# The 2+1 Lorentz Group and Its Representations

Author: Klas Sjöstedt klsj1231@student.su.se

Supervisor: Prof. Dr. Ingemar Bengtsson

Department of Physics, Stockholm University, AlbaNova University Center, 106 91 Stockholm, Sweden

> A thesis submitted for the degree of Bachelor of Science

# Abstract

The Lorentz group is a symmetry group on Minkowski space, and as such is central to studying the geometry of this and related spaces. The group therefore shows up also from physical considerations, such as trying to formulate quantum physics in anti-de Sitter space. In this thesis, the Lorentz group in 2+1 dimensions and its representations are investigated, and comparisons are made to the analogous rotation group. Firstly, all unitary irreducible representations are found and classified. Then, those representations are realised as the square-integrable, analytic functions on the unit circle and the unit disk, which turn out to correspond to the projective lightcone and the hyperbolic plane, respectively. Also, a way to realise a particular class of representations on 1+1-dimensional anti-de Sitter space is shown.

Keywords: The Lorentz group, representation theory, anti-de Sitter space, relativity.

# Sammanfattning

Lorentzgruppen är en symmetrigrupp på Minkowski-rum, och är således central för att studera geometrin i detta och relaterade rum. Gruppen dyker också därför upp från fysikaliska frågeställningar, såsom att försöka formulera kvantfysik i anti-de Sitter-rum. Denna uppsats undersöker Lorentzgruppen i 2+1 dimensioner och dess representationer, och jämför med den analoga rotationsgruppen. Först konstrueras och klassificeras alla unitära irreducibla representationer. Sedan realiseras dessa representationer som de analytiska funktioner på enhetscirkeln och enhetsskivan vars belopp i kvadrat är integrerbara. Det visar sig att denna cirkel respektive skiva svarar mot den projektiva ljuskonen respektive det hyperboliska planet. Dessutom visas att en särskild klass av representationer blir relevanta för att formulera kvantfysik i 1+1-dimensionellt anti-de Sitter-rum.

Nyckelord: Lorentzgruppen, representationsteori, anti-de Sitter-rum, relativitet.

# Contents

A	bstract	<b>2</b>
Sa	ammanfattning	2         3         Representations       5         the Lorentz Group       5         s       5         isomorphism       8         d Irreducibility       9         tiple Coverings       10         ation and Lorentz Groups       11         Both Groups       11         f the Rotation Group and the Lorentz Group       14         ns for the Rotation Group       16         ns for the Lorentz Group       17         of the Unitary Representations       19         the Projective Lightcone       19         Spaces       21
1	Introduction         1.1 Thesis Structure	<b>3</b> 4
2	A Few Groups and Group Representations2.1The Rotation Group and the Lorentz Group2.2Some More Matrix Groups2.3Establishing an Important Isomorphism2.4Group Representations and Irreducibility2.5Covering Groups and Multiple Coverings	<b>5</b> 5 8 9 10
3	Representations of the Rotation and Lorentz Groups         3.1       Linear Representations of Both Groups         3.2       Unitary Representations of the Rotation Group and the Lorentz Group         3.2.1       Unitarity Conditions for the Rotation Group         3.2.2       Unitarity Conditions for the Lorentz Group	<b>11</b> 11 14 16 17
4	Some Physical Realisations of the Unitary Representations4.1The Hyperbolic Plane and the Projective Lightcone4.2Representations in Hilbert Spaces4.3Möbius Transformations	<b>19</b> 19 21 24
5	The Discrete Series Represented on Anti-de Sitter Space5.1Lie Derivatives and Killing Vector Fields5.2Anti-de Sitter Space and the Discrete Series	<b>27</b> 27 28
6	Conclusions and Further Discussion	31
A	Acknowledgements	
Bi	Bibliography	
<b>A</b> ]	Appendix: Review of Group Theory         A.1: What Is a Group?         A.2: Lie Groups and Lie Algebras	

# 1 Introduction

The Lorentz group is the group of isometries on Minkowski space that keep the origin fixed. Intuitively, its elements are linear transformations of this space that preserve spacetime distances. The study of the Lorentz group was pioneered by mathematical physicists like Wigner [17] and Bargmann [1], and the groups are interesting not only for understanding the geometry of Minkowski and related spaces, but also because they have profound physical applications. For instance, close relatives of the Lorentz groups are the symmetry groups of *anti-de Sitter* space (indeed, the groups coincide with each other in 2+1 dimensions), a space studied extensively by theoretical physicists. The groups become highly relevant in trying to formulate a quantum field theory in anti-de Sitter space, as discussed in e.g. Fronsdal [7]. In these applications, it is important to find representations such that the "quantum number" corresponding to energy is bounded below (cf. Section 5 below). Similarly, the groups arise in the study of supersymmetry (Nicolai [15]), a proposed extension of the so-called standard model of particle physics. Also, there is no shortage of string theory papers where interest is taken in anti-de Sitter space, but the most famous is arguably Maldacena [14] (when this thesis was written, his paper had a modest 19,148 citations).

However, the Lorentz group also shows up in less expected places. For instance, Yurke *et al.* [18] characterised optical interferometers through a group-theoretic framework, and proposed a class of interferometers based on the group SU(1, 1) which, as we shall see, is closely related to the Lorentz group. An accessible introduction to this field is Han & Kim [10]. All these examples go to show that the study of the Lorentz group is likely worthwile - not only from a mathematical standpoint, but also from a physics perspective.

It is instructive to compare the Lorentz group and the *special orthogonal group*, or simply the *rotation group*, the group of rotations in standard euclidean space. This group is typically well-known by undergraduate physics students, and shows up in many different contexts. It arises naturally in e.g. Goldstein [8] in the study of rigid body motion in classical mechanics, but also when investigating angular momentum in quantum mechanics (section 3.3 of Sakurai [16] provides an excellent overview). More interestingly, quantum mechanics textbooks will often treat *representations* of the group, something to be discussed in this thesis, although this word is not always used in that context.

As it turns out, the unitary irreducible representations (the precise meaning of these words are defined further down) of the rotation group SO(3) have a deep physical interpretation, as they precisely describe possible spin states of a single particle in quantum mechanics. Given that the Lorentz group is the analogous group in Minkowski space, it therefore becomes relevant to study representations of the Lorentz group in 2 + 1 dimensions, and this is precisely the purpose of this thesis.

The over-arching structure of this thesis will follow that of Barut & Fronsdal [2] which, in turn, is a re-derivation of the original paper of Bargmann [1] by means of simpler algebraic methods. This paper is chosen because it very closely follows the methods to derive the representations of the rotation group, as they are given in standard quantum mechanical textbooks like Sakurai [16]. In developing the theory in this way, we will be able to see exactly when and why the two groups differ, and the implications that this has.

#### 1.1 Thesis Structure

The present thesis is structured with the undergraduate physics student in mind, and so the mathematical theory is developed at points where it becomes necessary. The only exception is the Appendix, to which introductory group theory has been appended. The reason for this is that although group theory is essential to this thesis, no fully satisfying account of the theory can be given in a thesis like this, and so only a brief review is given with references to further reading.

Section 2 defines a number of important groups, most notably the Lorentz group itself. Also in this section are a few mathematical proofs of isomorphisms between some groups, which enables us to work with not the Lorentz group itself, but instead the related group SU(1,1), as well as some definitions regarding group representations and covering groups. In Section 3 all irreducible representations of the rotation group and the Lorentz group are found. Firstly, the linear irreducible representations are constructed, and it is then investigated which of these can be made unitary. Then, Section 4 and Section 5 provide some physical contexts in which representations of the Lorentz group are highly relevant. Specifically, Section 4 discusses the hyperbolic plane and the projective lightcone, and the representations found are realised as functions on these spaces. Also discussed in this section are Möbius transformations, which turn out to connect some of the concepts used in this thesis. Section 5 follows up with an exploration of anti-de Sitter space and a class of representations viewed from this perspective. Finally, Section 6 gives a summary of the thesis, as well as discussing some questions for which there was not enough time.

# 2 A Few Groups and Group Representations

### 2.1 The Rotation Group and the Lorentz Group

Let us start by giving a definition of the *rotation group*, since it is perhaps easier to visualise.

**Definition 2.1.** The rotation group SO(3) is defined as the set of unit determinant, orientationpreserving linear transformations on  $\mathbb{R}^3$  that leave the form

$$X_1^2 + X_2^2 + X_3^2$$

invariant. The group operation, here, is composition of transformations.

Intuitively, the rotation group preserves spheres centered on the origin in 3-dimensional euclidean space. By changing an important sign, we are ready to give a more formal definition of the *Lorentz group*.

**Definition 2.2.** The 2+1-dimensional proper, orthochronous Lorentz group  $SO^+(2,1)$  is defined to be the set of unit determinant, orientation-preserving linear transformations on 2+1-dimensional Minkowski space that leave the form

$$X_1^2 + X_2^2 - X_3^2$$

invariant. Again, the group operation is composition of transformations.

When we write "the Lorentz group" from now on, we shall actually refer to the proper, orthochronous Lorentz group in 2+1 dimensions. It can be seen that the surfaces preserved by the Lorentz group are one family of one-sheeted hyperboloids, one family of two-sheeted hyperboloids, and the lightcone centered on the origin.

By extending the definitions slightly, one similarly defines the groups SO(n) and SO(n - k, k)for natural numbers n, k. However, this thesis will be restricted to n = 3 and k = 0, 1 since the most important aspects of the two groups are captured in this dimension. Furthermore, it so happens that the Lorentz group (on 2 + 1-dimensional Minkowski space) and the symmetry group for 1+1-dimensional anti-de Sitter space coincide, which is helpful for the discussion in Section 5. In fact, the reason that the 2+1 Lorentz group acts as symmetries on 1+1 Anti-de Sitter space is that this space can be viewed as a one-sheet hyperboloid embedded in 2+1 Minkowski space.

### 2.2 Some More Matrix Groups

Apart from the above groups, there are a few other matrix groups that are interesting to us. We shall define these below, starting with the most important group.

**Definition 2.3.** The general linear group of order n, denoted  $GL(n, \mathbb{C})$ , is the set of invertible  $n \times n$  complex matrices, with the group operation being matrix multiplication.

By restricting the determinant to unity, an equally important subgroup is obtained.

**Definition 2.4.** The special linear group of order n, denoted  $SL(n, \mathbb{C})$ , is the set of unit determinant  $n \times n$  complex matrices, again with matrix multiplication as the group operation.

The remaining groups in this subsection are all subgroups of the special linear group for some given n, as is the case below with n = 2.

**Definition 2.5.** The set of  $2 \times 2$  real matrices with unit determinant form the special linear group of order 2, denoted  $SL(2,\mathbb{R})$ . Thus, an element  $S \in SL(2,\mathbb{R})$  takes the form

$$S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ ad - bc = 1$$
(2.1)

with  $a, b, c, d \in \mathbb{R}$ .

Viewed as linear isomorphisms on  $\mathbb{R}^2$ , one sees that  $SL(2,\mathbb{R})$  precisely corresponds to areapreserving transformations.

**Definition 2.6.** The set of  $2 \times 2$  complex matrices of the form

$$U = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}, \ |\alpha|^2 + |\beta|^2 = 1$$
(2.2)

where  $\alpha, \beta \in \mathbb{C}$  and superscript \* denotes complex conjugation, form the **special unitary group**.

With a slight change of sign we obtain a closely related group, for which there is no common name as far as the author is aware.

**Definition 2.7.** The set of  $2 \times 2$  complex matrices of the form

$$U = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}, \ |\alpha|^2 - |\beta|^2 = 1$$
(2.3)

gives rise to another group, denoted SU(1,1).

Comparing with Definition 6.6 in the Appendix, we note that SU(2) is a compact Lie group, since it suffices to set C equal to any number greater than unity. On the contrary, SU(1,1) is non-compact, since the magnitudes of the parameters  $\alpha, \beta$  can be chosen arbitrarily large. This has consequences for the dimensions of the representations of the rotation group versus the Lorentz group, as discussed in Section 2.4.

It is a somewhat striking feature of matrix groups in small dimensions to coincide, and the above groups are prime examples.

**Theorem 2.1.** The groups  $SL(2,\mathbb{R})$  and SU(1,1) are isomorphic.

A proof of this theorem is presented below.

