identity may be identified with $n \times n$ matrices, as it is a vector space of that dimension. This means to each X , we can define a differential operator L_X via

$$L_X f(g) = \left. \frac{d}{dt} f(g \exp(tX)) \right|_{t=0}$$

for every square integrable function f on G . This defines a left-invariant derivation on the space of functions. If one studies vector fields, one would learn that they are derivations, and that the L_X are exactly the left-invariant vector fields. In practice, one computes the vector fields as we did it in the case of $SU(2)$. That is one tries to find a suitable parameterization of the Lie group, and then one computes derivatives that one can combine so that they act as the left-invariant vector fields. The right invariant vector fields are defined analogously

$$R_X f(g) = \left. \frac{d}{dt} f(\exp(-tX)g) \right|_{t=0}$$

In summary, we have seen how the action of the Lie group induces one of the underlying Lie algebra. The Lie algebra of a Lie group is the Lie algebra of infinitesimal transformation along the directions tangent to the Lie group. Whether we study the action of the Lie group, or its Lie algebra is thus very much related. Especially, every irreducible unitary representation of a compact Lie group is also an irreducible unitary one of its Lie algebra. The Peter Weyl theorem can thus be rephrased.

Theorem 41. (Peter and Weyl)

Let $\mathcal{H} = \mathcal{L}^2_\mu(G)$ be the Hilbert space of square integrable functions of a compact Lie group G . Then \mathcal{H} is a unitary representation of $\text{Lie}(G)$. Moreover, let \mathcal{R} be the set of all finite-dimensional unitary representations of $\text{Lie}(G)$, then under the left-right action of $\text{Lie}(G)$, the Hilbert space

decomposes as

$$\mathcal{L}_\mu^2(G) = \bigoplus_{\rho \in \mathcal{R}} \rho^L \otimes \bar{\rho}^R.$$

This is exactly our final result of the harmonic analysis on $SU(2)$. Recall, that once we knew all irreducible representations of the Lie algebra $su(2)$, we had to find functions that transform in these representations under the action of the vector fields. finding this action amounted to solving a set of first order differential equation. Our next goal is thus to understand the representation theory of the Lie algebras of compact Lie groups. Before we turn to this goal, we will look at another example of harmonic analysis. Unfortunately, for larger Lie groups than $SU(2)$ harmonic analysis becomes very cumbersome. However, there are also non-compact Lie groups and Lie supergroups. Studying them is usually much harder than compact Lie groups. But there is one Lie supegroup, that is the super-analogue of the commutative Lie group \mathbb{R} . This Lie supergroup can be nicely explicitly studied. In order to repeat and to illustrate some of the objects we have introduced we will study this Lie supergroup now.

5.4. Example of a Lie supergroup. Lie supergroups are a natural generalization of Lie groups. From a physics point of view, they describe physics in a somehow supersymmetric world. Whether such a world is realistic is of course very arguable. However, there is one application. The biggest success of string theory is the AdS/CFT conjecture of Juan Maldacena. It says that string theory on a supermanifold of Anti-de-Sitter type is dual to super Yang-Mills theory on the asymptotic boundary of the manifold. This conjecture is so important, because gauge theories like Yang-Mills only allow for a theoretical description within a weakly coupled regime. This duality gives then a description of a fairly realistic gauge theory beyond the weakly coupled regime. The supermanifolds appearing in this description are Lie supergroups and their cosets.

We are interested in the Lie supergroup $GL(1|1)$. Its Lie algebra is called $gl(1|1)$ and it is generated by four elements, two bosonic ones E, N and two fermionic ones ψ^\pm . The super-commutators are

$$\begin{aligned} [E, N] &= [N, E] = [E, \psi^\pm] = [\psi^\pm, E] = 0 \\ [N, \psi^\pm] &= -[\psi^\pm, N] = \pm \psi^\pm \\ [\psi^+, \psi^-] &= [\psi^-, \psi^+] = E, \quad [\psi^\pm, \psi^\pm] = 0. \end{aligned}$$

As Lie algebras can often be thought of as Lie algebras of matrices, we can do the same for this Lie superalgebra. Namely the identification

$$\begin{aligned} \rho_{e,n}(E) &= \begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix}, & \rho_{e,n}(N) &= \begin{pmatrix} n & 0 \\ 0 & n-1 \end{pmatrix} \\ \rho_{e,n}(\psi^+) &= \begin{pmatrix} 0 & e \\ 0 & 0 \end{pmatrix}, & \rho_{e,n}(\psi^-) &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

defines a Lie superalgebra homomorphism from $gl(1|1)$ into the space of 2×2 supermatrices. Here all the name super means is, that the commutator of two matrices A and B is defined as

$$[A, B] = A_0 B_0 - B_0 A_0 + A_0 B_1 - B_1 A_0 + A_1 B_0 - B_0 A_1 + A_1 B_1 + B_1 A_1,$$

where for

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A_0 = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}.$$

The even subalgebra is generated by E and N , and these two elements commute. Hence the even subalgebra is just the two-dimensional commutative Lie algebra and its Lie group is \mathbb{R}^2 (or if we compactify $S^1 \times S^1$). While a Lie superalgebra is not a Lie algebra, a Lie supergroup is a Lie group, but not over the complex numbers. We also have to allow fermionic numbers. Such numbers are called odd Grassmann numbers, and two such numbers η_{\pm} satisfy

$$\eta_{\pm}^2 = 0, \quad \eta_+ \eta_- = -\eta_- \eta_+.$$

We have learnt in the previous sections that it is a good idea to think about the Lie group of a Lie algebra as the exponentials of the Lie algebra. For our case the Lie algebra consists of elements of the form

$$aE + bN + \theta_+ \psi^+ + \theta_- \psi_-,$$

where a and b are Grassmann even numbers (like for example complex numbers) and η_{\pm} are Grassmann odd. Hence a Lie supergroup element is of the form

$$\exp(xE + yN + \theta_+ \psi^+ + \theta_- \psi_-). \quad (5.2)$$

If you prefer matrices, then our above Lie superalgebra homomorphism induces one on the Lie supergroup and the image are invertible matrices of the form

$$\begin{pmatrix} e^x & \theta_+ \\ \theta_- & e^y \end{pmatrix} \quad (5.3)$$

for Grassmann even numbers x, y and odd ones θ_{\pm} . We thus have two ways to parameterize the Lie supergroup. Let us now use both of them to compute the invariant vector fields. The left-invariant ones are defined by

$$L_X g = gX$$

for X in E, N, ψ^+ and $L_- g = -g\psi^-$, respectively by

$$L_X^{e,n} g = g\rho_{e,n}(X)$$

for X in E, N, ψ^+ and $L_-^{e,n} g = -g\rho_{e,n}(\psi^-)$. We start with the matrix form and proceed as in the $SU(2)$ case. We use the $\rho_{1,1/2}$ representation as it is fairly symmetric.

(1) We first modify our parameterization a little bit to

$$g(x, y, \theta_{\pm}) = e^x \begin{pmatrix} e^y & \theta_+ \\ \theta_- & e^{-y} \end{pmatrix}$$

(2) We then compute

$$\begin{aligned} \frac{d}{dx} g &= g, & \frac{d}{dy} g &= e^x \begin{pmatrix} e^y & 0 \\ 0 & -e^{-y} \end{pmatrix}, \\ \frac{d}{d\theta_+} g &= e^x \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, & \frac{d}{d\theta_-} g &= e^x \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \end{aligned} \quad (5.4)$$

(3) We then compute the left action of the corresponding matrices

$$\begin{aligned}
g \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= g \\
g \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &= \frac{1}{2} e^x \begin{pmatrix} e^y & -\theta_+ \\ \theta_- & -e^{-y} \end{pmatrix} \\
g \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} &= e^x \begin{pmatrix} \theta_- & e^{-y} \\ 0 & 0 \end{pmatrix} \\
g \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} &= e^x \begin{pmatrix} 0 & 0 \\ e^y & \theta_+ \end{pmatrix}
\end{aligned} \tag{5.5}$$

(4) Combining (5.4) and (5.5), we then get the invariant vector fields

$$\begin{aligned}
L_E^{1,1/2} &= \frac{d}{dx} \\
L_N^{1,1/2} &= \frac{1}{2} \left(\frac{d}{dy} - \theta_+ \frac{d}{d\theta_+} + \theta_- \frac{d}{d\theta_-} \right) \\
L_+^{1,1/2} &= e^{-y} \left(\frac{\theta_-}{2} \left(\frac{d}{dx} + \frac{d}{dy} \right) + \left(1 - \frac{\theta_- \theta_+}{2} \right) \frac{d}{d\theta_+} \right) \\
L_-^{1,1/2} &= -e^y \left(\frac{\theta_+}{2} \left(\frac{d}{dx} - \frac{d}{dy} \right) + \left(1 - \frac{\theta_+ \theta_-}{2} \right) \frac{d}{d\theta_-} \right)
\end{aligned} \tag{5.6}$$

In order to redo an analogous computation for the parameterization (5.2), it is again better to pass to a modified parameterization. In this case it is

$$g = e^{\theta_- \psi^-} e^{xE+yN} e^{\theta_+ \psi^+}.$$

One can now proceed as in the previous case in computing the invariant vector fields. We will only give the result.

$$\begin{aligned}
L_E &= \frac{d}{dx}, & L_N &= \frac{d}{dy} - \theta_+ \frac{d}{d\theta_+} \\
L_+ &= \frac{d}{d\theta_+}, & L_- &= -e^{-y} \frac{d}{d\theta_-} + \theta_+ \frac{d}{dx}.
\end{aligned} \tag{5.7}$$

It is a good exercise with Grassmann variables that indeed these differential operators satisfy the defining equations of the invariant vector fields. Another good exercise is to verify that they satisfy the commutation relations of the Lie superalgebra. Having these vector fields let us compute the Haar measure. We proceed in the following steps

(1) The ordinary differential operators are in terms of the invariant vector fields

$$\frac{d}{dx} = L_E, \quad \frac{d}{dy} = L_N + \theta_+ L_+, \quad \frac{d}{d\theta_+} = L_+, \quad \frac{d}{d\theta_-} = -e^y L_- - e^y \theta_+ L_E.$$

(2) The Maurer-Cartan one form is

$$\begin{aligned}
g^{-1} dg &= \left(g^{-1} \frac{d}{dx} g \right) dx + \left(g^{-1} \frac{d}{dy} g \right) dy + \left(g^{-1} \frac{d}{d\theta_+} g \right) d\theta_+ + \left(g^{-1} \frac{d}{d\theta_-} g \right) d\theta_- \\
&= g^{-1} (L_E dx + (L_N + \theta_+ L_+) dy + L_+ d\theta_+ - e^y (L_- + \theta_+ L_E) d\theta_-) g \\
&= E dx + (N + \theta_+ \psi^+) dy + \psi^+ d\theta_+ + e^y (\psi^- + \theta_+ E) d\theta_- \\
&= E(dx + \theta_+ d\theta_-) + N dy + \psi^+ d\theta_+ + \psi^- e^y d\theta_-
\end{aligned}$$

(3) Hence the dual one forms are $\omega(E) = (dx + \theta_+ d\theta_-)$, $\omega(N) = dy$, $\omega(\psi^+) = d\theta_+$ and $\omega(\psi^-) = e^y d\theta_-$.

(4) the Haar measure is then almost the exterior power of the dual one forms, it is

$$d\mu = e^{-y} dx dy d\theta_- d\theta_+.$$

Functions on the supergroup are then functions on \mathbb{R}^2 combined with odd functions. More precisely a basis of functions consists of the following

$$f_0(e, n); = e^{iex+iny}, \quad f_{\pm}(e, n) = \theta_{\pm} f_0(e, n), \quad f_2(e, n) = \theta_+ \theta_- f_0(e, n).$$

We see that all $f_i(e, n)$ have L_E eigenvalue ie and that the L_n eigenvalue of $f_0(e, n)$ and $f_-(e, n)$ is in , while the one of $f_+(e, n)$ and $f_2(e, n)$ is $i(n - 1)$. We say that the vector $i(e, n)$ and $i(e, n - 1)$ are the weight of $f_0(e, n)$ and $f_-(e, n)$ respectively of $f_+(e, n)$ and $f_2(e, n)$. We also see that $f_0(e, n)$ and $f_-(e, n)$ are annihilated by L_+ . We call L_+ an annihilation operator and L_- a creation operator. Then in analogy to $sl(2)$, we can call $f_0(e, n)$ and $f_-(e, n)$ highest-weight vectors of highest-weight $i(e, n)$. The action of L_- then maps as

$$L_- f_0(e, n) = -ief_+(e, n), \quad L_- f_-(e, n) = -ief_2(e, n) - f_0(e, n - 1)$$

and $f_0(e, n)$ together with $f_+(e, n)$ forms a two-dimensional representation of $gl(1|1)$. The same is true for $f_-(e, n)$ and $ief_2(e, n) + f_0(e, n - 1)$. These representations are irreducible unless $e = 0$. In that case we get four-dimensional indecomposable but reducible modules. We will terminate our analysis of this example here. The purpose was, that you see once again how to explicitly analyse a Lie (super)group.

6. LIE ALGEBRAS

In the last section, we have already encountered Lie algebras and its representation theory. We will now turn to a more general study of it. I use the book by Humphries [Hu] as reference. We have already had the definition of a Lie algebra

Definition 33. A Lie algebra \mathfrak{g} is a vector space \mathfrak{g} together with a bilinear operation

$$[\ , \] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$$

satisfying

$$[U, T] = -[T, U] \quad (\text{antisymmetry})$$

and

$$[U, [S, T]] + [T, [U, S]] + [S, [T, U]] = 0 \quad \text{Jacobi identity}$$

for all U, S, T in \mathfrak{g} .

Let $\phi : \mathfrak{g} \rightarrow \mathfrak{g}'$ be a vector space homomorphism. It is called a Lie algebra homomorphism if both \mathfrak{g} and \mathfrak{g}' are Lie algebras and if

$$\phi([x, y]) = [\phi(x), \phi(y)]$$

for all x, y in \mathfrak{g} . If one wants to work concretely with Lie algebras, a useful concept are the structure constants. For this let \mathfrak{g} be a finite-dimensional Lie algebra and $\{x_1, \dots, x_n\}$ be a basis of \mathfrak{g} . Its structure constants are defined as

$$[x_i, x_j] = \sum_{k=1}^n f_{ij}^k x_k,$$

so that antisymmetry implies

$$0 = [x_i, x_j] + [x_j, x_i] = \sum_{k=1}^n (f_{ij}^k + f_{ji}^k) x_k$$

and hence

$$f_{ij}^k + f_{ji}^k = 0$$

for all i, j, k in $1, \dots, n$. The Jacobi identity implies

$$0 = \sum_{k=1}^n (f_{ij}^k f_{kl}^m + f_{jl}^k f_{ki}^m + f_{li}^k f_{kj}^m),$$

which can be derived analogously to the previous case. The main examples of this section will be the special linear Lie algebras, that are the Lie algebras of traceless square matrices. For small n , explicit bases are

Example 11.

- (1) for $n = 2$, we already have encountered $sl(2; \mathbb{R})$. It is generated by e, f, h . they correspond to the matrices

$$f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

and their commutation relations are

$$[h, e] = 2e, \quad [h, f] = -2f, \quad [e, f] = h.$$

In order to have the same notation as the following two examples one needs to redefine

$$h = h_\alpha, \quad e = e_\alpha, \quad f = f_\alpha.$$

- (2) The Lie algebra $sl(3; \mathbb{R})$ is generated by eight elements, which we call $f_{\alpha_1}, f_{\alpha_2}, f_{\alpha_1+\alpha_2}, h_{\alpha_1}, h_{\alpha_2}, e_{\alpha_1}, e_{\alpha_2}, e_{\alpha_1+\alpha_2}$. In terms of matrices, the f 's are the lower triangular matrices

$$f_{\alpha_1} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad f_{\alpha_2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad f_{\alpha_1+\alpha_2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

the h 's are the traceless diagonal matrices

$$h_{\alpha_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad h_{\alpha_2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

and the e 's are the upper triangular matrices

$$e_{\alpha_1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_{\alpha_2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_{\alpha_1+\alpha_2} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The non-zero commutation relations are

$$\begin{aligned}
[f_{\alpha_1}, f_{\alpha_2}] &= -f_{\alpha_1+\alpha_2}, & [e_{\alpha_1}, e_{\alpha_2}] &= e_{\alpha_1+\alpha_2}, \\
[h_{\alpha_1}, f_{\alpha_1}] &= -2f_{\alpha_1}, & [h_{\alpha_1}, e_{\alpha_1}] &= 2e_{\alpha_1}, \\
[h_{\alpha_1}, f_{\alpha_2}] &= f_{\alpha_2}, & [h_{\alpha_1}, e_{\alpha_2}] &= -e_{\alpha_2}, \\
[h_{\alpha_1}, f_{\alpha_1+\alpha_2}] &= -f_{\alpha_1+\alpha_2}, & [h_{\alpha_1}, e_{\alpha_1+\alpha_2}] &= e_{\alpha_1+\alpha_2}, \\
[h_{\alpha_2}, f_{\alpha_1}] &= f_{\alpha_1}, & [h_{\alpha_2}, e_{\alpha_1}] &= -e_{\alpha_1}, \\
[h_{\alpha_2}, f_{\alpha_2}] &= -2f_{\alpha_2}, & [h_{\alpha_2}, e_{\alpha_2}] &= 2e_{\alpha_2}, \\
[h_{\alpha_2}, f_{\alpha_1+\alpha_2}] &= -f_{\alpha_1+\alpha_2}, & [h_{\alpha_2}, e_{\alpha_1+\alpha_2}] &= e_{\alpha_1+\alpha_2}, \\
[e_{\alpha_1}, f_{\alpha_1}] &= h_{\alpha_1}, & [e_{\alpha_2}, f_{\alpha_2}] &= h_{\alpha_2}, & [e_{\alpha_1+\alpha_2}, f_{\alpha_1+\alpha_2}] &= h_{\alpha_1} + h_{\alpha_2}, \\
[e_{\alpha_1}, f_{\alpha_1+\alpha_2}] &= -f_{\alpha_2}, & [e_{\alpha_2}, f_{\alpha_1+\alpha_2}] &= f_{\alpha_1}, \\
[f_{\alpha_1}, e_{\alpha_1+\alpha_2}] &= e_{\alpha_2}, & [f_{\alpha_2}, e_{\alpha_1+\alpha_2}] &= -e_{\alpha_1}.
\end{aligned} \tag{6.1}$$

You see that these are quite a lot of relations.

- (3) For $sl(n; \mathbb{R})$, let $E_{i,j}$ be the matrix that has entry zero everywhere except a one at the j -th position of row number i . Then define

$$h_{\alpha_i} = E_{i,i} - E_{i+1,i+1}, \quad e_{\alpha_i} = E_{i,i+1}, \quad f_{\alpha_i} = E_{i+1,i}$$

and further generators are defined similarly to the $sl(3; \mathbb{R})$ case.

A Lie algebra \mathfrak{g} acts on itself via the commutator. This action is called the adjoint representation

$$\text{ad} : \mathfrak{g} \rightarrow \text{End}(\mathfrak{g}), \quad \text{ad}(x)(y) = [x, y].$$

More generally a representation ρ of a Lie algebra \mathfrak{g} is a vector space V together with a Lie algebra homomorphism

$$\rho : \mathfrak{g} \rightarrow \text{End}(V).$$

Where a Lie algebra homomorphism is defined as follows. A vector space homomorphism (over our usual field \mathbb{C} or \mathbb{R}) is a map

$$\rho : V \rightarrow W$$

of vector spaces that is linear, meaning that

$$\rho(ax + by) = a\rho(x) + b\rho(y)$$

for all a, b in our field and all x, y in our vector space V . Both a finite-dimensional Lie algebra \mathfrak{g} and the space of linear operators (matrices) on a vector space are themselves vector spaces. A vector space homomorphism

$$\rho : \mathfrak{g} \rightarrow \text{End}(V)$$

is called a Lie algebra homomorphism, if \mathfrak{g} is a Lie algebra and if

$$[\rho(x), \rho(y)] = \rho([x, y])$$

for all x, y in \mathfrak{g} . If V is a finite-dimensional vector space, then $\text{End}(V)$ is a subalgebra of the algebra of square matrices acting on V . Matrices have a very natural bilinear form, the trace. The trace is denoted by tr and it is the sum of the diagonal entries of a given matrix.

Definition 34. A bilinear form $B : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{C}$ is called *non-degenerate* if for every non-zero x in \mathfrak{g} there exists y in \mathfrak{g} with $B(x, y) \neq 0$. The bilinear form is called *symmetric* if

$$B(x, y) = B(y, x)$$

and *invariant* if

$$B(x, [y, z]) = B([x, y], z)$$

for all x, y, z in \mathfrak{g} .

Proposition 42. Let A, B, C be three $n \times n$ matrix then

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

and

$$\operatorname{tr}(A[B, C]) = \operatorname{tr}([A, B]C).$$

Proof. Let a_{ij} be the entry on position (i, j) of the matrix A and b_{ij} and c_{ij} correspondingly for the matrices B and C . The trace is the sum over the diagonal elements, hence

$$\operatorname{tr}(AB) = \sum_{1 \leq i, j \leq n} a_{ij}b_{ji} = \sum_{1 \leq i, j \leq n} b_{ij}a_{ji} = \operatorname{tr}(BA).$$

Using this identity for the two matrices AC and B , we compute

$$\begin{aligned} \operatorname{tr}(A[B, C]) - \operatorname{tr}([A, B]C) &= \operatorname{tr}(A(BC - CB)) - \operatorname{tr}((AB - BA)C) \\ &= \operatorname{tr}(-ACB + BAC) \\ &= \operatorname{tr}(BAC - BAC) = 0. \end{aligned}$$

■

This is nice, it tells us that every matrix representation of a Lie algebra defines an invariant symmetry bilinear form

Corollary 43. Let $\rho : \mathfrak{g} \rightarrow \operatorname{End}(V)$ be a finite-dimensional representation of a finite-dimensional Lie algebra \mathfrak{g} , then

$$B(x, y) = \operatorname{tr}(\rho(x)\rho(y))$$

defines an invariant symmetric bilinear form on \mathfrak{g} .

A given Lie algebra surely has many finite-dimensional representations. The question is now, how are these related and if there is at least one such bilinear form that is non-degenerate.

Definition 35. Let \mathfrak{g} be a Lie algebra and I a subset of \mathfrak{g} . We call I an *ideal* if for every x in \mathfrak{g} and for every y in I also $[x, y]$ is in I . So especially every ideal is a sub Lie algebra.

A Lie algebra is called *simple* if it is not abelian and it has no non-trivial ideal.

Let B be a symmetric invariant bilinear form of \mathfrak{g} . The *radical* S_B is the set of all degenerate elements in the sense

$$S_B = \{x \in \mathfrak{g} \mid B(x, y) = 0 \text{ for all } y \in \mathfrak{g}\}.$$

Proposition 44. The radical S_B of a Lie algebra \mathfrak{g} with symmetric, invariant bilinear form B is an ideal of \mathfrak{g} .

Proof. Let x in S_B and y in \mathfrak{g} , we have to show that $[x, y]$ is in S_B . For this let z be an arbitrary element of \mathfrak{g} . By definition $B(x, z) = 0$, we have to show that also $B([x, y], z) = 0$. By invariance of the bilinear form $B([x, y], z) = B(x, [y, z])$ but the latter vanishes since x has zero product with every other element of the Lie algebra. ■

So we see that a bilinear form can only be non-degenerate if its radical vanishes. This especially happens for a simple Lie algebra as such a Lie algebra has no non-trivial ideal.