*Proof.* Fix the matrix

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}$$
(2.4)

and consider an arbitrary matrix  $U = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \in SU(1,1)$ . We want to show, firstly, that

$$TUT^{-1} \equiv S_U \in SL(2,\mathbb{R}) \tag{2.5}$$

Therefore,  $T^{-1}$  is computed. Since  $det(T) = \sqrt{2}$ , the inverse is found to be

$$T^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$
(2.6)

This allows us to compute  $S_U$  as

$$S_{U} = \frac{1}{2} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \beta^{*} & \alpha^{*} \end{pmatrix} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$
$$= \frac{1}{2} \begin{pmatrix} \alpha + i\beta + \alpha^{*} - i\beta^{*} & i\alpha + \beta - i\alpha^{*} + \beta^{*} \\ -i\alpha + \beta + i\alpha^{*} + \beta^{*} & \alpha - i\beta + \alpha^{*} + i\beta^{*} \end{pmatrix}$$
$$= \begin{pmatrix} \Re(\alpha) - \Im(\beta) & -\Im(\alpha) + \Re(\beta) \\ \Im(\alpha) + \Re(\beta) & \Re(\alpha) + \Im(\beta) \end{pmatrix}$$
(2.7)

Evidently, this matrix is real. To verify that it is contained in  $SL(2,\mathbb{R})$ , it remains to show that its determinant equals one. This will follow from the fact that  $|\alpha|^2 - |\beta|^2 = 1$ . Explicitly, we find

$$\det (S_U) = [\Re(\alpha)]^2 - [\Im(\beta)]^2 - (-[\Im(\alpha)]^2 + [\Re(\beta)]^2)$$
  
$$= [\Re(\alpha)]^2 + [\Im(\alpha)]^2 - ([\Re(\beta)]^2 + [\Im(\beta)]^2)$$
  
$$= |\alpha|^2 - |\beta|^2$$
  
$$= 1$$
(2.8)

Thus  $TUT^{-1} \in SL(2, \mathbb{R})$ , as was the claim.

We have found a function of SU(1,1) into  $SL(2,\mathbb{R})$ . It is clearly a group homomorphism since

$$TUVT^{-1} = TUT^{-1}TVT^{-1} = S_U S_V (2.9)$$

for all  $U, V \in SU(1, 1)$ . Moreover, it must be surjective; consider a matrix  $S \in SL(2, \mathbb{R})$  given by

$$S = \left(\begin{array}{cc} a & b \\ c & d \end{array}\right) \tag{2.10}$$

where the requirement is that ab - cd = 1. By defining  $\alpha$  and  $\beta$  such that

$$a = \Re(\alpha) - \Im(\beta) \tag{2.11a}$$

$$b = -\Im(\alpha) + \Re(\beta) \tag{2.11b}$$

$$c = \Im(\alpha) + \Re(\beta) \tag{2.11c}$$

$$d = \Re(\alpha) + \Im(\alpha) \tag{2.11d}$$

we are guaranteed that  $U = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \in SU(1,1)$ , and also  $TUT^{-1} = S$ .

Lastly, injectivity must be shown. For this, note that the inverse mapping  $U = T^{-1}S_UT$  is well-defined.

A bijective group homomorphism between SU(1,1) and  $SL(2,\mathbb{R})$  has been found, and it is so an isomorphism. Therefore,

$$SU(1,1) \cong SL(2,\mathbb{R}) \tag{2.12}$$

as was the claim.

### 2.3 Establishing an Important Isomorphism

This section is devoted to showing that the Lorentz group is isomorphic, up to a sign, to the group  $SL(2,\mathbb{R}) \cong SU(1,1)$ . Let us first define what "up to a sign" means in this context.

The subset  $\{1, -1\}$ , with 1 the 2 × 2 identity matrix, of  $SL(2, \mathbb{R})$  (or of SU(1, 1)) forms a normal subgroup in  $SL(2, \mathbb{R})$  (or in SU(1, 1)). By taking the quotient with this group, a subgroup is obtained in which we have identified each matrix with its corresponding negative matrix, and this group is given a special name.

**Definition 2.8.** The projective special linear group  $PSL(2, \mathbb{R})$  is defined as the quotient group

$$PSL(2,\mathbb{R}) = SL(2,\mathbb{R})/\{1,-1\}$$
(2.13)

Similarly, we define

$$PSU(1,1) = SU(1,1)/\{1,-1\}$$
(2.14)

Since  $SL(2,\mathbb{R})$  and SU(1,1) are isomorphic, so are  $PSL(2,\mathbb{R})$  and PSU(1,1). This group turns out to be closely related to  $SO^+(2,1)$  (isomorphic, in fact), which is suggested by re-writing position vectors as symmetric matrices. Recall that points in 2+1-dimensional Minkowski space are conveniently labelled by inertial coordinates  $(X_1, X_2, X_3)$ . Out of these coordinates, we form a symmetric matrix

$$X = \begin{pmatrix} X_1 + X_3 & X_2 \\ X_2 & -X_1 + X_3 \end{pmatrix}$$
(2.15)

The determinant of this matrix then satisfies

$$-\det(X) = X_1^2 + X_2^2 - X_3^2 \tag{2.16}$$

which is precisely the norm squared of the position vector  $\mathbf{X} = (X_1, X_2, X_3)$  in Minkowski space. Now, take a matrix S from  $SL(2, \mathbb{R})$  and consider the mapping

$$X \mapsto X' = SXS^T \tag{2.17}$$

where superscript T denotes matrix transposition. It is seen that transformations of this kind are linear, in the sense that if X, Y are two matrices constructed from inertial components and a, b are real numbers, then

$$aX + bY \mapsto aX' + bY' = aSXS^T + bSYS^T \tag{2.18}$$

Also, by properties of the determinant,

$$\det(X') = \det\left(SXS^T\right) = \det(X) \tag{2.19}$$

Viewed in this way, the matrices of  $SL(2,\mathbb{R})$  preserve the form  $X_1^2 + X_2^2 - X_3^2$ , which was how we defined the Lorentz group above. We note, though, that S and -S map X to the same matrix, and so we are led to believe the following theorem.

### **Theorem 2.2.** The groups $PSL(2, \mathbb{R})$ and $SO^+(2, 1)$ are isomorphic.

The above arguments provide an outline for a proof of this theorem. It remains to show that this map between  $PSL(2,\mathbb{R})$  and  $SO^+(2,1)$  is surjective, in the sense that every transformation in the Lorentz group is "hit" by a matrix from  $PSL(2,\mathbb{R})$ . Since this part of the proof does not provide much added insight, it has been omitted for brevity.

With this isomorphism established, we will be able to use the three groups  $PSL(2, \mathbb{R})$ , PSU(1, 1)and  $SO^+(2, 1)$  interchangeably to study properties of the Lorentz group. For instance, noting that SU(1, 1) is non-compact, we deduce that the same is true of PSU(1, 1), so by extension  $SO^+(2, 1)$ is non-compact. On the contrary, SU(2) is compact, and so PSU(2) (an analogously defined group) is and therefore also SO(3).

#### 2.4 Group Representations and Irreducibility

There are two more concepts that will be essential in our study of the Lorentz group, and one of them is the notion of a *group representation*. In principle, the properties of the groups defined above can all be deduced from their respective definitions, which are known as the *defining* representations of the groups. However, there exist many other possible representations, and it is not always the defining representations that arise from physical (or mathematical) investigations.

The definition of a group representation (slightly shortened from [6]) is given below.

**Definition 2.9.** Let G be a group,  $\mathbb{F}$  a field and n a positive integer. A **representation** of G is a homomorphism  $\rho : G \to GL(n, \mathbb{F})$ . If, furthermore, the homomorphism is injective, the representation is *faithful*.

In the above definition,  $GL(n, \mathbb{F})$  is the group of  $n \times n$  invertible matrices with elements from the field  $\mathbb{F}$ . For our purposes, the field  $\mathbb{F}$  will always be either  $\mathbb{R}$  or  $\mathbb{C}$ . Strictly speaking, the above definition defines a *matrix* representation, but such a representation is naturally identified with a *linear* representation [6], and we shall use the word "representation" to mean either kind.

If, in addition, all matrices in a representation of the above form are unitary, we shall say that *the representation itself is unitary*. This will be an important feature when we look for representations of physical significance in quantum mechanics.

Looking at the above definition, one observes that every group has a representation, given by the map  $g \mapsto 1$  for all  $g \in G$ , where 1 is the  $n \times n$  identity matrix. This representation is called the *trivial representation*. On the other hand, such a representation is never faithful, unless the group G contains only the identity element. Similarly, for a given representation, it could happen that one or more subspaces of  $\mathbb{F}^n$  is mapped to itself by all elements of  $\rho(G)$ , so that the representation can, in some sense, be decomposed into representations acting on those subspaces. This motivates the definition of irreducibility.

**Definition 2.10.** A representation  $\rho : G \to GL(n, \mathbb{F})$  is **irreducible** if the only subspaces  $V \subseteq \mathbb{F}^n$  such that  $\rho(G)$  sends V to itself is the zero vector space  $\{\mathbf{0}\}$  and  $\mathbb{F}^n$  itself. Otherwise, the representation is **reducible**.

A subspace  $V \subseteq \mathbb{F}^n$  mapped to itself by the elements of  $\rho(G)$  is said to be *invariant* under the action of G.

Now that we have established a language for discussing the representations of groups, there is a remark that can already be made on the qualitative difference between representations of the rotation group and of the Lorentz group. In general, the unitary irreducible representations of a non-compact Lie group are infinite-dimensional, something that is discussed in both [1] and [2]. With this in mind, it will not appear surprising that it is found in Section 3.2 that all such representations of the rotation group are finite-dimensional, while none of those for the Lorentz group are (apart from the trivial representation).

Finally, we want a notion of "sameness" of two representations, since two representations of the same group in different vector spaces may superficially look very different but can turn out to be identical.

**Definition 2.11.** Let G be a group, and suppose  $\rho : G \to GL(n, \mathbb{E})$  and  $\sigma : G \to GL(m, \mathbb{F})$  are two representations of G, with  $\mathbb{E}, \mathbb{F}$  fields and n, m natural numbers. Then  $\rho$  and  $\sigma$  are **equivalent** provided there exists a linear isomorphism  $L : \mathbb{E}^n \to \mathbb{F}^m$  such that

$$L \circ \rho(g) \circ L^{-1} = \sigma(g) \tag{2.20}$$

for all  $g \in G$ . If not, the two representations are **inequivalent**.

### 2.5 Covering Groups and Multiple Coverings

Finally, a discussion on *covering groups* will establish (among other things) something that has been hinted at in the previous sections, which is that the unitary irreducible representations of the rotation group will be finite-dimensional, while those of the Lorentz group will not.

We saw above that  $PSL(2, \mathbb{R})$  is isomorphic to the Lorentz group, rather than  $SL(2, \mathbb{R})$  itself. Therefore, the map  $SL(2, \mathbb{R}) \to SO^+(2, 1)$  is not one-to-one, but rather two-to-one. As another example of this, we could consider the map from the real line to SO(2) (whose group manifold is the circle) given by  $x \mapsto \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix}$ . Such a mapping is then infinitely many-to-one. The following definition seeks to formalise this idea.

**Definition 2.12.** A Lie group G is **covered** by a **covering group** C provided there exists a **covering map**  $p: C \to G$  with the following properties.

- The covering map p is continuous and surjective.
- The group C covers G evenly. That is to say, for every element  $g \in G$ , there exists an open neighbourhood  $U \subseteq G$  around g such that its pre-image  $p^{-1}(U)$  in C is the disjoint union of a set of open neighbourhoods in C, each of which is homeomorphic to U via p.
- The mapping p is a continuous group homomorphism.

The surjectivity ensures that, in some sense, the group is fully covered, while the second condition implies that  $p^{-1}(g)$  is a discrete set for each  $g \in G$ . If the third condition is removed, the more general definition of a *covering space* is obtained, valid for any *topological space*.

In particular, it is the second axiom that allows for the covering group C to cover G multiple times, provided any small region  $V \subseteq C$  "looks like"  $p(V) \subseteq G$ . If, for every  $U \subseteq G$  like in the above definition,  $p^{-1}(U)$  contains exactly two disjoint sets, then C is a *double cover* of G.

Equipped with the above terminology, we note that  $SL(2,\mathbb{R})$  is a double cover of the Lorentz group, as is SU(1,1). Similarly, the special unitary group SU(2) is a double cover of the rotation group.

# 3 Representations of the Rotation and Lorentz Groups

We now go on to construct representations for the rotation group and for the Lorentz group. The methods employed here mostly follow the standard discussion of angular momentum in advanced quantum mechanics textbooks like Sakurai [16].