Definition 36. A Lie algebra is called *semi-simple* if it is the direct sum of simple Lie algebras.

The *Killing form* of a Lie algebra is the invariant symmetric bilinear form defined by the trace in the adjoint representation, that is

$$\kappa(x, y) = \text{tr}(\text{ad}(x)\text{ad}(y)).$$

If you want to, you can express the Killing form in terms of the structure constants. That is in the adjoint representation, the basis vector x_i acts by multiplication with the matrix f_{ij}^k and hence

$$\kappa(x_i, x_j) = \sum_{1 \leq m, n \leq \dim(\mathfrak{g})} f_{im}^n f_{jn}^m.$$

For an abelian Lie algebra all structure constants are identical zero and hence the Killing form is identical zero. A nice structural theorem is

Theorem 45. A finite-dimensional Lie algebra is semi-simple if and only if its Killing form is non-degenerate.

Both having a non-degenerate bilinear form and being simple (having no ideals) are important properties that usually hold in questions of interest in physics.

Example 12. Let us look back at $s(3; \mathbb{R})$. We have seen that this Lie algebra is the Lie algebra of traceless 3×3 matrices. This means that the Lie algebra acts on three-dimensional real vector space. The corresponding invariant symmetric bilinear form is easy to read off:

$$\begin{aligned} B(h_{\alpha_1}, h_{\alpha_1}) &= B(h_{\alpha_2}, h_{\alpha_2}) = 2, & B(h_{\alpha_1}, h_{\alpha_2}) &= -1 \\ B(f_{\alpha_1}, e_{\alpha_1}) &= B(f_{\alpha_2}, e_{\alpha_2}) = B(f_{\alpha_1+\alpha_2}, e_{\alpha_1+\alpha_2}) = 1. \end{aligned}$$

Let us choose the ordered basis $\{f_{\alpha_1}, f_{\alpha_2}, f_{\alpha_1+\alpha_2}, h_{\alpha_1}, h_{\alpha_2}, e_{\alpha_1}, e_{\alpha_2}, e_{\alpha_1+\alpha_2}\}$ of $sl(3; \mathbb{R})$. Then in this basis

$$\text{ad}(h_{\alpha_1}) = \text{diag}(-2, 1, -1, 0, 0, 2, -1, 1), \quad \text{ad}(h_{\alpha_2}) = \text{diag}(1, -2, -1, 0, 0, -1, 2, 1)$$

and hence

$$\kappa(h_{\alpha_1}, h_{\alpha_1}) = \kappa(h_{\alpha_2}, h_{\alpha_2}) = 12, \quad \kappa(h_{\alpha_1}, h_{\alpha_2}) = -6.$$

The invariant bilinear form is uniquely determined by these three products. The reason is that the bilinear form is invariant and symmetric. It is an instructive exercise to show that. We observe that $\kappa = 6B$, the two bilinear forms only differ by a scalar.

This is actually not a coincidence, an important implication of the Lemma of Schur is

Theorem 46. Up to a scalar multiple the invariant symmetric non-degenerate bilinear form of a simple finite-dimensional Lie algebra is unique.

This statement is a Lie algebra analogue of the uniqueness of both left-right invariant Haar measure and Laplacian of compact Lie groups. Recall that the Laplacian is constructed from the Casimir element of the underlying Lie algebra. We will now discuss this special and important element.

6.1. The Casimir element of a representation. If you study algebra or representation theory you will often hear the statement this follows from Schur's Lemma, exactly as I just said for the uniqueness of the Killing form. Now I would like to tell you how it implies the uniqueness of the Casimir operator, which is the Lie algebra analogue of the Laplacian, which in turn played the role of the Hamiltonian of a system with compact Lie group symmetry. The important class of representations are irreducible representations and completely reducible ones.

Definition 37. Let $\rho : \mathfrak{g} \rightarrow \text{End}(V)$ be a representation of a Lie algebra. A *sub representation* of ρ is a sub vector space $W \subset V$ with the property that $\rho(x)w$ in W for every x in \mathfrak{g} and for every w in W .

A representation is called *irreducible* if there is no non-trivial subrepresentation, where V itself and the zero vector are the two trivial sub representations.

A representation is called *completely reducible*, if V is the direct sum

$$V = \bigoplus V_i$$

of vector spaces V_i such that each V_i is an irreducible sub representation of ρ .

Lemma 47. Schur's Lemma

Let $\rho : \mathfrak{g} \rightarrow \text{End}(V)$ be an irreducible representation of \mathfrak{g} . Then the only elements of $\text{End}(V)$ that commute with the image of \mathfrak{g} under ρ are the scalars, that are the scalar multiples of the identity matrix if V is finite-dimensional and if we identify the endomorphism ring with the space of square matrices.

Let \mathfrak{g} be a simple finite-dimensional Lie algebra. If you wish you can generalize to semi-simple Lie algebras, that are direct sums of simple ones. Let $\rho : \mathfrak{g} \rightarrow \text{End}(V)$ be an irreducible representation and let B_ρ the trace of matrices representing the Lie algebra in the endomorphism ring of V with respect to some basis of V . Let us choose a basis $\{x_1, \dots, x_n\}$ of \mathfrak{g} , then there exists a dual basis $\{y_1, \dots, y_n\}$ satisfying

$$B_\rho(x_i, y_j) = \delta_{i,j}.$$

We want to use the bilinear form to construct an operator, that is an element in the endomorphism ring of V that commutes with all x in \mathfrak{g} . So let x be an arbitrary element, then we use our two bases to define matrices via

$$[x, x_i] = \sum_{j=1}^n a_{ij}x_j, \quad [x, y_i] = \sum_{j=1}^n b_{ij}y_j.$$

The matrix b whose entries are the b_{ij} is minus one times the transpose of the matrix a with entries a_{ij} . This can be seen by the following computation

$$\begin{aligned} a_{ik} &= \sum_{j=1}^n a_{ij}B_\rho(x_j, y_k) = B_\rho([x, x_i], y_k) = B_\rho(-[x_i, x], y_k) \\ &= -B_\rho(x_i, [x, y_k]) = -\sum_{j=1}^n b_{kj}B_\rho(x_i, y_j) \\ &= -b_{ki}. \end{aligned}$$

Here we used anti-symmetry of the Lie bracket as well as invariance and linearity of the bilinear form. We define the *Casimir operator* of the representation ρ as

$$C_\rho = \sum_{i=1}^n \rho(x_i)\rho(y_i).$$

This is by definition an element of the endomorphism ring of V . It is not clear yet, that this definition is independent of our choice of basis.

Proposition 48. The Casimir operator commutes with $\rho(x)$ for every x in \mathfrak{g} .

Proof. In the endomorphism ring we have the identity

$$\begin{aligned} [\rho(x), \rho(y)\rho(z)] &= \rho(x)\rho(y)\rho(z) - \rho(y)\rho(z)\rho(x) \\ &= (\rho(x)\rho(y) - \rho(y)\rho(x))\rho(z) + \rho(y)(\rho(x)\rho(z) - \rho(z)\rho(x)) \\ &= [\rho(x), \rho(y)]\rho(z) + \rho(y)[\rho(x), \rho(z)]. \end{aligned}$$

for all x, y, z in \mathfrak{g} . Note, that in this argument we could have replaced $\rho(x), \rho(y), \rho(z)$ with any three endomorphisms. This equality holds because the endomorphism ring is associative, which is true because matrix multiplication is associative. Remember, you proved this statement in your very first homework problems.

We thus get

$$\begin{aligned} [\rho(x), C_\rho] &= \sum_{i=1}^n [\rho(x), \rho(x_i)\rho(y_i)] \\ &= \sum_{i=1}^n ([\rho(x), \rho(x_i)]\rho(y_i) + \rho(x_i)[\rho(x), \rho(y_i)]) \\ &= \sum_{1 \leq i, j \leq n} (a_{ij}\rho(x_j)\rho(y_i) + b_{ij}\rho(x_i)\rho(y_j)) \\ &= 0. \end{aligned}$$

So indeed the Casimir operator commutes with every element $\rho(x)$ for every x in \mathfrak{g} . ■

Corollary 49. Let $\rho : \mathfrak{g} \rightarrow \text{end}(V)$ be an irreducible finite-dimensional representation of a simple Lie algebra, then C_ρ is the endomorphism that acts by multiplication with the scalar

$$c_\rho = \frac{\dim(\mathfrak{g})}{\dim(V)}.$$

Especially the definition of the Casimir operator is independent of the choice of basis.

Proof. We are precisely in the situation of Schur's Lemma, so C_ρ must act as a scalar, that is if we choose a basis of V and let every endomorphism be represented by its matrix in this representation, then the Casimir operator is a scalar c_ρ times the identity matrix, $C_\rho = c_\rho Id$. In order to compute this number, we take the trace

$$\text{tr}(C_\rho) = \text{tr}(c_\rho Id) = \dim(V)c_\rho,$$

but also

$$\text{tr}(C_\rho) = \sum_{i=1}^n \text{tr}(\rho(x_i)\rho(y_i)) = \sum_{i=1}^n B_\rho(x_i, y_i) = \sum_{i=1}^n 1 = n.$$

The result follows since the dimension of \mathfrak{g} is n . ■

Example 13. Let \mathfrak{g} be $sl(2; \mathbb{R})$ and $V = \mathbb{R}^2$. Let ρ be the mapped we used to define the Lie algebra, that is

$$\rho(x) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \rho(h) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \rho(f) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Then the bilinear form is

$$B_\rho(e, f) = 1, \quad B_\rho(h, h) = 2.$$

So that a dual basis is given by $e, h/2, f$. The Casimir operator is thus

$$C = xy + \frac{1}{2}hh + yx.$$

Compare this result with the construction of the Laplacian in our harmonic analysis of $SU(2)$. Can you see that up to a scalar they are the same? The scalar is just a normalization. We thus find

$$\begin{aligned} C_\rho &= \rho(x)\rho(y) + \frac{1}{2}hh + \rho(y)\rho(x) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \frac{3}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Note, that indeed

$$\frac{3}{2} = \frac{\dim(\mathfrak{sl}(2; \mathbb{R}))}{\dim(\mathbb{R}^2)}.$$

6.2. Jordan Decomposition. In linear algebra, the main goal is to find procedures to bring matrices in a nice form. In general, the nicest possible form is the Jordan normal form. For complex matrices it reads

Theorem 50. Let A be an $n \times n$ matrix over the complex numbers, and let $P_A = (\lambda_1 - t)^{r_1} \dots (\lambda_m - t)^{r_m}$ be its characteristic polynomial. Then there exists an invertible matrix S , such that

$$SAS^{-1} = \begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & D_m \end{pmatrix}$$

where D_i is a $r_i \times r_i$ matrix of the form

$$D_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & \dots & 0 & \lambda_i \end{pmatrix}$$

Note, that especially SAS^{-1} splits into a diagonal matrix D and a nilpotent (upper triangular) matrix N such that $[D, N] = 0$. If we are working over another field \mathbb{K} than the complex numbers, as for example the real ones, the Jordan form is

Theorem 51. Let A be an $n \times n$ matrix over a field \mathbb{K} , and let $P_A = (\lambda_1 - t)^{r_1} \dots (\lambda_m - t)^{r_m}$ be its characteristic polynomial. Then there exists an invertible matrix S , such that

$$SAS^{-1} = \begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & D_m \end{pmatrix}$$

where D_i is a $r_i \times r_i$ matrix of the form

$$D_i = \begin{pmatrix} \lambda_i & * & * & \dots & * \\ 0 & \lambda_i & * & \dots & * \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & \dots & 0 & \lambda_i \end{pmatrix}$$

such that

$$SAS^{-1} = D + N$$

where D is a diagonal matrix and N is a nilpotent upper triangular matrix with $[D, N] = 0$.

An $n \times n$ matrix represents an endomorphism of n -dimensional vector space with respect to some chosen basis. The Jordan normal form tells us that there exists another basis, related to the first one by the basis change matrix S , such that the endomorphism is represented by a matrix in Jordan form. Especially this matrix (and hence the endomorphism) splits into a diagonal part and a nilpotent one. The preimage under the basis change of the diagonal part, $S^{-1}DS$, is usually called the semi-simple part of the endomorphism and the one of the upper triangular one $S^{-1}NS$ is called the nilpotent one. The decomposition

$$A = S^{-1}DS + S^{-1}NS$$

is then called the Jordan decomposition of A .

A Lie algebra acts on itself and this action is called the adjoint representation. Let \mathfrak{g} be a n -dimensional Lie algebra, then the adjoint representation is represented by $n \times n$ matrices in some given basis of \mathfrak{g} . Let x be an arbitrary element of \mathfrak{g} and

$$\text{ad}(x) = S + N$$

be the Jordan decomposition of $\text{ad}(x)$. It defines uniquely elements s and n of \mathfrak{g} such that $\text{ad}(s) = S$ and $\text{ad}(n) = N$. The decomposition

$$x = s + n$$

is called the abstract Jordan decomposition of x . The important theorem is

Theorem 52. Let $\mathfrak{g} \subset \mathfrak{gl}(\mathbb{R}^n)$ be a finite-dimensional Lie algebra.

- (1) Then \mathfrak{g} contains the semi-simple and nilpotent parts in $\mathfrak{gl}(\mathbb{R}^n)$ of all its elements. In particular the abstract and the usual Jordan decompositions in \mathfrak{g} coincide.
- (2) Let $\rho : \mathfrak{g} \rightarrow \text{End}(V)$ be a finite-dimensional representation of \mathfrak{g} . If $x = s + n$ is the abstract Jordan decomposition of x in \mathfrak{g} , then $\rho(x) = \rho(s) + \rho(n)$ is the usual Jordan decomposition of $\rho(x)$.

Example 14. Consider our prime example $sl(3; \mathbb{R})$. We know all Lie brackets of this algebra (6.1), but we also know the matrix form of elements in the three-dimensional representation. So that by the previous theorem, we know that h_{α_1} and h_{α_2} are the semi-simple elements and $f_{\alpha_1}, f_{\alpha_2}, f_{\alpha_1+\alpha_2}, e_{\alpha_1}, e_{\alpha_2}, e_{\alpha_1+\alpha_2}$ are the nilpotent ones. It is instructive to verify this for an example. For example, let us look at the action of e_{α_1} . We find

$$\begin{aligned} [e_{\alpha_1}, f_{\alpha_1+\alpha_2}] &= -f_{\alpha_2}, & [e_{\alpha_1}, f_{\alpha_2}] &= 0 \\ [e_{\alpha_1}, f_{\alpha_1}] &= h_{\alpha_1}, & [e_{\alpha_1}, h_{\alpha_1}] &= -2e_{\alpha_1}, & [e_{\alpha_1}, e_{\alpha_1}] &= 0 \\ [e_{\alpha_1}, h_{\alpha_2}] &= e_{\alpha_1} \\ [e_{\alpha_1}, e_{\alpha_2}] &= e_{\alpha_1+\alpha_2}, & [e_{\alpha_1}, e_{\alpha_1+\alpha_2}] &= 0 \end{aligned}$$

so that $\text{ad}(e_{\alpha_1})^3 = 0$ and e_{α_1} is indeed a nilpotent element. We see that we actually have constructed our basis of $sl(3; \mathbb{R})$ according to some Jordan decomposition.

6.3. Root Space Decomposition. We will now use the Jordan decomposition to find convenient decompositions and hence convenient bases of Lie algebras. For this, let \mathfrak{g} be a simple

Lie algebra. The semi-simple elements with respect to the abstract Jordan decomposition form a subalgebra.

Definition 38. Let \mathfrak{g} be a simple (or semi-simple) Lie algebra, then a subalgebra of semi-simple elements is called a *toral* subalgebra. A maximal toral subalgebra is called the *Cartan subalgebra* of \mathfrak{g} and it is usually denoted by \mathfrak{h} .

Proposition 53. A toral subalgebra is abelian, and the Cartan subalgebra is a maximal abelian subalgebra.

The reason for this statement is that semi-simple elements can be diagonalized and diagonal matrices commute. So think of the Cartan subalgebra as the subalgebra of the diagonal elements of your favourite matrix representation of your Lie algebra.

Example 15.

- (1) In the case of $sl(2; \mathbb{R})$ the Cartan subalgebra is the one-dimensional abelian Lie algebra spanned by h .
- (2) In the case of $sl(3; \mathbb{R})$ the Cartan subalgebra is the two-dimensional abelian Lie algebra spanned by h_{α_1} and h_{α_2} .

Of course a Cartan subalgebra is not unique, but for a simple Lie algebra it is unique up to automorphisms of the Lie algebra.

So the Cartan subalgebra is a maximal abelian subalgebra. Hence in the adjoint representation, the elements of the Cartan subalgebra form a set of commuting endomorphisms (matrices). What do we know from linear algebra? A very important statement is that commuting endomorphisms are simultaneously diagonalizable, especially using the adjoint action the Lie algebra decomposes into a direct sum of common eigenspaces of the elements of the Cartan subalgebra. Call our Lie algebra \mathfrak{g} , and its Cartan subalgebra \mathfrak{h} . Let \mathfrak{h}^* be its dual space and let e_α be an element in an eigenspace with eigenvalues defining a linear functional α in \mathfrak{h}^* such that

$$[h, e_\alpha] = \alpha(h)e_\alpha.$$

Compare this situation to our examples of $sl(2; \mathbb{R})$ and $sl(3; \mathbb{R})$. In other words, there exists a finite subset Δ of \mathfrak{h}^* such that

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha,$$

where

$$\mathfrak{g}_\alpha = \{x \in \mathfrak{g} \mid [h, x] = \alpha(h)x\}$$

and α in Δ if and only if \mathfrak{g}_α is non-trivial. The space \mathfrak{g}_α is called the *root space* of the *root* α and the decomposition of \mathfrak{g} into root spaces is called the *root space decomposition* of \mathfrak{g} .

Example 16. Let us look at our standard examples

- (1) For $\mathfrak{g} = sl(2; \mathbb{R})$ the Cartan subalgebra is one-dimensional and hence its dual space is also one-dimensional. Remember that we introduced the somehow stupid looking notation

$$h = h_\alpha, \quad e = e_\alpha, \quad f = f_\alpha$$

so that the Lie bracket reads

$$[h_\alpha, e_\alpha] = 2e_\alpha, \quad [h_\alpha, f_\alpha] = -2f_\alpha, \quad [e_\alpha, f_\alpha] = h_\alpha.$$

Define the element α in \mathfrak{h}^* by $\alpha(h_\alpha) = 2$, then the relations become

$$[h_\alpha, e_\alpha] = \alpha(h_\alpha)e_\alpha, \quad [h_\alpha, f_\alpha] = -\alpha(h_\alpha)f_\alpha, \quad [e_\alpha, f_\alpha] = h_\alpha,$$

so that e_α spans the one-dimensional roots space \mathfrak{g}_α and f_α spans the one-dimensional root space $\mathfrak{g}_{-\alpha}$.

- (2) For $\mathfrak{g} = sl(3; \mathbb{R})$ the situation becomes more complicated. The Cartan subalgebra is two-dimensional as it is spanned h_{α_1} and h_{α_2} . We define the roots α_1 and α_2 as the two linear functions on \mathfrak{h} defined by

$$\alpha_1(h_{\alpha_1}) = 2, \quad \alpha_1(h_{\alpha_2}) = -1, \quad \alpha_2(h_{\alpha_1}) = -1, \quad \alpha_2(h_{\alpha_2}) = 2.$$

If you look back at the commutation relations of $sl(3; \mathbb{R})$, we see that the definition is made to satisfy

$$[h, e_\beta] = \beta(h)e_\beta, \quad [h, f_\beta] = -\beta(h)f_\beta$$

for β in $\{\alpha_1, \alpha_2, \alpha_1 + \alpha_2\}$. Especially, we see that each e_β spans the one-dimensional root space \mathfrak{g}_β and each f_β the one-dimensional root space $\mathfrak{g}_{-\beta}$. But we see more if we look very carefully at the relations of $sl(3; \mathbb{R})$. Namely, we see that

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}.$$

This is actually a property that has to follow from the Jacobi identity. Moreover looking at the Killing form, we observe (with $\mathfrak{g}_0 = \mathfrak{h}$) that if $\alpha + \beta \neq 0$, then \mathfrak{g}_α is orthogonal to \mathfrak{g}_β relative to the Killing form. Finally, we also observed previously that all e_α and f_α are nilpotent.

The properties illustrated in the example all hold in generality:

Proposition 54. Let \mathfrak{g} be a simple finite-dimensional Lie algebra with Cartan subalgebra \mathfrak{h} . Let α and β be roots, then the following are true:

- (1) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$
- (2) x in \mathfrak{g}_α for $\alpha \neq 0$ is nilpotent
- (3) if $\alpha + \beta \neq 0$, then \mathfrak{g}_α is orthogonal to \mathfrak{g}_β relative to the Killing form
- (4) the restriction of the Killing form to \mathfrak{h} is non-degenerate.

The non-degeneracy of the Killing form restricted to \mathfrak{h} is very nice, as it allows us to identify \mathfrak{h} with \mathfrak{h}^* via h_α is the unique element with the property that $\kappa(h_\alpha, h) = \alpha(h)$ for all h in \mathfrak{h} . It is a nice exercise that indeed this holds (depending on your normalization of Killing form of course) in our examples of $sl(2; \mathbb{R})$ and $sl(3; \mathbb{R})$.

Theorem 55. Let \mathfrak{g} be a simple Lie algebra with Cartan subalgebra \mathfrak{h} and root system Δ , then

- (1) Δ spans \mathfrak{h}^*
- (2) If α in Δ then $-\alpha$ in Δ
- (3) For x in \mathfrak{g}_α and y in $\mathfrak{g}_{-\alpha}$ we have $[x, y] = \kappa(x, y)h_\alpha$.
- (4) If α is in Δ , then $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is one-dimensional with basis h_α
- (5) $\alpha(h_\alpha) = \kappa(h_\alpha, h_\alpha) \neq 0$ for α in Δ
- (6) If α in Δ and x_α is any nonzero element in \mathfrak{g}_α , then there exists y_α in $\mathfrak{g}_{-\alpha}$ such that x_α, y_α and $t_\alpha = [x_\alpha, y_\alpha]$ span the three-dimensional subalgebra isomorphic to $sl(2; \mathbb{R})$
- (7) $t_\alpha = \frac{2h_\alpha}{\kappa(h_\alpha, h_\alpha)}$

Proof.

- (1) If Δ does not span \mathfrak{h}^* , then there exists (by duality) non-zero h in \mathfrak{h} such that $\alpha(h) = 0$ for all α in Δ , but this forces h to commute with all x in \mathfrak{g} which cannot be true since \mathfrak{g} is simple.
- (2) Assume that α in Δ but $-\alpha$ not, then $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}) = 0$ and we get a contradiction to the non-degeneracy of the Killing form.
- (3) Let x in \mathfrak{g}_α and y in $\mathfrak{g}_{-\alpha}$ and h an arbitrary element of \mathfrak{h} . Then

$$\begin{aligned}\kappa(h, [x, y]) &= \kappa([h, x], y) = \alpha(h)\kappa(x, y) = \kappa(h_\alpha, h)\kappa(x, y) = \kappa(\kappa(x, y)\mathfrak{h}_\alpha, h) \\ &= \kappa(h, \kappa(x, y)\mathfrak{h}_\alpha)\end{aligned}$$

so that by the non-degeneracy of the restriction of the Killing form to \mathfrak{h} the claim $[x, y] - \kappa(x, y)\mathfrak{h}_\alpha = 0$ follows.