#### 3.1 Linear Representations of Both Groups

Following [2], we shall study the representations of the Lie algebra of the group SU(1,1), instead of directly studying the Lorentz group. The reason for studying the algebra, as was mentioned in the Appendix, is that those are typically easier to work with. Also, we might as well work with SU(1,1) first, and later find the representations of  $SO^+(2,1)$  as a special case. Its algebra has three linearly independent generators, i.e. a basis, denoted by  $J_{12}, J_{13}, J_{23}$ . They are defined to satisfy the relations [2]

$$J_{\alpha\beta} + J_{\beta\alpha} = 0 \tag{3.1a}$$

$$[J_{\alpha\gamma}, J_{\gamma\beta}] = -ig_{\gamma\gamma}J_{\alpha\beta} \tag{3.1b}$$

(with no summation performed over  $\gamma$ ). Here,  $g_{\alpha\beta}$  refers to the Minkowski metric, which is diagonal and in this convention is given by  $g_{11} = g_{22} = -g_{33} = 1$ . The reason that the generators have two indices, unlike the single index objects seen for the case of SO(3), is that this notation more easily generalises to higher dimensions. Specifically, it is true in any dimension that the action of a rotation can be decomposed using a family of parallel planes, such that points in a given plane are transformed to other points in the same plane. The same also holds for Lorentz boosts. On the contrary, it is a special property of 3-dimensional space that there exists a unique (up to sign) unit normal vector to each plane, which is why rotations can be said to be defined by a single axis direction.

Now, since we *are* studying the special case of three dimensions, we could define a set of singleindex operators by

$$J_{\alpha\beta} = \epsilon_{\alpha\beta\gamma} J^{\gamma} \tag{3.2}$$

where  $\epsilon_{\alpha\beta\gamma}$  is the totally antisymmetric Levi-Civita symbol with  $\epsilon_{123} = 1$ . Also, here the Einstein summation convention is employed, which states that an index is implicitly summed over when it appears once upstairs and once downstairs. We shall continue to employ the Einstein convention throughout this thesis.

We could then re-write the equations (3.1b) in a way that more resembles the familiar case:

$$[J_2, J_3] = iJ_1 \tag{3.3a}$$

$$[J_3, J_1] = iJ_2 \tag{3.3b}$$

$$[J_1, J_2] = -iJ_3 \tag{3.3c}$$

where the index-lowering property of the metric tensor was used, namely  $J_{\alpha} = g_{\alpha\beta}J^{\beta}$ . We see that the commutation relations are almost identical, but not equivalent, to those for the rotation group. Although it is illustrative to work with the single-index operators in three dimensions, they will not be pursued further since they do not generalise to higher dimensions like the double-index operators do. We can go even further in our attempts to make the algebras look similar. For this purpose, the operators  $M^+$  and  $M^-$  are defined, called the *raising* and *lowering* operators respectively, via

$$M^{\pm} = \frac{1}{\sqrt{2}} \left( \pm J_{23} + iJ_{13} \right) \tag{3.4}$$

This is to be compared with equation (3.5.5) in [16], where the corresponding operators for SU(2) are  $M^{\pm} = \frac{1}{\sqrt{2}} (J_{23} \pm i J_{13})$ . Then, the commutation relations for  $J_{12}, M^+, M^-$  are identical for both groups:

$$[J_{12}, M^{\pm}] = \pm M^{\pm} \tag{3.5a}$$

$$[M^+, M^-] = J_{12}$$
 (3.5b)

This allows us to find the linear representations for both groups simultaneously. Given this, it is important to realise where the two groups differ. In a unitary representation, it will be required that  $J_{ij}^{\dagger} = J_{ij}$ , so that elements of the Lie algebra are hermitian. Then, looking at (3.4),

$$\left(M^{\pm}\right)^{\dagger} = -M^{\mp} \tag{3.6}$$

while, for the rotation group,  $(M^{\pm})^{\dagger} = M^{\mp}$ . We see that the groups must necessarily differ in their unitary representations.

Let us return to the search for linear representations, by considering the action of these operators on a vector space. We define a set of vectors by

$$|a,b\rangle = \xi_1^a \xi_2^b \tag{3.7}$$

for a two-component spinor  $(\xi_1, \xi_2)$  and where, a priori, a and b may take arbitrary complex values (this thesis will contain no deeper discussion of spinors - for us, they are simply the vectors in the space on which these operators act). In this representation, the operators  $J_{12}$  and  $M^{\pm}$  are given by [2]

$$J_{12} = \frac{1}{2} \left( \xi_1 \frac{\partial}{\partial \xi_1} - \xi_2 \frac{\partial}{\partial \xi_2} \right)$$
(3.8a)

$$M^{+} = \frac{1}{\sqrt{2}} \xi_1 \frac{\partial}{\partial \xi_2} \tag{3.8b}$$

$$M^{-} = \frac{1}{\sqrt{2}} \xi_2 \frac{\partial}{\partial \xi_1} \tag{3.8c}$$

These equations may seem to appear out of nowhere, but it can be shown that the operators indeed satisfy the imposed commutation relations. As an example, we compute  $[J_{12}, M^+]$  below:

$$\begin{bmatrix} J_{12}, M^+ \end{bmatrix} = J_{12}M^+ - M^+ J_{12}$$

$$= \frac{1}{2\sqrt{2}}\xi_1 \frac{\partial}{\partial\xi_1} \left(\xi_1 \frac{\partial}{\partial\xi_2}\right) - \frac{1}{2\sqrt{2}}\xi_2 \frac{\partial}{\partial\xi_2} \left(\xi_1 \frac{\partial}{\partial\xi_2}\right) - \frac{1}{2\sqrt{2}}\xi_1 \frac{\partial}{\partial\xi_2} \left(\xi_1 \frac{\partial}{\partial\xi_1} - \xi_2 \frac{\partial}{\partial\xi_2}\right)$$

$$= \frac{1}{2\sqrt{2}}\xi_1 \frac{\partial}{\partial\xi_2} + \frac{1}{2\sqrt{2}}\xi_1^2 \frac{\partial^2}{\partial\xi_1 \partial\xi_2} - \frac{1}{2\sqrt{2}}\xi_1 \xi_2 \frac{\partial^2}{\partial\xi_2^2} - \frac{1}{2\sqrt{2}}\xi_1^2 \frac{\partial^2}{\partial\xi_2 \partial\xi_1} + \frac{1}{2\sqrt{2}}\xi_1 \frac{\partial}{\partial\xi_2} + \frac{1}{2\sqrt{2}}\xi_1 \xi_2 \frac{\partial^2}{\partial\xi_2^2}$$

$$= \frac{1}{\sqrt{2}}\xi_1 \frac{\partial}{\partial\xi_2} = M^+$$
(3.9)

as was the claim. By letting the operators defined in (3.8a)-(3.8c) act on the vectors  $|a, b\rangle$ , it is found that

$$J_{12} |a, b\rangle = \frac{1}{2} (a - b) |a, b\rangle$$
(3.10a)

$$M^{+}|a,b\rangle = \frac{1}{\sqrt{2}}b|a+1,b-1\rangle$$
 (3.10b)

$$M^{-}|a,b\rangle = \frac{1}{\sqrt{2}}a|a-1,b+1\rangle$$
 (3.10c)

By observing the above action, one finds a few invariants. Firstly, the complex number

$$j = \frac{1}{2}(a+b)$$
(3.11)

is seen to be left unchanged by the operators. In any irreducible representation, therefore, its value must be fixed. This value is related to the operator Q, called the *Casimir operator*, given by

$$Q = 2M^{-}M^{+} + J_{12}\left(J_{12} + 1\right) \tag{3.12}$$

since it can be seen that

$$Q|a,b\rangle = j(j+1)|a,b\rangle \tag{3.13}$$

so that Q is a scalar in a given representation. For the case of the rotation group, Q turns out to equal  $\mathbf{J}^2$ , the total angular momentum squared. For the Lorentz group, it is found that

$$Q = -(J_{23})^2 - (J_{13})^2 + (J_{12})^2$$
(3.14)

We see that the different signs in the metric give rise to some relative signs in the Casimir operator, but there is not much more to say about this.

Next, one finds that  $\frac{1}{2}(a-b)$  can only change its value by integers, and so the fractional part is another invariant, and the permissible values of the expression are

$$\frac{1}{2}(a-b) = E_0 + m \tag{3.15}$$

for integers *m*. Without loss of generality, we may impose  $-\frac{1}{2} \leq \Re(E_0) < \frac{1}{2}$ .

If j and  $E_0$  are fixed, then the labelling of the basis vectors by m is unambiguous. The only problem could be that the representation is reducible still, after having fixed these values. This will happen when either a, b, or both take integer values. This is because at least one of the relations (3.10b)-(3.10c) vanishes when either a or b equals 0, and it follows that there will exist an invariant subspace. With this observation, we are able to classify all possible linear representations by investigating the cases separately.

(A) If neither a nor b are integers, the representations above are irreducible. It remains to identify equivalent representations, which occur when  $E'_0 = E_0$  and j' = j or j' = -j - 1 [2]. For this reason, all representations with different Q and/or  $E_0$  are inequivalent.

This family of representations will be denoted  $D(Q, E_0)$ , following the convention in [2].

(B) If a, and possibly b, takes integer values, then  $M^{-}|0,b\rangle = 0$ , and so the subspace  $a \ge 0$  is invariant. Therefore, to obtain an irreducible representation, the vector space must firstly be restricted to  $a \ge 0$ .

Next, we must consider if restrictions have to be put on b. The claim is that the values of b must be such that 2j is not a non-negative integer, i.e. that j is not zero or a positive half-integer. To see why, note that b = 2j - a is bounded above by 2j. Hence, if 2j is a non-negative integer, then for a = 2j we find that b = 0. Noting that  $M^+ |a, 0\rangle = 0$ , the subspace  $b \ge 0$  would be an invariant subspace, and the representation would be reducible. It is for this reason b is restricted such that 2j is not a non-negative integer.

Lastly, we see that there is a natural smallest "fractional part" of  $\frac{1}{2}(a-b)$  obtained when a = 0. Denoting the corresponding value by  $E_0$  (i.e. dropping the restriction that  $-\frac{1}{2} \leq E_0 < \frac{1}{2}$ ), it is found that  $E_0 = -j$  since, for the *b* such that a = 0, we have  $j = \frac{1}{2}(a+b) = \frac{1}{2}b$ . It follows that the spectrum of  $\frac{1}{2}(a-b)$  is

$$\frac{1}{2}(a-b) = -j+m, \ m \ge 0 \tag{3.16}$$

where only non-negative integers  $m \ge 0$  are allowed because  $a \ge 0$ . Equivalently, the spectrum is obtained by repeated applications of  $M^+$  to the state  $|0, 2j\rangle$ .

We denote this family of representations by  $D^+(j)$ .

(C) If b, and possibly a, takes integer values, then for the same reasons irreducibility is met if  $b \ge 0$  and 2j is not a non-negative integer. This time, there is a natural largest "fractional part"  $E_0 = j$ , and the spectrum of  $\frac{1}{2}(a-b)$  is

$$\frac{1}{2}(a-b) = j+m, \ m \le 0 \tag{3.17}$$

as is seen by applying  $M^-$  repeatedly to  $|2j, 0\rangle$ .

This family of representations is written  $D^{-}(j)$ , and collectively the representations  $D^{\pm}(j)$  are called the *discrete series*.

(D) Lastly, if both a and b are integers, two cases must be considered. If a + b < 0, then either of the cases (B) or (C) above are recovered. We therefore restrict our attention to when  $a+b \ge 0$ . To obtain irreducibility, we must set  $a \ge 0$  and  $b \ge 0$ . Then, 2j is a positive integer or zero, so let us fix 2j to be a non-negative integer. We note that a and b are both bounded above by 2j, and this gives rise to the finite spectrum

$$\frac{1}{2}(a-b) = -j, -j+1, \dots, j-1, j$$
(3.18)

We write D(j) for this family of representations.

### 3.2 Unitary Representations of the Rotation Group and the Lorentz Group

We now go on to consider which, if any, of the above representations can be made unitary by some transformation of the basis vectors. The claim is that it can be done while keeping  $J_{12}$  diagonal, and the argument for why that is true is, to some extent, that it works. However, it is perhaps not so surprising that it is so, since the general strategy in quantum mechanics is to find a maximal set of diagonalisable, mutually commuting operators, and we know this to be possible for the rotation group. With this being said, it turns out [2] that the operators  $J_{13}$  and  $J_{23}$  have continuous spectra, which makes for a significantly tougher analysis than what is given in this thesis. For a more thorough discussion on this matter, the reader is referred to Lindblad & Nagel [13], where the analysis is carried out for  $J_{13}$  diagonal.