- (4) The last statement says that $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is spanned by \mathfrak{h}_α as long as $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \neq 0$. But if for non-zero x in \mathfrak{g}_α , we have $\kappa(x, y) = 0$ for all y in $\mathfrak{g}_{-\alpha}$ then we get a contradiction to the non-degeneracy of the Killing form.
- (5) The proof of this statement will be omitted because we lack some preparation for it
- (6) Let x be a non-zero element of \mathfrak{g}_α , then by the last statement and since $\kappa(x, \mathfrak{g}_{-\alpha}) \neq 0$, we can find an element y in $\mathfrak{g}_{-\alpha}$ with $\kappa(x, y) = \frac{2}{\kappa(h_\alpha, h_\alpha)}$. Define

$$t_\alpha = \frac{2h_\alpha}{\kappa(h_\alpha, h_\alpha)}$$

then by (3) $[x, y] = t_\alpha$. Moreover,

$$[t_\alpha, x] = \frac{2}{\alpha(h_\alpha)}[h_\alpha, x] = \frac{2\alpha(h_\alpha)}{\alpha(h_\alpha)}x = 2x$$

and similarly $[t_\alpha, y] = -2y$.

- (7) This follows directly from the proof of the previous statement. ■

We have seen that roots are very efficient in describing a simple Lie algebra, and we also have seen that each vector x in a root space \mathfrak{g}_α comes with distinguished vectors y in $\mathfrak{g}_{-\alpha}$ and h in the Cartan subalgebra such that these three elements generate a copy of $sl(2; \mathbb{R})$. We now want to look at the representation theory. For this both roots and these $sl(2; \mathbb{R})$ subalgebras are important.

6.4. Finite-dimensional irreducible representations of $sl(2; \mathbb{R})$. Recall the commutation relations of $\mathfrak{g} = sl(2; \mathbb{R})$

$$[h, e] = 2e, \quad [h, f] = -2f, \quad [e, f] = h.$$

Recall that h spans the Cartan subalgebra, that e spans the root space \mathfrak{g}_α and f the one of $-\alpha$, where $\alpha(h) = 2$. Let $\rho : \mathfrak{g} \rightarrow \text{End}(V)$ be a finite-dimensional representation of \mathfrak{g} , then the action of h on V can be diagonalized

$$V = \bigoplus_{\lambda} V_{\lambda}$$

with

$$V_{\lambda} = \{v \in V \mid hv = \lambda(h)v\}.$$

The eigenvalues $\lambda(h)$ are described by a linear function λ on \mathfrak{h} . These linear functions are called *weights* and the eigenspaces V_{λ} are called *weight spaces*. The set of all λ such that

$V_\lambda \neq 0$ is called the set of weights of the representation V . We identify weights with complex numbers via $\lambda(h)$

Lemma 56. If v in V_λ , then $\rho(e)v$ in $V_{\lambda+2}$ and $\rho(f)v$ in $V_{\lambda-2}$.

Proof.

$$\rho(h)\rho(e)v = \rho([h, e])v + \rho(e)\rho(h)v = 2\rho(e)v + \lambda\rho(e)v = (\lambda + 2)\rho(e)v$$

and similarly for $\rho(f)$. ■

So for a finite-dimensional representation V , there must exist a weight λ such that $V_{\lambda+2} = 0$. Such a weight is called a *highest-weight*.

Lemma 57. Let $\rho : \mathfrak{g} \rightarrow \text{End}(V)$ be a finite-dimensional irreducible representation of \mathfrak{g} and let v_0 in V_λ be a highest-weight vector. Define

$$v_i = \frac{1}{i!} \rho(f)^i v_0$$

and $v_{-1} = 0$ then

- (1) $\rho(h)v_i = (\lambda - 2i)v_i$
- (2) $\rho(f)v_i = (i + 1)v_{i+1}$
- (3) $\rho(e)v_i = (\lambda - i + 1)v_{i-1}$

Proof. The first statement follows by iterative application of the previous lemma, while the second one is the definition of the vectors v_i . The last statement is proven by induction. The case $i = 0$ is true since $v_{-1} = 0$. The induction step follows by the following computation

$$\begin{aligned} i\rho(e)v_i &= \rho(e)\rho(f)v_{i-1} \\ &= \rho([e, f])v_{i-1} + \rho(f)\rho(e)v_{i-1} \\ &= \rho(h)v_{i-1} + \rho(f)\rho(e)v_{i-1} \\ &= (\lambda - 2(i-1))v_{i-1} + (\lambda - i + 2)\rho(f)v_{i-2} \\ &= (\lambda - 2(i-1))v_{i-1} + (i-1)(\lambda - i + 2)v_{i-1} \\ &= i(\lambda - i + 1)v_{i-1} \end{aligned}$$

dividing both sides by i finishes the proof. ■

The first point of this lemma tells us that all v_i are linearly independent as they have different $\rho(h)$ eigenvalues. Since V is finite-dimensional, there must be a m with $v_m \neq 0$ but $v_{m+1} = 0$. Hence all v_{m+i} must vanish too. Above lemma shows that v_0, \dots, v_m form a basis of a submodule of ρ and since the representation is irreducible they form a basis of V . The matrix $\rho(h)$ is a diagonal matrix, while $\rho(e)$ is upper triangular and $\rho(f)$ is lower triangular. The third statement of above lemma tells us something very interesting. For $i = m + 1$ the left-hand side is zero, but for the right-hand side this can only be true if $\lambda = m$. In other words, the highest weight is a non-negative integer m and the dimension of the representation is $m + 1$. We summarize

Theorem 58. Let $\rho : \mathfrak{g} \rightarrow \text{End}(V)$ be a $m + 1$ -dimensional irreducible representation of \mathfrak{g} , then the weight space decomposition of V is

$$V = V_{-m} \oplus V_{-m+2} \oplus \dots \oplus V_{m-2} \oplus V_m$$

the representation has a highest-weight vector of highest-weight m , so every $m + 1$ -dimensional irreducible representation of $sl(2; \mathbb{R})$ is isomorphic to the irreducible highest-weight representation of highest-weight m .

We remark, that when we studied functions on the Lie group $SU(2)$, we have introduced representations of $sl(2; \mathbb{R})$ in a *natural* way. Namely the Lie algebra acts on the polynomial ring in two variables in a degree preserving way. But the vector space of polynomials of degree m is $m + 1$ -dimensional and one can show that it carries the highest-weight representation of highest-weight m . Such natural representations also exist for other Lie algebras and for example $sl(3; \mathbb{R})$ acts naturally on homogeneous polynomials in three-variables.

6.5. Representation theory. We have seen, that $sl(2; \mathbb{R})$ and also $sl(3; \mathbb{R})$ allow a triangular decomposition

$$\mathfrak{g} = \mathfrak{g}_- \oplus \mathfrak{h} \oplus \mathfrak{g}_+,$$

where \mathfrak{g}_- is spanned by the creation operators, the f_α and \mathfrak{g}_+ are the annihilation operators, the e_α . This picture can be generalized. Let \mathfrak{g} be a simple finite-dimensional Lie algebra. A subset Π of the root system Δ is called *simple* if Π is a basis of \mathfrak{h}^* and if each root can be written as

$$\beta = \sum_{\alpha \in \Pi} k_\alpha \alpha$$

with integral coefficients k_α , either all non-positive or all non-negative. A root is then called positive if all coefficients are non-negative and negative if all coefficients are non-positive. The space of all positive respectively negative roots is denoted by Δ_\pm , and we have

$$\mathfrak{g} = \mathfrak{g}_- \oplus \mathfrak{h} \oplus \mathfrak{g}_+,$$

with

$$\mathfrak{g}_\pm = \bigoplus_{\alpha \in \Delta_\pm} \mathfrak{g}_\alpha.$$

We want to think of the elements in \mathfrak{g}_- as creation operators and of those in \mathfrak{g}_+ as annihilation operators.

Example 17. For $\mathfrak{g} = sl(3; \mathbb{R})$ the simple roots are α_1 and α_2 . The six roots of \mathfrak{g} are

$$\pm\alpha_1, \pm\alpha_2, \pm(\alpha_1 + \alpha_2)$$

so that $\alpha_1, \alpha_2, (\alpha_1 + \alpha_2)$ are positive roots and $-\alpha_1, -\alpha_2, -(\alpha_1 + \alpha_2)$ are negative ones.

Let $\rho : \mathfrak{g} \rightarrow \text{End}(V)$ be an irreducible finite-dimensional representation of \mathfrak{g} , then since abstract and usual Jordan decomposition agree the Cartan subalgebra \mathfrak{h} must act diagonalizable on V . Hence V decomposes into eigenspaces

$$V = \bigoplus_{\lambda} V_\lambda,$$

where the direct sum is over λ in \mathfrak{h}^* , and if $V_\lambda \neq 0$ then we call λ a *weight* and V_λ a *weight space*. We say that a weight λ is larger than another weight μ if

$$\lambda - \mu = \sum_{\alpha \in \Pi} k_\alpha \alpha$$

and all $k_\alpha \leq 0$. Of course there are weights that are neither larger nor smaller, a comparison is not always possible. We say that a vector v_λ in V_λ is a *highest-weight state* of *highest-weight* λ if $\mu < \lambda$ for all weights μ of the representation ρ .

Theorem 59. Every irreducible \mathfrak{g} module has a unique highest-weight state, it is a highest-weight representation.

This means that

$$\begin{aligned} \rho(e_\alpha)v_\lambda &= 0, & \text{for all } e_\alpha \text{ in } \mathfrak{g}_+ \\ \rho(h)v_\lambda &= \lambda(h)v_\lambda. \end{aligned} \tag{6.2}$$

and all other states of the representation can be written as a linear combinations of monomials of the form

$$\prod_{\alpha \in \Delta_-} \rho(f_\alpha)^{n_\alpha} v_\lambda$$

with non-negative integers n_α and f_α an element of \mathfrak{g}_α . Such a vector has weight

$$\mu = \lambda - \sum_{\alpha \in \Delta_+} n_{-\alpha} \alpha$$

and hence clearly obeys the relation $\mu < \lambda$.

6.6. Highest-weight representations of $sl(3; \mathbb{R})$. Let us conclude this section with a detailed description of some irreducible highest-weight representations of $sl(3; \mathbb{R})$. Let $\mathbb{R}[x, y, z]$ be the ring of polynomials over the real numbers in three variables. Define a linear map via

$$\begin{aligned} \rho(h_{\alpha_1}) &= x \frac{d}{dx} - y \frac{d}{dy}, & \rho(h_{\alpha_2}) &= y \frac{d}{dy} - z \frac{d}{dz}, \\ \rho(e_{\alpha_1}) &= x \frac{d}{dy}, & \rho(e_{\alpha_2}) &= y \frac{d}{dz}, & \rho(e_{\alpha_1 + \alpha_2}) &= x \frac{d}{dz}, \\ \rho(f_{\alpha_1}) &= y \frac{d}{dx}, & \rho(f_{\alpha_2}) &= z \frac{d}{dy}, & \rho(f_{\alpha_1 + \alpha_2}) &= z \frac{d}{dx}, \end{aligned} \tag{6.3}$$

then it is a computation that ρ defines a Lie algebra homomorphism, for example

$$[\rho(e_{\alpha_1}), \rho(e_{\alpha_2})] = x \frac{d}{dy} y \frac{d}{dz} - y \frac{d}{dz} x \frac{d}{dy} = x \frac{d}{dz} = \rho(e_{\alpha_1 + \alpha_2}) = \rho([e_{\alpha_1}, e_{\alpha_2}]).$$