Proceeding with  $J_{12}$  diagonal, we introduce a set of normalising factors  $\{N_m\}$  together with the set of basis vectors

$$|j,m\rangle = N_m |a,b\rangle = N_m \xi_1^a \xi_2^b \tag{3.19}$$

The question is whether there exist normalising factors  $N_m$  such that

$$\langle j, m' | j, m \rangle = \delta_{m'm} \tag{3.20}$$

can consistently be imposed, where  $\delta_{nm}$  is the Kronecker delta. One finds in this representation that, analogously to (3.10a)-(3.10c), the actions of our operators are found by substituting a and b for j and m, remembering to take the factors  $N_m$  into account. The new relations are

$$J_{12}|j,m\rangle = (E_0 + m)|j,m\rangle$$
 (3.21a)

$$M^{+}|j,m\rangle = \frac{1}{\sqrt{2}} \left(j - E_{0} - m\right) \frac{N_{m}}{N_{m+1}} |j,m+1\rangle$$
(3.21b)

$$M^{-}|j,m\rangle = \frac{1}{\sqrt{2}} \left(j + E_{0} + m\right) \frac{N_{m}}{N_{m-1}} |j,m-1\rangle$$
(3.21c)

It is at this point that the groups start to differ, since the operators for the respective groups satisfy slightly different unitarity relations. For both, we have  $J_{12}^{\dagger} = J_{12}$ . Therefore, noting that  $J_{12}|j,m\rangle \propto |j,m\rangle$ , a particular inner product can be evaluated in two ways. Namely, consider  $\langle j,m|J_{12}|j,m\rangle$ . On the one hand, applying  $J_{12}$  to the right gives

$$\langle j, m | J_{12} | j, m \rangle = E_0 + m$$
 (3.22)

On the other, since  $J_{12}^{\dagger} = J_{12}$ , we can have  $J_{12}$  act to the left instead. Then,

$$\langle j, m | J_{12} | j, m \rangle = \langle j, m | J_{12}^{\dagger} | j, m \rangle = E_0^* + m$$
(3.23)

Equating the two, we require that

$$E_0^* + m = E_0 + m \tag{3.24}$$

Since m takes integer values, it must be that  $\Im(E_0) = 0$ .

Now, the notation  $E_0$  seems to suggest a possible interpretation as a ground state energy, and we shall see in Section 5 that it indeed takes this role, for the special case of the discrete series  $D^+(j)$ . However, the value of  $E_0$  has a more general connection to group coverings. Consider the one-parameter subgroup G of the Lorentz group (or the rotation group, equivalently) generated by  $J_{12}$ , whose elements are

$$q_{\phi} = e^{i\phi J_{12}} \tag{3.25}$$

for real  $\phi$ , and where the group operation is multiplication. We note, in particular, that  $g_0 = 1$ , the identity operator. In a natural way, this group acts on the  $|j, m\rangle$  vectors by

$$g_{\phi} \left| j, m \right\rangle = e^{i\phi(E_0 + m)} \left| j, m \right\rangle \tag{3.26}$$

Recalling that the complex exponential  $f(x) = e^{ix}$  is periodic in its real argument x, one could ask whether the same is true of  $g_{\phi}$ . That is, is there a smallest positive  $\phi$  such that  $g_{\phi} = 1$ ? This

must depend on the eigenvalues of  $J_{12}$ , namely the value of  $E_0$ . Restricting ourselves to unitary representations, we consider real  $E_0$ .

We start by noting that if there exists a smallest positive number  $\tau$  such that  $g_{\tau} = 1$ , then the group G is still generated if  $\phi$  is restricted onto  $[0, \tau)$ . For such a number to exist,  $\tau$  must necessarily be an integer multiple of  $2\pi$ , since we require  $e^{im\tau} = 1$  for all (or at least infinitely many) integers m. This, in turn, forces

$$\tau E_0 = 2\pi p E_0 = 2\pi q \tag{3.27}$$

for some positive integer p and any integer q, or

$$E_0 = \frac{q}{p} \tag{3.28}$$

That is, whenever  $E_0$  is a rational number, then there exists a period  $\tau$  and we may restrict  $\phi$  onto  $[0,\tau)$ . Conversely, if  $E_0$  is irrational, then no such period exists, and  $g_{\phi} \neq g_{\psi}$  whenever  $\phi \neq \psi$  (so the map  $\phi \mapsto g_{\phi}$  with domain  $\mathbb{R}$  is one-to-one).

Now, it can be asked which of these coverings correspond to the groups discussed above. For the rotation group, one must set  $E_0$  to an integer or a half-integer. The first choice gives a faithful representation of SO(3), and the other a faithful representation of SU(2). For the Lorentz group, on the other hand, many more values for  $E_0$  are accessible, and in this way one could construct faithful representations of other coverings of the Lorentz group than the ones discussed in this thesis.

### 3.2.1 Unitarity Conditions for the Rotation Group

For the rotation group, recall that  $(M^+)^{\dagger} = M^-$ . Hence, noting that  $M^-M^+ |j,m\rangle \propto |j,m\rangle$ , we a similar trick as previously. On the one hand,

$$\langle j, m | M^{-}M^{+} | j, m \rangle = \frac{1}{\sqrt{2}} \left( j - E_{0} - m \right) \frac{N_{m}}{N_{m+1}} \left\langle j, m | M^{-} | j, m+1 \right\rangle$$
$$= \frac{1}{2} \left( j - E_{0} - m \right) \frac{N_{m}}{N_{m+1}} \left( j + E_{0} + m+1 \right) \frac{N_{m+1}}{N_{m}}$$
$$= \frac{1}{2} \left( j - E_{0} - m \right) \left( j + E_{0} + m+1 \right)$$
(3.29)

On the other,

$$\langle j, m | M^{-} M^{+} | j, m \rangle = \langle j, m | (M^{+})^{\dagger} M^{+} | j, m \rangle$$
  
=  $\frac{1}{\sqrt{2}} (j^{*} - E_{0} - m) \frac{N_{m}^{*}}{N_{m+1}^{*}} \cdot \frac{1}{\sqrt{2}} (j - E_{0} - m) \frac{N_{m}}{N_{m+1}}$   
=  $\frac{1}{2} (j^{*} - E_{0} - m) (j - E_{0} - m) \left| \frac{N_{m}}{N_{m+1}} \right|^{2}$  (3.30)

where we have used the fact that  $E_0$  is real. Again, equating the two gives

$$(j - E_0 - m)(j + E_0 + m + 1) = (j^* - E_0 - m)(j - E_0 - m)\left|\frac{N_m}{N_{m+1}}\right|^2$$
(3.31)

Requiring any kind of recursion relation,  $j - E_0 - m$  is assumed non-zero, and the equation is divided through by it. Lastly, solving for the normalising factors, the analogue of equation (4.4) in [2] is recovered:

$$\left|\frac{N_{m+1}}{N_m}\right|^2 = \frac{j^* - E_0 - m}{m + j + E_0 + 1} \tag{3.32}$$

We require the left-hand side to be real and non-negative, and so we must do the same of the right-hand side. It is clear from looking at the equation, however, that for sufficiently large m the numerator is negative while the denominator is positive, and vice versa when m is sufficiently small. In any unitary irreducible representation of the rotation group, then, m must be bounded above and below.

So, a familiar result from quantum mechanics is recovered. Namely, we see that when j is interpreted as the total angular momentum of a particle, then a component m thereof cannot take arbitrarily large or small values.

#### 3.2.2 Unitarity Conditions for the Lorentz Group

Recall that  $J_{12}^{\dagger} = J_{12}$  for this group, too, and so  $\Im(E_0) = 0$  by the same arguments. The difference, this time, is that  $(M^+)^{\dagger} = -M^-$ , and so when  $\langle j, m | M^- M^+ | j, m \rangle$  is evaluated like above, a relative sign between the left- and right-hand side is introduced. Therefore,

$$\left|\frac{N_{m+1}}{N_m}\right|^2 = -\frac{j^* - E_0 - m}{m + j + E_0 + 1} = \frac{m + E_0 + \frac{1}{2} - \left(j^* + \frac{1}{2}\right)}{m + E_0 + \frac{1}{2} + \left(j + \frac{1}{2}\right)}$$
(3.33)

We can no longer apply the argument above to deduce that m must be bounded above and below. Instead, we investigate each family of representations separately.

(A) For  $D(Q, E_0)$ , the recursion relation must hold for all integers m. We see that the right-hand side is necessarily positive for large enough or small enough m, so it is the intermediate case that needs investigation. Also, it must be considered whether j is allowed to take complex values or not. Let us investigate the case of m = 0. We then require that

$$\frac{E_0 + \frac{1}{2} - (j^* + \frac{1}{2})}{E_0 + \frac{1}{2} + (j + \frac{1}{2})} > 0$$
(3.34)

Since it is not known whether j is purely real or more generally complex, the latter is assumed. It is convenient to split j up into its real and complex parts, say, j = k + il. We find

$$\frac{E_0 + \frac{1}{2} - \left(k + \frac{1}{2}\right) + il}{E_0 + \frac{1}{2} + \left(k + \frac{1}{2}\right) + il} > 0$$
(3.35)

Assume, firstly, that  $l \neq 0$ . For the fraction to be positive at all, it must certainly be real. Since the complex parts of the numerator and denominator are equal, so too must the real parts be. This happens only if  $k = -\frac{1}{2}$ , and so one set of permissible values for j is

$$j = -\frac{1}{2} + il, \ l \in \mathbb{R}$$

$$(3.36)$$

With this choice of j, it is also clear that the fraction is positive for all integers m, since the fraction equals 1 identically.

On the other hand, assume now that j is purely real. Then, for m = 0 we have

$$\frac{E_0 + \frac{1}{2} - (j + \frac{1}{2})}{E_0 + \frac{1}{2} + (j + \frac{1}{2})} > 0$$
(3.37)

Recalling that, for this representation  $-\frac{1}{2} \leq E_0 < \frac{1}{2}$ , no representation can be possible for  $E_0 = -\frac{1}{2}$ . For this reason, we restrict the search to  $-\frac{1}{2} < E_0 < \frac{1}{2}$ .

We may multiply by the denominator squared on both sides of the above inequality to find

$$\left(\frac{1}{2} + E_0\right)^2 - \left(j + \frac{1}{2}\right)^2 > 0 \tag{3.38}$$

or

$$\left(j+\frac{1}{2}\right)^2 < \left(\frac{1}{2}+E_0\right)^2$$
 (3.39)

Since  $\frac{1}{2} + E_0 > 0$ , taking the square root on both sides yields

$$\left| j + \frac{1}{2} \right| < \frac{1}{2} + E_0 \tag{3.40}$$

Repeating the procedure for the case of m = -1, one finds

$$\left| j + \frac{1}{2} \right| < \frac{1}{2} - E_0 \tag{3.41}$$

For both equations to be true simultaneously, we need

$$\left| j + \frac{1}{2} \right| < \frac{1}{2} - |E_0| \tag{3.42}$$

which is the equation given in [2]. It remains to show that j has been restricted enough to be guaranteed a representation. However, we see from the above equation that the numerator and denominator are both positive for m = 0, and both negative for m = -1. It follows that the fraction is positive for arbitrarily large integers m, and for arbitrarily small integers m. We are done.

(B) For the representation  $D^+(j)$ ,  $m \ge 0$  and  $E_0 = -j$ . Since  $E_0$  is real, so is j, and in fact j must be real for the remaining representations. We therefore require

$$\frac{m-2j}{m+1} > 0 \tag{3.43}$$

for all non-negative integers m. This is satisfied precisely when j < 0.

(C) For  $D^{-}(j)$ , instead,  $E_0 = j$  and  $m \leq 0$ , and so we impose

$$\frac{m}{m+1+2j} > 0 \tag{3.44}$$

The largest integer m for which  $\left|\frac{N_{m+1}}{N_m}\right|^2$  is well-defined is m = -1. Inserting this above, it is seen that the restriction j < 0 is again sufficient.

(D) Lastly, for the representation D(j), then  $E_0 = 0$  and m = -j, -j + 1, ..., j - 1, j, where we have allowed m to take half-integer values too. If we try to impose

$$\frac{m-j}{m+1+j} > 0 \tag{3.45}$$

for m in the stated range, this fails because the numerator is non-positive, while the denominator is positive. The only possibility is that j = 0, for which there is no recursion in the first place since the basis for the vector space contains the single  $|0,0\rangle$  vector. Thus, the only unitary D(j) representation is the trivial one.