We thus get an action on the space of homogeneous polynomials. Let us analyze those degree by degree. Degree zero polynomials are just the real numbers and they carry the one-dimensional representation of \mathfrak{g} . Degree one polynomials form \mathbb{R}^3 , and we get the three-dimensional defining representation spanned by x, y and z . What are the weights of these states? For example for x , we get

$$\rho(h_{\alpha_1})x = x, \quad \rho(h_{\alpha_2})x = 0,$$

let $\lambda = a\alpha_1 + b\alpha_2$, then $\lambda(h_{\alpha_1}) = 2a - b$ and $\lambda(h_{\alpha_2}) = 2b - a$, so that the weight of x is

$$\lambda_x = \frac{1}{3}(2\alpha_1 + \alpha_2).$$

Similarly one finds that the weight of y is

$$\lambda_y = \frac{1}{3}(-\alpha_1 + \alpha_2),$$

and

$$\lambda_z = \frac{1}{3}(-\alpha_1 - \alpha_2).$$

We observe that

$$\lambda_y = \lambda_x - \alpha_1, \quad \lambda_z = \lambda_x - \alpha_1 - \alpha_2$$

so that x is our highest-weight-vector of highest-weight $\lambda = \lambda_x$. This is in perfect agreement with the action of the creation operators

$$\rho(f_{\alpha_1})x = y \frac{d}{dx}x = y, \quad \rho(f_{\alpha_2})y = z \frac{d}{dy}y = z.$$

Let us now look at a monomial of degree n , for example

$$x^a y^b z^c$$

with non-negative integers a, b, c satisfying $a + b + c = n$. Then by induction and by the Leibniz rule of the derivative one can compute that the weight of this monomial is

$$a\lambda_x + b\lambda_y + c\lambda_z = n\lambda_x - ((n-a)\lambda_x + b(\alpha_1 - \lambda_x) + c(\alpha_1 + \alpha_2 - \lambda_x)) = n\lambda_x - (c\alpha_2 + (b+c)\alpha_1).$$

In other words, $n\lambda_x$ is a highest-weight with highest-weight vector x^n . The monomial can then be rewritten as

$$x^a y^b z^c = \frac{a!}{n!} \left(y \frac{d}{dx} \right)^b \left(z \frac{d}{dx} \right)^c x^n$$

so that

$$x^a y^b z^c = \frac{a!}{n!} (\rho(f_{\alpha_1}))^b (\rho(f_{\alpha_1 + \alpha_2}))^c x^n.$$

So that we can summarize

Theorem 60. The polynomial ring in three variables decomposes as a direct sum

$$\mathbb{R}[x, y, z] = \bigoplus_{n=0}^{\infty} V_{n\lambda}$$

of irreducible highest-weight representations $V_{n\lambda}$ of $\mathfrak{g} = \mathfrak{sl}(3; \mathbb{R})$ of highest-weight

$$n\lambda = \frac{n}{3}(2\alpha_1 + \alpha_2)$$

Representations of this type appear frequently in physics, and they are termed oscillator representations, where you should think of $\frac{d}{dx}$ as the annihilation operator a_x and as x as the creation operator a_x^\dagger .

More generally this result (with appropriate ordering of simple roots) generalizes as follows

Theorem 61. The polynomial ring in m -variables ($m > 1$) decomposes as a direct sum

$$\mathbb{R}[x_1, \dots, x_m] = \bigoplus_{n=0}^{\infty} V_{n\lambda}$$

of irreducible highest-weight representations $V_{n\lambda}$ of $\mathfrak{g} = \mathfrak{sl}(m; \mathbb{R})$ of highest-weight

$$n\lambda = \frac{n}{m}((m-1)\alpha_1 + (m-2)\alpha_2 + \dots + \alpha_{m-1})$$

6.7. Exercises.

(1) Show that the cross product of vectors

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_2 y_3 - y_2 x_3 \\ x_3 y_1 - y_3 x_1 \\ x_1 y_2 - y_1 x_2 \end{pmatrix}$$

defines a Lie algebra structure on \mathbb{R}^3 .

(2) Determine all non-abelian Lie algebras \mathfrak{g} (up to isomorphism) of dimension two.

- (3) Compute the Casimir of $sl(2; \mathbb{R})$ in the adjoint representation without using Corollary 49.
- (4) Compute the Casimir of $sl(3; \mathbb{R})$ in the three-dimensional representation without using Corollary 49.

6.8. Solutions.

- (1) we choose as a basis of \mathbb{R}^3 the standard unit vectors

$$\sigma_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The cross-product is by definition anti-symmetric. Further, we compute

$$\sigma_1 \times \sigma_2 = \sigma_3, \quad \sigma_2 \times \sigma_3 = \sigma_1, \quad \sigma_3 \times \sigma_1 = \sigma_2.$$

These are exactly the commutation relations of the Pauli-matrices, which can be compactly written as

$$[\sigma_i, \sigma_j] = \varepsilon_{ijk} \sigma_k$$

with the totally anti-symmetric tensor ε . These define the Lie algebra $su(2)$, see(4.5).

- (2) Let \mathfrak{g} be a two-dimensional Lie algebra, and let x, y be a basis of \mathfrak{g} . Then the most general form of the possible commutation relations are

$$[x, x] = [y, y] = 0, \quad [x, y] = ax + by$$

for some constants a and b . If both these numbers are zero, then the Lie algebra is abelian. Otherwise we can assume that $a \neq 0$ (if $a = 0$, then $b \neq 0$ and we can rename). The aim is to find another basis, such that the commutation relations look very nice. Let $y' = \frac{y}{a}$, then

$$[x, y'] = x + by'$$

Let $x' = x + by'$, then

$$[x', y'] = x'$$

We thus have shown, that every non-abelian two-dimensional Lie algebra has a basis x', y' such that the commutation relations take the simple form as above. Note, that this Lie algebra is a Lie algebra of two by two matrices, and it can be viewed as a subalgebra of $sl(2, \mathbb{R})$ generated by f and $\frac{1}{2}h$.

- (3) The strategy to compute a Casimir is as follows:

- Choose an ordered basis x_1, \dots, x_n of \mathfrak{g}
- Find the corresponding representation matrices $\rho(x_i)$.
- Compute the traces $\text{tr}(\rho(x_i)\rho(x_j))$ and use them to find a dual basis y_1, \dots, y_n defined by $\text{tr}(\rho(x_i)\rho(y_j)) = \delta_{i,j}$.
- The Casimir is the matrix

$$C_\rho = \sum_{i=1}^n \rho(x_i)\rho(y_i).$$

So let's do that. As a basis, we choose $x_1 = e, x_2 = h, x_3 = f$. Then the representation matrix for $\text{ad}(e)$ is defined using

$$\text{ad}(e)(e) = [e, e] = 0, \quad \text{ad}(e)(h) = [e, h] = -2e, \quad \text{ad}(e)(f) = [e, f] = h,$$

so that in our ordered basis we have

$$\operatorname{ad}(e) = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Similarly one computes

$$\operatorname{ad}(h) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad \operatorname{ad}(f) = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

The non-zero traces are then

$$\operatorname{tr}(\operatorname{ad}(e)\operatorname{ad}(f)) = \operatorname{tr}(\operatorname{ad}(f)\operatorname{ad}(e)) = 4, \quad \operatorname{tr}(\operatorname{ad}(h)\operatorname{ad}(h)) = 8,$$

so that a dual basis is $y_1 = \frac{1}{4}f, y_2 = \frac{1}{8}h, y_3 = \frac{1}{4}e$, and the Casimir is

$$\begin{aligned} C_{\operatorname{ad}} &= \frac{1}{4}\operatorname{ad}(e)\operatorname{ad}(f) + \frac{1}{8}\operatorname{ad}(h)\operatorname{ad}(h) + \frac{1}{4}\operatorname{ad}(f)\operatorname{ad}(e) \\ &= \frac{1}{4} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{8} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

- (4) We already know the representation matrices. As we naturally identify $sl(3; \mathbb{R})$ with its three-dimensional matrix representation, I will omit the ρ indicating the representation. We choose our basis to be $x_1 = e_{\alpha_1}, x_2 = e_{\alpha_2}, x_3 = e_{\alpha_1+\alpha_2}, x_4 = h_{\alpha_1}, x_5 = h_{\alpha_2}, x_6 = f_{\alpha_1}, x_7 = f_{\alpha_2}, x_8 = f_{\alpha_1+\alpha_2}$. We know from the lecture, that the trace can only be non-zero for vectors in root spaces \mathfrak{g}_{α} and \mathfrak{g}_{β} with $\alpha + \beta = 0$. So that there are only few possibilities for non-zero traces. They then can be compactly written as

$$\operatorname{tr}(e_{\alpha}, f_{\beta}) = \delta_{\alpha, \beta}, \quad \operatorname{tr}(h_{\alpha_1}, h_{\alpha_1}) = \operatorname{tr}(h_{\alpha_2}, h_{\alpha_2}) = 2, \quad \operatorname{tr}(h_{\alpha_1}, h_{\alpha_2}) = -1.$$

So that the dual vectors $y_1 = f_{\alpha_1}, y_2 = f_{\alpha_2}, y_3 = f_{\alpha_1+\alpha_2}, y_6 = e_{\alpha_1}, y_7 = e_{\alpha_2}, y_8 = e_{\alpha_1+\alpha_2}$ are obvious. It remains to find y_4 and y_5 . We make the Ansatz $y_4 = ah_{\alpha_1} + bh_{\alpha_2}$, and we are looking for real numbers a and b such that

$$\operatorname{tr}(h_{\alpha_1}y_4) = 1, \quad \operatorname{tr}(h_{\alpha_2}y_4) = 0.$$

But

$$\operatorname{tr}(h_{\alpha_1}y_4) = 2a - b, \quad \operatorname{tr}(h_{\alpha_2}y_4) = 2b - a$$

so that the second equation tells us that $a = 2b$, while inserting this in the first one gives $3b = 1$. Hence

$$y_4 = \frac{1}{3}(2h_{\alpha_1} + h_{\alpha_2}).$$

Similarly, we find

$$y_5 = \frac{1}{3}(h_{\alpha_1} + 2h_{\alpha_2}).$$

So that the Casimir becomes

$$\begin{aligned}
 C &= \sum_{i=1}^8 x_i y_i \\
 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \\
 &\quad \frac{1}{3} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \frac{8}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
 \end{aligned}$$

7. THE BOSONIC STRING

In this final section, we illustrate features of what we have learnt in the example of the bosonic string. String theory is a quantum field theory of strings, that is of one-dimensional objects, instead of point-like particles. The original motivation of string theory was to find a quantum field theory that consistently incorporates gravity. As one-dimensional objects have more structure than points, string theory (if formulated correctly) is guaranteed to refine quantum theory of point-like objects. The original string theory development started with purely bosonic strings. By now it is realized that this is not sufficient and supersymmetry has to be incorporated. Nonetheless, the bosonic string by itself has many nice features and it will be our interesting example to terminate this course.

There are both many science and popular science books on string theory. A very enjoyable popular science book (to me) is the book the elegant universe by Brian Greene. We won't really follow any textbook, but [P] is a suitable reference.

7.1. The free boson compactified on a circle. A string propagating in time sweeps out a two-dimensional surface in space-time. This surface is called the world-sheet of the string. The world-sheet quantum field theory is a two-dimensional conformal quantum field theory (CFT), so that these two-dimensional CFTs are the building blocks of string theory. CFTs are intimately connected to Lie groups and Lie algebras. In some sense the symmetry algebra of a CFT is a quantization of a Lie algebra and its representation theory a quantization of harmonic analysis on the corresponding Lie group. We learnt that the Circle S^1 is the one-dimensional abelian Lie group $U(1)$ with Lie algebra $u(1)$ represented by $\frac{d}{dx}$. It acts on periodic functions $f: \mathbb{R} \rightarrow \mathbb{R}$. Periodicity means that $f(x+2\pi) = f(x)$ and a basis of functions has been given by $f_n(x) = e^{inx}$. Multiplication of functions is then easily performed

$$f_n(x)f_m(x) = f_{n+m}(x)$$

and seen to be the same as addition in the integers. The inner product of functions is (we take an appropriate normalization of the measure)

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f_n(x)f_m(x)dx = \delta_{n+m,0}.$$

The free boson compactified on a circle is a quantization of this data. Let us denote the generator of $u(1)$ by u , then we consider its loop algebra

$$u(1) \otimes \mathbb{R}[t, t^{-1}]$$

of formal Laurent polynomials in one variable with coefficients in $u(1)$. This is an infinite-dimensional abelian Lie algebra. Abelian Lie algebras are not too interesting. However, this Lie algebra allows for a central extension

$$\hat{u}(1) = u(1) \otimes \mathbb{R}[t, t^{-1}] \oplus \mathbb{R}K \oplus \mathbb{R}d$$

with non-zero commutation relations

$$[u_n, u_m] = n\delta_{n+m,0}K, \quad [d, u_n] = nu_n$$

with the short-hand notation $u_n = u \otimes t^n$. Note, that d acts as $t \frac{d}{dt}$. It acts as a differential operator and it is actually called a derivation. The Lie algebra $\hat{u}(1)$ is called the affinization of $u(1)$, or the Kac-Moody algebra of $u(1)$. It is the symmetry algebra of the uncompactified free boson CFT. What does this mean? A quantum field theory is in the first place a theory of fields. So what are these fields? The Heisenberg field or free boson is the formal Laurent polynomial

$$X(z) = \sum_{n \in \mathbb{Z}} u_n z^{-n-1}$$

with coefficients in the symmetry algebra $\hat{u}(1)$. It acts naturally on representations of the symmetry algebra. The most important representation is the vacuum representation V_0 . It is generated by a highest-weight state $|0\rangle$, the vacuum. The vacuum satisfies

$$u_0|0\rangle = 0, \quad u_n|0\rangle = 0, \quad K|0\rangle = |0\rangle$$

for all $n > 0$, while the u_n for negative n create the descendent states of the infinite-dimensional vacuum representation. Its dual $\langle 0|$ is defined by interchanging the role of creation and annihilation operators. We define a norm by $\langle 0|0\rangle = 1$. The algebraic structure of the free boson is encoded in correlation functions

$$\begin{aligned} \langle 0|X(z)X(w)|0\rangle &= \langle 0| \sum_{n \in \mathbb{Z}} u_n z^{-n-1} \sum_{m \in \mathbb{Z}} u_m w^{-m-1} |0\rangle \\ &= \langle 0| \frac{1}{zw} \sum_{n=1}^{\infty} u_n z^{-n} \sum_{m=1}^{\infty} u_{-m} w^m |0\rangle \\ &= \langle 0| \frac{1}{zw} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} [u_n, u_{-m}] z^{-n} w^m |0\rangle \\ &= \langle 0| \frac{1}{zw} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} n\delta_{n,m} K |0\rangle \\ &= \frac{1}{zw} \sum_{n=1}^{\infty} n z^n w^{-n} \\ &= \frac{1}{(z-w)^2} \end{aligned}$$

In the last equation, we used the derivative of the geometric series

$$\frac{1}{(1 - \frac{z}{w})} = \sum_{n=0}^{\infty} z^n w^{-n}.$$

So that the identity is only true if $|z| < |w|$. Correlation functions are neatly summarized in the algebraic structure called operator product algebra. For the free boson, the operator product algebra is completely determined by

$$X(z)X(w) \sim \frac{1}{(z-w)^2}.$$

The free boson X should be viewed as a quantization of the Lie algebra generator $\frac{d}{dx}$. This operator acts naturally on the square integrable functions on the circle, the basis of these functions are the e^{inx} , and $\frac{d}{dx}$ acts as

$$\frac{d}{dx}e^{inx} = ine^{inx}.$$

We now would like to find fields that quantize this action. For this define the field $\phi(z)$ via

$$\frac{d}{dz}\phi(z) = X(z)$$

and define the normally ordered exponential

$$V_\alpha(z) =: e^{\alpha\phi(z)} :$$

(the factor of i is hidden in the definition of ϕ). In quantum field theory, due to non-commutativity, there is an issue in how to write products of operators. Normally ordering means that all annihilation operators (the u_n with $n > 0$) are to the left of the other operators. This procedure ensures that the action of fields on states is well-defined. The operator product of $X(z)$ with the field V_α resembles very much the action of the Lie algebra $u(1)$ on its square integrable functions, namely

$$X(x)V_\alpha(w) \sim \frac{\alpha V_\alpha(w)}{(z-w)}.$$

Moreover, the product of functions translates to the operator product of fields as

$$V_\alpha(z)V_\beta(w) =: V_\alpha(z)V_\beta(w) : (z-w)^{\alpha\beta} \sim (z-w)^{\alpha\beta} (V_{\alpha+\beta}(w) + \dots).$$

Here, we see that the operator product is multi-valued (we have to decide for a choice of root) if $\alpha\beta$ is not an integer. So that if we restrict to square root of two times integer α we get a well-defined operator product algebra. The integers however are an example of a lattice and the conformal field theory generated by X and the V_n for integer n is called the free boson compactified on $\mathbb{R}/\mathbb{Z}\sqrt{2} = S^1$, or just the lattice CFT of the lattice $\sqrt{2}\mathbb{Z}$.

The interaction of physical quantities in quantum field theory is given by expectation values for interaction, these are correlation functions. Correlation functions should be viewed as the quantum analog of the inner product. For the free boson, two-point functions are

$$\langle V_\alpha(z)V_\beta(w) \rangle = \delta_{\alpha+\beta,0}(z-w)^{-\alpha^2}$$

and three-point functions are

$$\langle V_\alpha(z)V_\beta(w)V_\gamma(x) \rangle = \delta_{\alpha+\beta+\gamma,0}(z-w)^{-\alpha^2}.$$

You can go on with this and for n -field insertions you get

$$\langle V_{\alpha_1}(z_1) \dots V_{\alpha_n}(z_n) \rangle = \delta_{\alpha_1+\dots+\alpha_n,0} \prod_{i<j} (z_i - z_j)^{\alpha_i\alpha_j}.$$

Note, that the corresponding quantum mechanical correlation function is

$$\langle e^{i\alpha_1 x} \dots e^{i\alpha_n x} \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\alpha_1 x} \dots e^{i\alpha_n x} dx = \delta_{\alpha_1 + \dots + \alpha_n, 0}.$$

Recall that the Laplacian on the circle is just the second derivative $\Delta = \frac{d}{dx} \frac{d}{dx}$, and also recall that our free boson analogue of the ordinary derivative is the free boson $X(z)$ itself, while its normally ordered product is (up to a factor of $1/2$ the Virasoro field). The Laplacian was something like the Hamiltonian, and the Virasoro field is actually also called the energy-momentum field of the CFT. In general, there is a nice analogy between harmonic analysis on a Lie group and a CFT on a Lie group.

Harmonic analysis on a Lie group is called the semi-classical (or quantum mechanical) limit of two-dimensional CFT on the same Lie group. In this limit

- fields of the CFT become square integrable functions on the Lie group.
- the product of fields becomes ordinary multiplication of functions.
- correlation functions become inner products.
- the infinite-dimensional symmetry Lie algebra degenerates to the Lie algebra of invariant vector fields.
- the Virasoro field (the energy-momentum tensor) becomes the Laplacian.

In other words, one application of what you have learnt in this course is, you have learnt the quantum mechanics limit of a string propagating on a Lie group.

Before we turn to more general lattice CFTs, we have to discuss the most important structure of every world-sheet CFT of a string. Its Virasoro Lie algebra.

7.2. The Virasoro algebra. The Virasoro algebra is a central extension of the Lie algebra of continuous derivations on Laurent series in one variable t . The Lie algebra of continuous derivations is also called the Witt algebra and it is generated by $-t^{n+1} \frac{d}{dt}$ with commutation relations

$$\left[-t^{n+1} \frac{d}{dt}, -t^{m+1} \frac{d}{dt} \right] = -(n-m)t^{n+m+1} \frac{d}{dt}$$

The Virasoro algebra has generators L_n for n in \mathbb{Z} plus a central element C . The commutation relations are almost as those of the Witt algebra (up to setting C to zero),

$$[L_n, L_m] = (n-m)L_{n+m} + \frac{n^3 - n}{12} \delta_{n+m, 0} C.$$

The central element C acts in representations of CFT as multiplication by a number c . This number is called the central charge c , and the quotient of the Virasoro algebra by the ideal generated by the relation $C = c$ is called the Virasoro algebra of central charge c . The importance of this algebra is that it appears as the symmetry algebra of every two-dimensional CFT. In other words, every two-dimensional CFT and hence every world-sheet theory of a string has an infinite-dimensional Lie algebra as symmetry. This is a very helpful and powerful structure.

How do we find this structure in the free boson CFT? We define the Virasoro field

$$T(z) = \frac{1}{2} : X(z)X(z) := \frac{1}{2} \sum_{n \in \mathbb{Z}} z^{-n-2} \sum_{m=0}^{\infty} (u_{n-m} u_m + u_{-m-1} u_{n+m+1}) = \sum_{n \in \mathbb{Z}} L_n z^{-n-2}.$$

It is an instructive and laborious computation to verify that

$$[L_n, L_m] = (n-m)L_{n+m} + \frac{n^3 - n}{12} \delta_{n+m, 0},$$

we indeed get the Virasoro algebra relations at central charge one. This laborious computation can be circumvented knowing that the operator product algebra of the central charge c Virasoro algebra is

$$T(z)T(w) \sim \frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\frac{d}{dw}T(w)}{(z-w)}$$

and verifying that the operator product of $\frac{1}{2} : X(z)X(z) :$ indeed is the same for $c = 1$. The power of the operator product is that the complete algebraic structure of the infinite-dimensional symmetry algebra of a CFT is encoded in its generating fields. In the case of the Virasoro algebra this means it is all encoded in this innocent looking operator product. In order to get to the bosonic string, we have to replace the integers by other lattices.

7.3. Lattice CFT. The first example of a bosonic string is given by starting with a CFT of a self-dual even rank 24 lattice. What does this mean. A lattice L of rank n is a free \mathbb{Z} -module of dimension n . This means there are n generators x_1, \dots, x_n , such that the set

$$L = \{m_1x_1 + \dots + m_nx_n \mid m_i \in \mathbb{Z}\}$$

is the lattice L . It is closed under addition and it carries an action of \mathbb{Z} via multiplication. Further, we require that this set has a quadratic form

$$Q : L \times L \rightarrow \mathbb{Q}.$$

We call a lattice integral if this quadratic form takes values in the integers and even if it takes values in the even integers. Even integral lattices lift to conformal field theories very much like the free boson compactified via the lattice $\sqrt{2}\mathbb{Z}$. Namely to each element γ in L , we can associate a vertex operator V_γ , such that the operator product algebra is defined by lattice addition. Namely

$$V_\gamma(z)V_\mu(w) =: V_\gamma(z)V_\mu(w) : (z-w)^{Q(\gamma,\mu)} \sim (z-w)^{Q(\gamma,\mu)} (V_{\gamma+\mu}(w) + \dots).$$

We see as in the free boson that integrality ensures that we don't have to take any roots, so there is no multi-valuedness. Modules of a lattice theory are parameterized by elements of the coset L'/L , where L' is the dual lattice with respect to Q . so that if a lattice is self-dual, then the only module of the lattice theory is the lattice theory itself. In that case, we are talking about the lattice CFT of the self-dual lattice. If the quadratic form takes only values in the even integers, then we ensure that it is a bosonic theory. Otherwise there will also be fermions. Geometrically a lattice CFT describes a string propagation on the n -dimensional torus \mathbb{R}^n/L . A lattice CFT of rank n , that is its complexification has dimension n as a complex vector space has n copies of the Heisenberg algebra, the free boson, as a subalgebra. Its Virasoro field is the sum of the Virasoro fields of these subalgebras.

Example 18. There are two-important examples for us

- (1) The two-dimensional lattice $L_{1,1} = \mathbb{Z}x \oplus \mathbb{Z}y$ with quadratic form $(x,y) = 1$ and $(x,x) = (y,y) = 0$.
- (2) A self-dual even lattice Λ of rank 24. These lattices are called Niemeier lattices and the most prominent example is the Leech lattice.