To distinguish between the two families of representations of the form  $D(Q, E_0)$ , we call (3.36) and (3.42) the *principal* and the *supplementary series*, respectively, denoted by  $D_P(Q, E_0)$  and  $D_S(Q, E_0)$ .

Lastly, let us find explicitly the normalising factors  $N_m$ . For the principal series, the recursion holds for  $N_m = 1$  identically, or indeed any other constant. For the remaining representations, the claim is that the choice

$$N_m = \left[\frac{(m+E_0-1-j)!}{(m+E_0+j)!}\right]^{1/2} = \left[\frac{\Gamma(m+E_0-j)}{\Gamma(m+E_0+1+j)}\right]^{1/2}$$
(3.46)

where  $\Gamma(z)$  is the gamma function, gives the correct recursion. The reason this works is that the gamma function satisfies the relation

$$\Gamma(z+1) = z\Gamma(z) \iff \frac{\Gamma(z+1)}{z} = \Gamma(z)$$
 (3.47)

Computing the ratio  $\left|\frac{N_{m+1}}{N_m}\right|^2$  using the above relation and that j is purely real in these cases indeed reproduces the recursion relation (3.33).

# 4 Some Physical Realisations of the Unitary Representations

This section is devoted to ascribing some physical meaning to the above unitary representations, by representing them as functions on suitable Hilbert spaces. After this, a discussion of Möbius transformations follows, which will connect many of the concepts introduced in the previous sections.

#### 4.1 The Hyperbolic Plane and the Projective Lightcone

We turn our attention to some spaces relevant for the various representations found above, starting with the *hyperbolic plane*.

**Definition 4.1.** The hyperboloid model for the hyperbolic plane  $\mathbb{H}^2$  is the hyperboloid

$$X_1^2 + X_2^2 - X_3^2 = -R^2, \ X_3 > 0 \tag{4.1}$$

with R > 0 a constant, embedded in 3-dimensional Minkowski space, together with the (standard Minkowski) metric

$$ds^2 = dX_1^2 + dX_2^2 - dX_3^2$$
(4.2)

It can be shown that the hyperbolic plane has constant gaussian curvature  $-\frac{1}{R^2}$ , and thus it is in a sense a non-compact analogue of the 2-sphere (for which the curvature is  $\frac{1}{R^2}$ ). From now on, R will be set to 1 for convenience. Also, we note that the Lorentz group is precisely the symmetry group on this hyperboloid.

In a similar way to the hyperboloid model, the lightcone  $X^{\alpha}X_{\alpha} = 0$  is mapped to itself by the Lorentz group. Since, moreover, the spacetime distance between any two points on the same lightray is zero, it is natural to view every lightray as a "point" in some other space, which results in a space called the *projective lightcone*. One way to go about this is to, for every lightray, choose a representative point of that lightray, e.g. the one with  $X_3 = 1$ . This lets us view the projective lightcone as the circle  $X_1^2 + X_2^2 = 1, X_3 = 1$ .

Coming back to the hyperboloid model, it should be stressed that it really defines a particular model of  $\mathbb{H}^2$ , and that there are many others. In particular, another model called the *Poincaré disk* model is obtained by stereographically projecting the hyperboloid sheet above to the  $X_3 = 0$  plane via the point  $(X_1, X_2, X_3) = (0, 0, -1)$ . That is to say, each point on the hyperboloid is connected to (0, 0, -1) by a straight line segment, and the intersection of that line with the plane  $X_3 = 0$  defines the mapped point. See Figure 1 for a depiction of this projection.



Figure 1: The projection from the hyperboloid  $X_3 = \sqrt{1 + X_1^2 + X_2^2}$  onto the disk  $X_1^2 + X_2^2 < 1, X_3 = 0$ . The point **a** is mapped by a straight line, connecting **a** with the point (0, 0, -1), to the point **a'** in the disk. The hyperboloid should be thought of as extending infinitely upwards.

It is seen that the hyperboloid is mapped to the unit disk  $X_1^2 + X_2^2 < 1, X_3 = 0$ . After some computation, the induced metric in this space can be found, and it is given in e.g. Cannon *et al.* [4]. We can make the formal definition.

**Definition 4.2.** The **Poincaré disk model** of the **hyperbolic plane**  $\mathbb{H}^2$  is the set of points  $(x_1, x_2) \in \mathbb{R}^2$  inside the unit disk

$$x_1^2 + x_2^2 < 1 \tag{4.3}$$

together with the metric

$$ds^{2} = 4 \frac{dx_{1}^{2} + dx_{2}^{2}}{\left(1 - x_{1}^{2} - x_{2}^{2}\right)^{2}}$$

$$(4.4)$$

This model of the hyperbolic plane has the advantage of using as many coordinates as the dimension of the space, and also that any point in the space is described by finite coordinates. We shall see this unit disk, as well as the projective lightcone, appearing from other arguments in the next section.

Another model of  $\mathbb{H}^2$  is the *half-plane* model, given below.

**Definition 4.3.** The half-plane model of the hyperbolic plane  $\mathbb{H}^2$  is the set of points  $(x_1, x_2) \in \mathbb{R}^2$  in the upper-half plane

$$x_2 > 0 \tag{4.5}$$

together with the metric

$$ds^{2} = \frac{1}{x_{2}^{2}} \left( dx_{1}^{2} + dx_{2}^{2} \right)$$
(4.6)

Although this model does not directly show up in our investigations in the next section, it turns out to very naturally connect the groups PSU(1,1) and  $PSL(2,\mathbb{R})$  that were discussed above. The connection comes from Möbius transformations, discussed in Section 4.3.

### 4.2 Representations in Hilbert Spaces

Informally, a *Hilbert space* is an infinite-dimensional vector space equipped with some inner product, and as such it requires a heavier mathematical apparatus than that for ordinary vector spaces. For the discussion in this thesis, though, we will take much of this theory for granted, in the same way that (say) the position basis is used without much hesitation in quantum mechanics textbooks like [16].

In the  $|j,m\rangle$  basis, any vector  $|\Psi\rangle$  can be written as a linear combination of the basis vectors in the following way:

$$|\Psi\rangle = \sum_{m} C_m |j,m\rangle \tag{4.7}$$

for some set of coefficients  $\{C_m\}$ , and where the summation is performed over either (all) the integers, the non-negative integers, or the non-positive integers, depending on which unitary irreducible representation is taken (the different cases will be discussed shortly). The inner product between two such vectors is then

$$\langle \Psi | \Psi' \rangle = \sum_{m} C_m^* C_m \tag{4.8}$$

and so for a vector  $|\Psi\rangle$  to have finite norm, the series  $\sum_m |C_m|^2$  must converge.

We now define a continuous basis for this vector space by

$$|\phi\rangle = \sum_{m} N_{m} e^{-im\phi} |j,m\rangle \tag{4.9}$$

where  $\phi$  is allowed to take real values. We can translate between the two bases by using some Fourier analysis. It is found that

$$|j,m\rangle = \frac{1}{2\pi N_m} \int_0^{2\pi} e^{im\phi} |\phi\rangle \, \mathrm{d}\phi \tag{4.10}$$

by noting that  $\frac{1}{2\pi} \int_0^{2\pi} e^{i(n-m)\phi} d\phi$  equals 0 when  $n \neq m$  and 1 when n = m. Therefore, in the  $|\phi\rangle$  basis, an arbitrary vector  $|\Psi\rangle$  is written

$$\begin{split} |\Psi\rangle &= \sum_{m} \frac{C_{m}}{2\pi N_{m}} \int_{0}^{2\pi} e^{im\phi} |\phi\rangle \, \mathrm{d}\phi \\ &= \frac{1}{2\pi} \int_{0}^{2\pi} \sum_{m} \frac{C_{m}}{N_{m}} e^{im\phi} |\phi\rangle \, \mathrm{d}\phi \\ &= \frac{1}{2\pi} \int_{0}^{2\pi} \psi(\phi) |\phi\rangle \, \mathrm{d}\phi \end{split}$$
(4.11)

where, in the last line, we defined the "wavefunction"

$$\psi(\phi) = \sum_{m} \frac{C_m}{N_m} e^{im\phi} \tag{4.12}$$

Given a wavefunction  $\psi$ , the coefficients  $C_m$  can also be found by Fourier analysis as

$$C_m = \frac{N_m}{2\pi} \int_0^{2\pi} e^{-im\phi} \psi(\phi) \,\mathrm{d}\phi \tag{4.13}$$

It is possible to have the operators  $J_{12}, M^+, M^-$  act on this wavefunction for a given representation, by defining e.g.  $M^+\psi$  such that

$$M^{+} |\Psi\rangle = \frac{1}{2\pi} \int_{0}^{2\pi} \left( M^{+} \psi(\phi) \right) |\phi\rangle \, \mathrm{d}\phi$$
(4.14)

and similarly for the other operators. Initially, we shall consider the principal and supplementary series, and so currently m runs from  $-\infty$  to  $\infty$  in all series. As an example, we compute  $M^+\psi$ , by first noting

$$M^{+} |\Psi\rangle = \sum_{m} C_{m} M^{+} |j,m\rangle = \sum_{m} C_{m} \cdot \frac{1}{\sqrt{2}} (j - E_{0} - m) \frac{N_{m}}{N_{m+1}} |j,m+1\rangle$$
(4.15)

If equation (4.10) is now inserted, it is found that

$$M^{+} |\Psi\rangle = \sum_{m} C_{m} \cdot \frac{1}{\sqrt{2}} \left(j - E_{0} - m\right) \frac{N_{m}}{N_{m+1}} \cdot \frac{1}{2\pi N_{m+1}} \int_{0}^{2\pi} e^{i(m+1)\phi} |\phi\rangle \, \mathrm{d}\phi$$
$$= \frac{1}{2\pi} \int_{0}^{2\pi} \left(\frac{1}{\sqrt{2}} e^{i\phi} \sum_{m} \frac{C_{m}}{N_{m}} \left(j - E_{0} - m\right) \left|\frac{N_{m}}{N_{m+1}}\right|^{2} e^{im\phi}\right) |\phi\rangle \, \mathrm{d}\phi$$
(4.16)

From the above, the transformed wavefunction  $M^+\psi$  can be read off. If we also insert the recursion relation (3.33) it is found that

$$M^{+}\psi(\phi) = \frac{1}{\sqrt{2}}e^{i\phi}\sum_{m}\frac{C_{m}}{N_{m}}\left(j - E_{0} - m\right)\left(-\frac{j + E_{0} + m + 1}{j - E_{0} - m}\right)e^{im\phi}$$
$$= \frac{1}{\sqrt{2}}e^{i\phi}\sum_{m}\frac{C_{m}}{N_{m}}\left(-j - E_{0} - m - 1\right)e^{im\phi}$$
$$= \frac{1}{\sqrt{2}}e^{i\phi}\left(-j - E_{0} - 1\right)\sum_{m}\frac{C_{m}}{N_{m}}e^{im\phi} - \frac{1}{\sqrt{2}}\sum_{m}\frac{C_{m}}{N_{m}}me^{im\phi}$$
(4.17)

The first series is recognised to be the wavefunction  $\psi(\phi)$  (cf. equation (4.12)). As for the latter series, observe that

$$\frac{\mathrm{d}\psi}{\mathrm{d}\phi}(\phi) = \frac{\mathrm{d}}{\mathrm{d}\phi} \left( \sum_{m} \frac{C_m}{N_m} e^{im\phi} \right) = \sum_{m} \frac{C_m}{N_m} \frac{\mathrm{d}}{\mathrm{d}\phi} \left( e^{im\phi} \right) = i \sum_{m} \frac{C_m}{N_m} m e^{im\phi}$$
(4.18)

Collecting these results,

$$M^{+}\psi(\phi) = \frac{1}{\sqrt{2}} \left(-j - E_{0} - 1\right) e^{i\phi}\psi(\phi) + \frac{i}{\sqrt{2}} e^{i\phi}\frac{\mathrm{d}\psi}{\mathrm{d}\phi}(\phi)$$
(4.19)

Finally, an application of the chain rule backwards yields

$$M^{+}\psi(\phi) = \frac{1}{\sqrt{2}} \left( -j - E_0 + i\frac{\mathrm{d}}{\mathrm{d}\phi} \right) e^{i\phi}\psi(\phi)$$
(4.20)

The other equations are found similarly, and the results are

$$J_{12}\psi(\phi) = \left(E_0 - i\frac{\mathrm{d}}{\mathrm{d}\phi}\right)\psi(\phi) \tag{4.21a}$$

$$M^{+}\psi(\phi) = \frac{1}{\sqrt{2}} \left( -j - E_0 + i\frac{\mathrm{d}}{\mathrm{d}\phi} \right) e^{i\phi}\psi(\phi)$$
(4.21b)

$$M^{-}\psi(\phi) = \frac{1}{\sqrt{2}} \left( -j + E_0 - i\frac{\mathrm{d}}{\mathrm{d}\phi} \right) e^{-i\phi}\psi(\phi)$$
(4.21c)

These are the reported equations (5.6) in [2].

Given two vectors  $|\Psi\rangle$  and  $|\Psi'\rangle$  with respective wavefunctions  $\psi$  and  $\psi'$ , their inner product can be computed from (4.11) as

$$\langle \Psi | \Psi' \rangle = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \psi^*(\phi) \psi'(\phi') \langle \phi | \phi' \rangle \, \mathrm{d}\phi \mathrm{d}\phi' \tag{4.22}$$

The inner product  $\langle \phi | \phi' \rangle$  is evaluated using (4.9):

$$\langle \phi | \phi' \rangle = \sum_{m} \sum_{m'} N_m^* N_{m'} e^{im\phi} e^{-im'\phi} \langle j, m | j, m' \rangle$$

$$= \sum_{m} \sum_{m'} N_m^* N_{m'} e^{-im\phi} e^{im'\phi} \delta_{mm'}$$

$$= \sum_{m} |N_m|^2 e^{im(\phi - \phi')}$$

$$(4.23)$$

We now recall that for the principal series  $N_m = 1$  identically, and so

$$\langle \phi | \phi' \rangle = \sum_{m} e^{im\phi - \phi'} = 2\pi \delta(\phi - \phi') \tag{4.24}$$

where  $\delta$  is the Dirac delta function. Therefore,

$$\langle \Psi | \Psi' \rangle = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \psi^*(\phi) \psi'(\phi') \cdot 2\pi \delta(\phi - \phi') \, \mathrm{d}\phi \mathrm{d}\phi'$$
  
=  $\frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} \psi^*(\phi) \psi'(\phi) \, \mathrm{d}\phi$  (4.25)

In particular, requiring any vector  $|\Psi\rangle$  to have finite norm, we impose

$$\langle \Psi | \Psi \rangle = \frac{1}{2\pi} \int_0^{2\pi} |\psi(\phi)|^2 \, \mathrm{d}\phi < \infty \tag{4.26}$$

Hence, for the principal series the Hilbert space is the space of square-integrable functions on the circle.

A similar analysis can be performed for the supplementary series, and also for the discrete series, but for the latter case the derivation is trickier because  $\phi$  must be allowed to take complex values. A discussion on these cases is provided in Bargmann's paper [1], and it is found that the discrete series is realised on the Hilbert space of analytic functions on the unit disk |z| < 1 in the complex plane.

From these investigations we have again encountered a unit circle and a unit disk, which are meaningful having considered the hyperbolic plane and the projective lightcone above. In fact, with this interpretation we can consider the respective Hilbert spaces to be the analytic, squareintegrable functions on the hyperbolic plane and on the projective lightcone themselves. What we have then obtained is an orthonormal basis for these sets of functions, in the sense that any such function on the space of interest is a linear combination of the basis functions on that space. Therefore, we have found the analogues of the spherical harmonics on the 2-sphere, and another connection to the rotation group has been established.

#### 4.3 Möbius Transformations

Möbius transformations are a powerful tool to study geometry on the complex plane, and in particular on what is known as the *Riemann sphere*. In addition, they will give some geometrical intuition to matrix groups we have already discussed. Firstly, we define the Riemann sphere, and its interpretation in terms of projections.

**Definition 4.4.** The **Riemann sphere** is the union of the complex plane with the point at complex infinity, denoted by  $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ . Accordingly, sequences that diverge in magnitude in the complex plane are said to converge to the point  $\infty$  on the Riemann sphere.

The reason it is called a sphere is seen from the following visualisation. Consider the unit 2sphere embedded in  $\mathbb{R}^3$  whose center lies on the origin, i.e. the surface  $X_1^2 + X_2^2 + X_3^2 = 1$ . We call the point  $(X_1, X_2, X_3) = (0, 0, 1)$  the "north pole", and similarly (0, 0, -1) the "south pole". Next, we visualise the plane  $X_3 = -1$  tangent to the sphere at the south pole. Then, any line segment drawn from a point on the plane to the north pole will intersect the sphere in a unique point, and furthermore every point on the sphere is hit by one such a line segment, except for the north pole itself. If we view the plane projected from as the complex plane, this gives a mapping from  $\mathbb{C}$  to the unit sphere minus the north pole. Therefore, if we furthermore define the point  $\infty$  to be mapped to the north pole, this gives a bijective mapping between the Riemann sphere and the unit two-dimensional sphere. This motivates the naming of the Riemann sphere. For a diagrammatic overview of this projection, called the *stereographic projection*, see Figure 2.



Figure 2: A diagram depicting the projection from the Riemann sphere onto the 2-sphere. Drawn is the sphere  $X_1^2 + X_2^2 + X_3^2 = 1$  and part of the plane  $X_3 = -1$ . The point **b** in the plane is mapped by a straight line, connecting **b** with the "north pole"  $(X_1, X_2, X_3) = (0, 0, 1)$ , to the point **b**' on the 2-sphere. The plane should be visualised as extending infinitely in its directions.

It is reasonable to ask what the group of symmetries on the Riemann sphere looks like, and it turns out that these are given by a set of functions called *Möbius transformations*.

**Definition 4.5.** A Möbius transformation  $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$  is a fractional linear transformation of the form

$$f(z) = \frac{az+b}{cz+d} \tag{4.27}$$

with  $a, b, c, d \in \mathbb{C}$  such that  $ad - bc \neq 0$ .

The group of Möbius transformations on  $\hat{\mathbb{C}}$  is denoted  $\operatorname{Aut}(\hat{\mathbb{C}})$ , the group of automorphisms on the Riemann sphere. These transformations have many interesting properties. Firstly, they are analytic on the Riemann sphere, and this will be shown below.

Consider a Möbius transformation of the above form. Since it is the ratio of two differentiable functions, it is differentiable on  $\mathbb{C}$  whenever the numerator does not equal zero. For such points we find that

$$f'(z) = \frac{a(cz+d) - (az+b)c}{(cz+d)^2} = \frac{ad-bc}{(cz+d)^2}$$
(4.28)

If c = 0, f' is constant and so f is differentiable on  $\mathbb{C}$ . If not, the derivative f' goes to complex infinity at the point  $z = -\frac{d}{c}$ , and so it actually converges to  $\infty$  in  $\hat{\mathbb{C}}$ . Since we consider f as a function of the Riemann sphere into itself, it is necessarily differentiable everywhere on  $\mathbb{C}$ . It remains to show differentiability at  $\infty$ , and the standard procedure is to consider the function  $f(\frac{1}{z})$ and show that its derivative is well-defined around the origin. We find

$$f'\left(\frac{1}{z}\right) = \frac{-\frac{a}{z^2}\left(\frac{c}{z}+d\right) - \left(\frac{a}{z}+b\right)\left(-\frac{c}{z^2}\right)}{\left(\frac{c}{z}+d\right)^2} = -\frac{ad-bc}{(c+dz)^2}$$
(4.29)

If  $d \neq 0$ , investigating a disk of radius smaller than  $\left|\frac{c}{d}\right|$  shows that f is differentiable around infinity. If d = 0, the derivative is constant (and so well-defined around infinity). So f is analytic on the Riemann sphere.

It is a well-known result of complex analysis (see e.g. theorem 3.4 in [5]) that an analytic function f is angle-preserving, or *conformal*, everywhere on its domain where  $f' \neq 0$ . Hence, all Möbius transformations are everywhere conformal, since  $ad - bc \neq 0$  by definition.

There is a beautiful connection between subgroups of the group of Möbius transformations and the matrix groups we have defined. Consider the mapping  $\pi : GL(2, \mathbb{C}) \to \operatorname{Aut}(\hat{\mathbb{C}})$  defined by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto f_{abcd} \tag{4.30}$$

where  $f_{abcd}$  is the Möbius transformation such that  $f_{abcd}(z) = \frac{az+b}{cz+d}$ . Noting that the same Möbius transformation is obtained if a, b, c, d are all re-scaled by a non-zero complex number, the coefficients can always be chosen such that det  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1$ , and so we can consider the domain of the mapping to be  $SL(2, \mathbb{C})$  instead. Even so, there are exactly two matrices (differing by a sign) that map to each given transformation, and the domain can be restricted further to  $PSL(2, \mathbb{C})$ . The resulting function  $\pi : PSL(2, \mathbb{C}) \to \operatorname{Aut}(\hat{\mathbb{C}})$  is then a bijection. Moreover, it can be shown to define a group homomorphism, and it follows that  $\operatorname{Aut}(\hat{\mathbb{C}}) \cong PSL(2, \mathbb{C})$ .

With this result, we can consider subgroups of  $PSL(2, \mathbb{C})$  and interpret them as groups of Möbius transformations. For instance, it is seen that matrices in  $PSL(2, \mathbb{R})$  leave the real line invariant, and so the upper-half plane (and the lower-half plane) is left invariant. Therefore, if we view the upper-half plane model as embedded in  $\mathbb{C}$  instead of  $\mathbb{R}^2$ , then  $PSL(2, \mathbb{R})$  is the symmetry group on the hyperbolic plane.

Also interesting is the case of PSU(1,1), for which the induced Möbius transformations take the form

$$f(z) = \frac{\alpha z + \beta}{\beta^* z + \alpha^*} \tag{4.31}$$

These transformations will, in fact, preserve the unit circle  $|z|^2 = 1$  in the complex plane. To see this, we note that

$$|f(z)|^{2} = \frac{|\alpha z + \beta|^{2}}{|\beta^{*}z + \alpha^{*}|^{2}} = \frac{|\alpha|^{2} + \alpha z\beta^{*} + \alpha^{*}z^{*}\beta + |\beta|^{2}}{|\beta|^{2} + \alpha z\beta^{*} + \alpha^{*}z^{*}\beta + |\alpha|^{2}} = 1$$
(4.32)

where, in the middle step, we used the fact that |z| = 1. By extension, Möbius transformations with coefficients from PSU(1, 1) map the unit disk |z| < 1 to itself. Therefore, if we in the same way view the Poincaré disk model as embedded in  $\mathbb{C}$ , then PSU(1, 1) is its symmetry group. With this and the above result for  $PSL(2, \mathbb{R})$ , it becomes immediately evident that the two groups must be

isomorphic, the proof of which we established in Section 2.2. In fact, we can even give a geometric interpretation of the matrix T, defined in equation (2.4), that was used in this proof. Viewed as a Möbius transformation, the matrix T takes the real line to the unit circle |z| = 1, and the upper-half plane  $\Im(z) > 0$  to the unit disk |z| < 1. To see this, we observe that for z = x + iy with x real and y > 0,

$$|f(x+iy)|^{2} = \frac{|x+iy-i|^{2}}{|-ix+y+1|^{2}} = \frac{x^{2}+(y-1)^{2}}{x^{2}+(y+1)^{2}} < 1$$
(4.33)

and so T acts as a chart between the two models of the hyperbolic plane.

## 5 The Discrete Series Represented on Anti-de Sitter Space

In this section, we explore another setting in which the class of representations  $D^+(j)$ , the discrete series, that was found above takes on a very natural physical meaning. We only need a few more concepts from differential geometry, and the definition of 1+1-dimensional anti-de Sitter space, to see this.

#### 5.1 Lie Derivatives and Killing Vector Fields

A *Killing vector field* is a vector field on some spacetime that describes the flow of an isometry of that spacetime, in a way that will be made precise below. The study of such vector fields will help us obtain a physical intuition for some of the representations found above.

Before giving the precise definition of a Killing vector field, the notion of a *Lie derivative* is needed. Intuitively, the Lie derivative measures the initial rate of change of an object (e.g. a function, vector field, tensor field) as points on the spacetime move in the direction of a given vector field. More formally, we have the following definition, taken from Huggett & Tod [12].