You might have heard that the bosonic string is only consistent in 26 dimensions. This statement is actually false, it must (probably) say in dimension at most 26. We will come to that. The full world-sheet CFT is not the bosonic string, but its states are described by a semi-infinite cohomology, called BRST-cohomology.

7.4. Fermionic Ghosts. The next ingredient we need is a fermionic CFT. This is a theory based on a Lie superalgebra. The Lie superalgebra is generated by fermionic elements c_n, b_n for n in \mathbb{Z} and only one bosonic element K . The non-zero commutation relations are

$$[b_n, c_m] = [c_m, b_n] = \delta_{n,m} K.$$

These look very similar to those of $\hat{u}(1)$. The associated fields are called the bc -ghosts, and they are

$$b(z) = \sum_{n \in \mathbb{Z}} b_n z^{-n-1}, \quad c(z) = \sum_{n \in \mathbb{Z}} c_n z^{-n}.$$

The Lie superalgebra structure is encoded in the operator product

$$b(z)c(w) \sim c(z)b(w) \sim \frac{1}{(z-w)}.$$

This CFT has a Virasoro field

$$T_{\text{ghost}}(z) =: \left(\frac{d}{dz} b(z) \right) c(z) : - 2 \frac{d}{dz} : (b(z)c(z)) :$$

of central charge -26 . This central charge is the reason people are claiming that the bosonic string is only consistent in 26 dimensions. Why are these ghosts so important?

7.5. BRST quantization of the bosonic string. Consider a CFT, that has as commuting subalgebras the fermionic ghosts and a Virasoro algebra of central charge c . Then one can define the BRST-current

$$J_{\text{BRST}}(z) =: c(z)T(z) : + \frac{1}{2} : c(z)T_{\text{ghost}}(z) + \frac{3}{2} \frac{d^2}{dz^2} c(z) :$$

This is a fermionic field. It satisfies the following important operator products

$$\begin{aligned} J_{\text{BRST}}(z)b(w) &\sim \frac{3}{(z-w)^3} + \frac{:b(w)c(w):}{(z-w)^2} + \frac{T(w) + T_{\text{ghost}}(w)}{(z-w)}, \\ J_{\text{BRST}}(z)J_{\text{BRST}}(w) &\sim \frac{(c/2 - 9) : c(w) \frac{d}{dw} c(w) :}{(z-w)^3} - \\ &\quad \frac{(c/2 - 9)/2 : c(w) \frac{d^2}{dw^2} c(w) :}{(z-w)^2} - \frac{(c - 26)/12 : c(w) \frac{d^3}{dw^3} c(w) :}{(z-w)}. \end{aligned}$$

Here it is very important that the first order pole of the operator product of the BRST-current with itself vanishes if and only if $c = 26$. Define the BRST-differential

$$Q_{\text{BRST}} = \frac{1}{2\pi i} \oint J_{\text{BRST}}(z) dz$$

such that it picks up the residuum of the operator product of a field with the BRST-current. We thus see, that

$$Q_{\text{BRST}} b(w) = T(w) + T_{\text{ghost}}(w), \quad Q_{\text{BRST}} J_{\text{BRST}}(w) = (c - 26)/12 : c(w) \frac{d^3}{dw^3} c(w) :$$

especially if and only if $c = 26$, we have

$$Q_{\text{BRST}} J_{\text{BRST}}(w) = 0$$

and hence also

$$Q_{\text{BRST}} Q_{\text{BRST}} = 0.$$

This is very important. Let V be a vector space on which Q_{BRST} acts, that is we have a map

$$Q_{\text{BRST}} : V \rightarrow V$$

and since $Q_{\text{BRST}}^2 = 0$, the image of this map is contained in the kernel, so that we can look at the vector space of equivalence classes

$$\begin{aligned} \mathcal{H}_V &= \frac{\ker(Q_{\text{BRST}} : V \rightarrow V)}{\text{im}(Q_{\text{BRST}} : V \rightarrow V)} \\ &= \{[v] \mid v \in V, Q_{\text{BRST}}v = 0, [v] = [w] \leftrightarrow Q_{\text{BRST}}x = v - w \text{ for some } x \in V\} \end{aligned}$$

This set of equivalence classes is called the BRST-cohomology on V .

Definition 39. Let V be a two-dimensional conformal field theory of central charge 26 that is graded by $L_{1,1}$ and that contains a rational unitary CFT of central charge 24 as subalgebra, such that the Virasoro field of V is the sum of the Virasoro fields of the two subalgebras. Further let the only module of V be V itself. Let W be the product of this CFT with the fermionic ghost CFT and let X be the kernel of b_0 acting on the vacuum module of W . The space of physical states of a bosonic string propagating on a world-sheet described by V is the semi-infinite BRST-cohomology \mathcal{H}_X .

An example then would be if we take for V the product of the lattice CFT of a Niemeier lattice times the lattice CFT of the lattice $L_{1,1}$. In this case $V = W$. This describes the original bosonic string and it is a string propagating on a 26-dimensional torus. There are a few important theorems, most importantly the no-ghost theorem that tells us that the states of the bosonic string live in a Hilbert space [P].

Let us outline the mathematical importance of the bosonic string. Lian and Zuckerman showed that the states of the bosonic string have the structure of an infinite-dimensional Lie algebra. Richard Borcherds realized that these are actually very nice Lie algebras, and he called them generalized Kac-Moody algebras. Frenkel Lepowsky and Meurman were able to construct a CFT of central charge with automorphism group the largest finite simple sporadic group, the monster. The Lie algebra of the bosonic string had the favourable nature, that its graded dimensions were counted by an automorphic product. Borcherds was able to use the properties of this automorphic product to relate Hauptmoduls of genus zero subgroups of the modular group to twisted partition functions of the monster CFT. He received the fields medal for his work, that is the highest possible award for a mathematician. Terry Gannon, here of the University of Alberta is a leading expert on this subject and he wrote the standard textbook [G].

There is a generalization of this story, which my master advisor Nils Scheithauer has pursued. Namely there are very few other Lie algebras (less or equal than 12) that behave as the Lie algebras of the two known bosonic strings. They also connect to automorphic forms and finite sporadic groups. The conjecture is that these are all Lie algebras of physical states of bosonic strings. The conjecture is still open, and in my master thesis I took care of a family of four cases.

8. POSSIBLE EXAM QUESTIONS

Exercise 4. Let \mathbf{H} be the Lie algebra generated by p, q, z, d with commutation relations

$$[d, p] = p, \quad [d, q] = -q, \quad [p, q] = z$$

and all others vanish. Especially p, q, z generated the three-dimensional Heisenberg Lie algebra of the introduction and also of the harmonic oscillator in quantum mechanics.

- (1) A quadratic Casimir operator is a polynomial of exactly degree two in the generators of the Lie algebra with the property that it commutes with every element of the Lie algebra. This Lie algebra \mathbf{H} has two quadratic Casimir operators, one of them is $C_1 = z^2$. Find the other one.

Hint: Make the Ansatz $C_2 = \alpha dz + \beta(pq + qp)$ and determine α and β such that C_2 commutes with d, p, q and z .

- (2) Let

$$g(v, w, x, y) = e^{vp} e^{xd+yz} e^{wq},$$

the left Maurer-Cartan one form is defined as

$$\omega(g) = g^{-1} dg = g^{-1} \frac{d}{dv} g dv + g^{-1} \frac{d}{dw} g dw + g^{-1} \frac{d}{dx} g dx + g^{-1} \frac{d}{dy} g dy.$$

it takes values in \mathbf{H} , so it can be written as

$$\omega(g) = \omega(d)d + \omega(z)z + \omega(p)p + \omega(q)q$$

with the dual one-forms $\omega(X)$ for X in \mathbf{H} . Compute $\omega(g)$ and the right Maurer-Cartan one form $\omega(g^{-1})$. The left Haar measure is the wedge product of the dual one forms of the left Maurer-Cartan one form, and the right Haar measure correspondingly. Do these two measures agree?

The following formula might be helpful,

$$e^X Y e^{-X} = \sum_{n=0}^{\infty} \frac{(\text{ad}(X))^n(Y)}{n!} = Y + [X, Y] + \frac{1}{2}[X, [X, Y]] + \frac{1}{6}[X, [X, [X, Y]]] + \dots$$

- (3) $sl(3; \mathbb{R})$ contains a copy of $sl(2; \mathbb{R})$ as subalgebra generated by $e_{\alpha_1}, h_{\alpha_1}, f_{\alpha_1}$. So that $sl(3; \mathbb{R})$ can be viewed as a $sl(2; \mathbb{R})$ -module via the adjoint action. Decompose $sl(3; \mathbb{R})$ in a direct sum of irreducible $sl(2; \mathbb{R})$ -modules under this action.
- (4) $sl(3; \mathbb{R})$ contains a copy of $sl(2; \mathbb{R})$ as subalgebra generated by $e_{\alpha_1+\alpha_2}, h_{\alpha_1} + h_{\alpha_2}, f_{\alpha_1+\alpha_2}$. So that $sl(3; \mathbb{R})$ can be viewed as a $sl(2; \mathbb{R})$ -module via the adjoint action. Decompose $sl(3; \mathbb{R})$ in a direct sum of irreducible $sl(2; \mathbb{R})$ -modules under this action.
- (5) What is the highest-weight of the adjoint representation of $sl(3; \mathbb{R})$?
- (6) Consider the polynomial ring $\mathbb{R}[\eta, v]$ in two odd variables η, v . This means η and v satisfy $\eta^2 = v^2 = 0$ and $\eta v + v \eta = 0$. So that the polynomial ring has real basis $1, \eta, v, \eta v$. It is four dimensional. The derivatives $\frac{d}{dv}$ and $\frac{d}{d\eta}$ act on the basis vectors as follows

$$\frac{d}{dv} 1 = 0, \quad \frac{d}{dv} \eta = 0, \quad \frac{d}{dv} v = 1, \quad \frac{d}{dv} v \eta = \eta, \quad \frac{d}{dv} \eta v = -\eta,$$

and

$$\frac{d}{d\eta} 1 = 0, \quad \frac{d}{d\eta} \eta = 1, \quad \frac{d}{d\eta} v = 0, \quad \frac{d}{d\eta} v \eta = -v, \quad \frac{d}{d\eta} \eta v = v.$$

Further these operators satisfy themselves the rules

$$\frac{d}{d\eta} \frac{d}{d\eta} = \frac{d}{dv} \frac{d}{dv} = 0, \quad \frac{d}{d\eta} \frac{d}{dv} + \frac{d}{dv} \frac{d}{d\eta} = 0.$$

Show

- The map $\rho : sl(2; \mathbb{R}) \rightarrow \text{End}(\mathbb{R}[\eta, v])$ defined by

$$\rho(e) = \eta \frac{d}{dv}, \quad \rho(h) = \eta \frac{d}{d\eta} - v \frac{d}{dv}, \quad \rho(f) = v \frac{d}{d\eta}$$

defines a representation of $sl(2; \mathbb{R})$.

- Decompose $\mathbb{R}[\eta, \nu]$ into irreducible representations of $sl(2; \mathbb{R})$.

REFERENCES

- [B] V Bouchard. Ma Ph 451 – Mathematical Methods for Physics I. Winter 2013.
[Bu] D Bump. Lie Groups. Springer Graduate Texts in Mathematics.
[G] T. Gannon. Moonshine Beyond the Monster. Cambridge University Press.
[H] S Hassani. Mathematical Physics: A Modern Introduction to Its Foundations.
[Hu] J E Humphreys. Introduction to Lie Algebras and Representation Theory. Springer Graduate Texts in Mathematics.
[P] J. Polchinski. String Theory. Cambridge University Press.

(T Creutzig) 573 CAB, UNIVERSITY OF ALBERTA

E-mail address: creutzig@ualberta.ca