**Definition 5.1.** Let  $\mathbf{V} = V^{\alpha} \partial_{\alpha}$  and  $\mathbf{W} = W^{\alpha} \partial_{\alpha}$  be tangent vector fields on some spacetime S, and let f be a function on S. Then, the **Lie derivative** of f in the direction of  $\mathbf{V}$  is

$$\mathcal{L}_{\mathbf{V}}f = V^{\alpha}\partial_{\alpha}f \tag{5.1}$$

Similarly, the **Lie derivative** of **W** in the direction of **V** is

$$\mathcal{L}_{\mathbf{V}}W^{\alpha} = V^{\beta}\partial_{\beta}W^{\alpha} - W^{\beta}\partial_{\beta}V^{\alpha} \tag{5.2}$$

Observing the antisymmetry in  $\mathbf{V}$  and  $\mathbf{W}$  in the definition of the Lie derivative, it makes sense to define the *commutator* of two tangent vector fields to be the new vector field whose components are

$$[\mathbf{V},\mathbf{W}]^{\alpha} = V^{\beta}\partial_{\beta}W^{\alpha} - W^{\beta}\partial_{\beta}V^{\alpha} \quad (=\mathcal{L}_{\mathbf{V}}W^{\alpha} = -\mathcal{L}_{\mathbf{W}}V^{\alpha}) \tag{5.3}$$

By requiring the Lie derivative to be linear and satisfy leibnizian properties, this derivative can be extended uniquely to arbitrary tensor fields on a spacetime. Also, for those familiar with differential geometry, the partial derivatives could all be replaced by covariant derivatives without altering the definitions, as is done in Huggett & Tod [12] (this can be done since all the arising Christoffel symbols cancel). However, to avoid a discussion on more differential geometry than is needed, and also to stress that the Lie derivative can be defined even in the absence of a metric, we shall use the partial derivatives instead.

In particular, consider what happens when the Lie derivative is applied to the *metric*  $g_{\alpha\beta}$  of a spacetime. For some tangent vector field **V**, we have

$$\mathcal{L}_{\mathbf{V}}g_{\alpha\beta} = V^{\gamma}\partial_{\gamma}g_{\alpha\beta} + g_{\alpha\gamma}\partial_{\beta}V^{\gamma} + g_{\gamma\beta}\partial_{\alpha}V^{\gamma}$$
(5.4)

An interesting special case arises, when such a derivative vanishes identically.

**Definition 5.2.** Let  $g_{\alpha\beta}$  be a metric on a spacetime S, and  $\mathbf{V} = V^{\alpha}\partial_{\alpha}$  a tangent vector field on S. Then,  $\mathbf{V}$  is a **Killing vector field** provided

$$\mathcal{L}_{\mathbf{V}}g_{\alpha\beta} = 0 \tag{5.5}$$

Thus, a Killing vector field has the property that spacetime distances and angles are unchanged as points in the spacetime move in the direction of the vector field, and so Killing vector fields describe symmetries (isometries) on a spacetime.

### 5.2 Anti-de Sitter Space and the Discrete Series

There is one more space left for us to define which, like the hyperboloid model, is most easily visualised as embedded in Minkowski space.

**Definition 5.3.** The hyperboloid model for 1+1-dimensional anti-de Sitter space  $AdS_2$  is the hyperboloid

$$X_1^2 + X_2^2 - X_3^2 = 1 (5.6)$$

together with the (standard Minkowski) metric

$$\mathrm{d}s^2 = \mathrm{d}X_1^2 + \mathrm{d}X_2^2 - \mathrm{d}X_3^2 \tag{5.7}$$

In this space, one direction is periodic while the other is not, and since we are free to swap the meaning of space and time in 1+1 dimensions, the periodicity can be ascribed to either coordinate. However, in accordance with how higher-dimensional anti-de Sitter spaces are defined, time is interpreted to be the periodic coordinate. (Actually, instead interpreting the periodic direction as the spatial direction gives a model of 1+1-dimensional *de Sitter space*,  $dS_2$ , but the close relation between anti-de Sitter space and de Sitter space is lost in higher dimensions. Also, the interpretation of  $D^+(j)$  made below cannot be made for the case of de Sitter space.)

We can construct three linearly independent Killing vector fields on the hyperboloid (such that any Killing vector field is a linear combination of the three) by

$$\mathbf{J}_{12} = -X_2\partial_1 + X_1\partial_2 \tag{5.8a}$$

$$\mathbf{J}_{13} = X_3 \partial_1 + X_1 \partial_3 \tag{5.8b}$$

$$\mathbf{J}_{23} = X_3 \partial_2 + X_2 \partial_3 \tag{5.8c}$$

These preserve  $AdS_2$  in the sense that  $\mathcal{L}_{\mathbf{J}_{\alpha\beta}}\left(X_1^2 + X_2^2 - X_3^2\right) = 0$ . For instance,

$$\mathcal{L}_{\mathbf{J}_{13}}\left(X_1^2 + X_2^2 - X_3^2\right) = J_{13}^\beta \partial_\beta \left(X_1^2 + X_2^2 - X_3^2\right) = 2X_3X_1 - 2X_1X_3 = 0$$
(5.9)

Specifically,  $\mathbf{J}_{12}$  corresponds to a rotation of the hyperboloid, and the other two give Lorentz boosts in the respective directions. It is no coincidence that we have used the symbols  $\mathbf{J}_{\alpha\beta}$  to describe these vector fields; indeed, they satisfy the commutation relations

$$[\mathbf{J}_{13}, \mathbf{J}_{12}] = \mathbf{J}_{23} \tag{5.10a}$$

$$[\mathbf{J}_{12}, \mathbf{J}_{23}] = \mathbf{J}_{13} \tag{5.10b}$$

$$[\mathbf{J}_{13}, \mathbf{J}_{23}] = -\mathbf{J}_{12} \tag{5.10c}$$

As an example, we compute the components of  $[\mathbf{J}_{12}, \mathbf{J}_{23}]$ .

$$[\mathbf{J}_{12}, \mathbf{J}_{23}]^1 = J_{12}^\beta \partial_\beta J_{23}^1 - J_{23}^\beta \partial_\beta J_{12}^1 = X_3$$
(5.11a)

$$\left[\mathbf{J}_{12}, \mathbf{J}_{23}\right]^2 = J_{12}^\beta \partial_\beta J_{23}^2 - J_{23}^\beta \partial_\beta J_{12}^2 = 0$$
(5.11b)

$$[\mathbf{J}_{12}, \mathbf{J}_{23}]^3 = J_{12}^\beta \partial_\beta J_{23}^3 - J_{23}^\beta \partial_\beta J_{12}^3 = X_1$$
(5.11c)

and so  $[\mathbf{J}_{12}, \mathbf{J}_{23}] = \mathbf{J}_{13}$ . Therefore, the Killing vector fields  $\mathbf{J}_{\alpha\beta}$  give a representation for the algebra of  $SO^+(2, 1)$ . Now, the observant reader will have noticed that there is a factor *i* missing between equations (3.3a)-(3.3c) and equations (5.10a)-(5.10c). This is not an error in our calculations; it comes from how the Lie algebra is defined. Instead of defining the algebra as the elements X such that  $e^{itX}$  is an element of the group for all real *t*, one could equivalently define the algebra as the elements such that  $e^{tX}$  is in the group for all real *t*. This is why we can still rightfully say that the vector fields  $\mathbf{J}_{\alpha\beta}$  represent a basis for the Lie algebra of the Lorentz group.

More conveniently, there exists a coordinate system for  $AdS_2$  in which both the temporal and the spatial coordinate take finite values, called "sausage coordinates". (The name has to do with the analogous coordinates for 2+1 anti-de Sitter space, for which the coordinate space informally looks like a sausage.) These coordinates are discussed in e.g. Holst [11], and are defined by the relations

$$X_1 = \frac{1+x^2}{1-x^2}\cos(t)$$
(5.12a)

$$X_2 = \frac{1+x^2}{1-x^2}\sin(t)$$
(5.12b)

$$X_3 = \frac{2x}{1 - x^2} \tag{5.12c}$$

In this space it is seen explicitly that the temporal coordinate is periodic, and this is slightly problematic but not a major hindrance. When it is important that time be non-periodic, we can mathematically go to the *covering space* where this is fulfilled (cf. Definition 2.12). Also, the induced metric in sausage coordinates can be computed, and is done in e.g. Bengtsson *et al.* [3]. We can make a formal definition.

**Definition 5.4.** The sausage model for 1+1-dimensional anti-de Sitter space,  $AdS_2$ , is the set of points  $(x, t) \in \mathbb{R}^2$  in the rectangle  $-1 < x < 1, 0 \le t \le 2\pi$  together with the metric

$$ds^{2} = -\left(\frac{1+x^{2}}{1-x^{2}}\right)^{2} dt^{2} + \frac{4}{\left(1+x^{2}\right)^{2}} dx^{2}$$
(5.13)

where the points  $(x, 2\pi)$  are identified with the points (x, 0).

Like previous models of spacetimes, this model has the convenient property that all points are represented by finite coordinates. It does have the drawback that lightcones are not given by 45 degree lines, but nonetheless it will suffice for the discussion to follow.

We now wish to write the vector fields  $\mathbf{J}_{\alpha\beta}$  in the sausage model, and one of them is easily found since, by the chain rule,

$$\partial_t = \frac{\partial t}{\partial X_\alpha} \partial_\alpha = -X_2 \partial_1 + X_1 \partial_2 = \mathbf{J}_{12}$$
(5.14)

Thus  $\mathbf{J}_{12}$  precisely generates time translation, and the time translational symmetry of  $AdS_2$  is evident.

Before writing down the two other vector fields, the coordinate relations (5.12a)-(5.12c) will be inverted. We note that x is a function of  $X_3$  alone, and that

$$\tan(t) = \frac{X_2}{X_1} \tag{5.15}$$

so that

$$t = \operatorname{atan2}\left(X_2, X_1\right) \tag{5.16}$$

where atan2 is the function such that if  $t = \operatorname{atan2}(X_2, X_1)$ , then  $\operatorname{tan}(t) = \frac{X_2}{X_1}$  and  $X_2 = \sqrt{X_1^2 + X_2^2} \sin(t)$ . That is,  $\operatorname{atan2}(X_2, X_1)$  is similar to  $\operatorname{arctan}\left(\frac{X_2}{X_1}\right)$ , but it also uses the signs of the two arguments to return the correct value in  $0 \le t \le 2\pi$ , as opposed to the inverse tangent function whose codomain is e.g.  $-\frac{\pi}{2} \le t \le \frac{\pi}{2}$ .

We also need some partial derivatives of the coordinate relations. We find

$$\frac{\partial t}{\partial X_1} = \frac{1}{1 + \frac{X_2^2}{X_1^2}} \cdot \left(-\frac{X_2}{X_1^2}\right) = -\frac{X_2}{X_1^2 + X_2^2}$$
(5.17a)

$$\frac{\partial t}{\partial X_2} = \frac{1}{1 + \frac{X_2^2}{X_1^2}} \cdot \frac{1}{X_1} = \frac{X_1}{X_1^2 + X_2^2}$$
(5.17b)

Moreover, by a standard result from single variable calculus,

$$\frac{\partial x}{\partial X_3} = \frac{\mathrm{d}x}{\mathrm{d}X_3} = \frac{1}{\frac{\mathrm{d}X_3}{\mathrm{d}x}} = \left(\frac{2\left(1-x^2\right)-2x(-2x)}{\left(1-x^2\right)^2}\right)^{-1}$$
$$= \frac{\left(1-x^2\right)^2}{2\left(1+x^2\right)} \tag{5.18}$$

Therefore, we can find  $\mathbf{J}_{13}$  and  $\mathbf{J}_{23}$  by similar chain rule considerations. As an example,  $\mathbf{J}_{13}$  is computed:

$$\mathbf{J}_{13} = X_3\partial_1 + X_1\partial_3 = X_3\left(\frac{\partial x}{\partial X_1}\partial_x + \frac{\partial t}{\partial X_1}\partial_t\right) + X_1\left(\frac{\partial x}{\partial X_3}\partial_x + \frac{\partial t}{\partial X_3}\partial_t\right)$$
$$= \frac{2x}{1-x^2} \cdot \left(-\frac{1-x^2}{1+x^2}\sin(t)\right)\partial_t + \frac{1+x^2}{1-x^2}\cos(t) \cdot \frac{\left(1-x^2\right)^2}{2\left(1+x^2\right)}\partial_x$$
$$= \frac{1}{2}\left(1-x^2\right)\cos(t)\partial_x - \frac{2x}{1+x^2}\sin(t)\partial_t \tag{5.19}$$

The calculation for  $\mathbf{J}_{23}$  is comparable, and one finds

$$\mathbf{J}_{23} = \frac{1}{2} \left( 1 - x^2 \right) \sin(t) \partial_x + \frac{2x}{1 + x^2} \cos(t) \partial_t \tag{5.20}$$

We return, now, to a physical application of the representations found in previous sections. It was observed that  $\mathbf{J}_{12}$  corresponds to infinitesimal time translation in the sausage model for  $AdS_2$ . This gives a new interpretation of the spectrum of  $J_{12}$  as the possible quantum mechanical energies of a single particle in this space. In any reasonable quantum theory, these energies must be bounded from below, and therefore the discrete series representation  $D^+(j)$  is singled out as the physically interesting one here. Indeed, this is the reason why it is this class of unitary irreducible representations that is studied in e.g. Fronsdal [7] and Nicolai [15]. Finally, for this class of representations the lowest allowable energy is  $E_0 = -j$ , which motivates the choice of notation.

# 6 Conclusions and Further Discussion

In this thesis, we have discussed the Lorentz group in 2+1 dimensions and its representations, with the over-arching goal of comparing to and contrasting with the more familiar rotation group. In this pursuit, we first saw a number of matrix groups that initially seemed different, but were later seen to be related in various ways. We went on to find and classify all linear, and then all unitary, irreducible representations of both the rotation group and the Lorentz group. It was found that while all the unitary irreducible representations of the rotation group are finite-dimensional, none of those for the Lorentz group are. After this, we saw that the representations of the Lorentz group can be viewed as the set of analytical, square-integrable functions on suitable spaces, in analogy with the spherical harmonics defined on the 2-sphere. Finally, a particular class of representations was found to be relevant for quantum theories in Anti-de Sitter space. Along the way, a number of mathematical concepts were introduced, to support these arguments.

As it stands, this thesis has in many ways been a mathematical endeavor. That is not to say that mathematics is something that should be used sparingly, but there are many questions of physical nature that can be raised from this thesis. Firstly, how does the approach presented in this thesis generalise to higher dimensions? What new difficulties arise, and what new physics can be done in these higher dimensions?

Also, how do the methods developed apply to symmetry groups of spacetimes other than Minkowski? For instance, what can be found for SO(n, 2) with n > 1, the symmetry groups of higher-dimensional anti-de Sitter spaces? Secondly, given the relevance that the representations for the Lorentz group seems to have for quantum theories in curved spacetimes, one might ask what such a theory actually looks like. With the incorporation of relativity, the Schrödinger equation can no longer be the governing equation, since it treats temporal and spatial derivatives differently.

Finally, the connection between SU(2), SU(1,1) and quantum optics, as investigated in Yurke *et al.* [18], would be interesting to study in detail. Due to the time constraints of this thesis, there was unfortunately not enough time to pursue this direction very far. These are all questions that would be interesting to pursue in a future study.

# Acknowledgements

There are many people who helped make this thesis possible. Firstly, I would like to thank my supervisor, Ingemar Bengtsson, for his guidance in researching the topic, as well as writing the final thesis. I am grateful for his willingness to discuss with me whenever I had doubts or questions, as well as taking the time to proof-read many of the drafts for this thesis. Secondly, I am thankful for my family, and for their love and support.

# Bibliography

- V. Bargmann. Irreducible unitary representations of the lorentz group. Annals of Mathematics, 48(3):568-640, 1947. ISSN 0003486X. URL http://www.jstor.org/stable/1969129.
- [2] A. O. Barut and C. Fronsdal. On non-compact groups. II. Representations of the 2+1 Lorentz group. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 287(1411):532-548, 1965.
- [3] I. Bengtsson, S. Holst, and E. Jakobsson. Classics illustrated: limits of spacetimes. Classical and Quantum Gravity, 31(20):205008, 2014.

- [4] J. W. Cannon, W. J. Floyd, R. Kenyon, and W. R. Parry. Hyperbolic geometry. In S. Levy, editor, *Flavors of geometry*, pages 59–115. Cambridge University Press, 1997.
- [5] J. B. Conway. Functions of One Complex Variable I, pages 33–57. Graduate Texts in Mathematics, Springer Verlag, 2nd edition, 1978.
- [6] D. Dummit and R. Foote. Abstract algebra, pages 16–117 and pages 840–847. Wiley Hoboken, 3rd edition, 2004.
- [7] C. Fronsdal. Elementary Particles in a Curved Space. Reviews of Modern Physics, 37(1): 221–224, 1965.
- [8] H. Goldstein, C. Poole, and J. Safko. *Classical mechanics*, pages 139–151. American Association of Physics Teachers, 3rd edition, 2002.
- [9] B. C. Hall. Lie Groups, Lie Algebras, and Representations, pages 3–43. Graduate Texts in Mathematics, Springer Verlag, 2003.
- [10] D. Han and Y. S. Kim. Special relativity and interferometers. *Physical Review A*, 37(11): 4494–4496, 1988.
- S. Holst. Horizons and time machines global structures in locally trivial spacetimes. PhD thesis, Stockholm University, 2000.
- [12] S. A. Huggett and K. P. Tod. An introduction to twistor theory, pages 5–9. Cambridge University Press, 2nd edition, 1994.
- [13] G. Lindblad and B. Nagel. Continuous bases for unitary irreducible representations of SU(1,1). Annales de l'I.H.P. Physique théorique, 13(1):27-56, 1970. URL http://www.numdam.org/ item/AIHPA\_1970\_13\_1\_27\_0.
- [14] J. Maldacena. The Large N limit of Superconformal field theories and supergravity. International journal of theoretical physics, 38(4):1113–1133, 1999.
- [15] H. Nicolai. Representations of Supersymmetry in Anti-de Sitter Space. In Spring School on Supergravity and Supersymmetry, pages 368–399, 4 1984.
- [16] J. J. Sakurai and J. Napolitano. Modern quantum mechanics, pages 172–196. Pearson Harlow, 2nd edition, 2014.
- [17] E. Wigner. On unitary representations of the inhomogeneous lorentz group. Annals of Mathematics, 40(1):149-204, 1939. ISSN 0003486X. URL http://www.jstor.org/stable/1968551.
- [18] B. Yurke, S. McCall, and J. Klauder. SU(2) and SU(1,1) interferometers. *Physical Review* A, 33(6):4033–4054, 1986.

# Appendix: Review of Group Theory

## A.1: What Is a Group?

Basic results of group theory are ubiquitous to this thesis, and so the author chose to append a review of the most important definitions and results. However, as an introduction to the subject it will certainly be lacking, and the reader may wish to consult standard mathematics textbooks for this, e.g. Dummit & Foote [6]. This subsection is largely written following this book.

So we start with the definition of a group.

**Definition 6.1.** A set *G* equipped with a function  $\star : G \times G \to G$  is called a **group** if the following axioms hold (we write  $a \star b \equiv \star(a, b)$  for all  $a, b \in G$ ):

- 1. Associativity: the relation  $a \star (b \star c) = (a \star b) \star c$  holds for all  $a, b, c \in G$ .
- 2. Identity: there exists an element  $e \in G$  such that  $e \star a = a \star e = a$  for all  $a \in G$ .
- 3. Inverses: for each  $a \in G$  there exists an element  $a^{-1} \in G$ , called the inverse of a, such that  $aa^{-1} = a^{-1}a = e$ .

When it is clear what the operation  $\star$  is from the context, we shall write  $ab = a \star b$  for brevity. This definition may sound vague, but therein lies its power, since the abstract concept of a group allows for proofs of properties of a very large class of objects.

Familiar examples of groups include the integers  $\mathbb{Z}$  under addition, the non-zero complex numbers  $\mathbb{C} \setminus \{0\}$  under multiplication, and many more. These groups both have the property that ab = ba for all elements in the group, which is the defining property of an *abelian group*. Groups also frequently appear as symmetries of mathematical objects, as is the case for the Lorentz group. For such cases, the group elements are the symmetry transformations, and the group operation is composition of transformations.

In a similar fashion, a subgroup H of a group G is a subset of G such that the group properties hold for H, and we have closure, which is to say that  $hg \in H$  for all  $h, g \in H$ . In this way we see, for instance, that the rational numbers  $\mathbb{Q}$  for a subgroup of the reals  $\mathbb{R}$  under addition.

Next, we need a notion of when two groups are "equal", and for this, we define a group isomorphism.

**Definition 6.2.** Let G and H be two groups with respective operations  $\cdot$  and  $\star$ . A function  $\varphi: G \to H$  is called a **group isomorphism** if it is a bijection and

$$\varphi\left(g_1 \cdot g_2\right) = \varphi\left(g_1\right) \star \varphi\left(g_2\right) \tag{6.1}$$

for all  $g_1, g_2 \in G$ . If such a function exists, we say that G and H are *isomorphic*, and write  $G \cong H$ .

A function that satisfies (6.1) (bijective or not) is called a group homomorphism. The set of elements in G that map to the identity in H is called the kernel of  $\varphi$ , denoted ker  $\varphi$ , while the image of  $\varphi$ , written im  $\varphi$ , is the set  $\varphi(G) = \{\varphi(g) \mid g \in G\}$ .

We now need two more notions from group theory, namely normal subgroups and quotient groups. A normal subgroup N of some group G is a subgroup such that

$$gNg^{-1} = \{gng^{-1} \mid n \in N\} = N$$
(6.2)

for all  $g \in G$ . Given such a normal subgroup, we define the *quotient group* G/N (read "G mod N") as the set

$$G/N = \{gN \mid g \in G\} \tag{6.3}$$

where  $gN = \{gn \mid n \in N\}$ . It can be shown that the operation

$$(gN)(hN) = (gh)N \tag{6.4}$$

is well-defined, and this turns the set G/N into a group, which is why it deserves its name. The requirement that N is normal is needed, because only then is the above operation well-defined.

As a concluding remark on group theory, we shall state a famous result without proof. The interested reader is again referred to [6].

**Theorem 6.1.** The first isomorphism theorem. Let  $\phi : G \to H$  be a group homomorphism. Then

$$G/\ker\varphi \cong \operatorname{im}\varphi \tag{6.5}$$

In particular, if  $\varphi$  is surjective, then  $G/\ker \varphi \cong H$ .

Finally, for completeness we shall formally define a *field*.

**Definition 6.3.** A field  $\mathbb{F}$  is a set equipped with two operations + and  $\cdot$ , called addition and multiplication respectively, with respective identities 0 and 1 such that  $0 \neq 1$ . Also, the following properties must all hold:

- $\mathbb{F}$  is an abelian group under addition.
- $\mathbb{F} \setminus \{0\}$  is an abelian group under multiplication.
- The distributive law holds, so for all  $a, b, c \in \mathbb{F}$ ,

$$a \cdot (b+c) = (a \cdot b) + (a \cdot c) \tag{6.6}$$

Most of the familiar number systems are fields, like the rationals  $\mathbb{Q}$ , the reals  $\mathbb{R}$  and the complex numbers  $\mathbb{C}$ .

#### A.2: Lie Groups and Lie Algebras

Groups come in all varieties, both finite and infinite. Among the infinite ones, some further classifications can be made. Of special importance, both mathematically and physically, are *Lie groups*. They are groups that also have a manifold structure, and so they are in some sense smooth. As such, the general theory of Lie groups require a reasonable amount of topology and differential geometry. Fortunately for us, though, we will only work with matrix groups, for which the theory can be substantially simplified.

The exposition in this subsection will closely follow the textbook of Hall [9], which is an approachable treatment for physics students, too. Before defining a matrix Lie group, we need a definition of convergence of matrices in  $M_n(\mathbb{C})$ , the set of  $n \times n$  complex matrices.

**Definition 6.4.** Let  $A_m$  be a sequence of complex matrices in  $M_n(\mathbb{C})$ . Then  $A_m$  converges to a matrix  $A \in M_n(\mathbb{C})$  if each entry in  $A_m$  converges to the corresponding entry in A, as  $m \to \infty$ . That is to say, each sequence of complex numbers  $(A_m)_{ij}$  converges to  $A_{ij}$ , for all  $1 \le i, j \le n$ .

We can now state the definition of matrix Lie groups, which indeed are Lie groups [9].

**Definition 6.5.** A matrix Lie group G is any subgroup of  $GL(n, \mathbb{C})$  such that if  $A_m$  is a sequence in G that converges to a matrix  $A \in M_n(\mathbb{C})$ , then either  $A \in G$  or A is not invertible. Between different matrix Lie groups, there is one more distinction that is interesting to make, which is that of *compactness*. It is essentially compactness that sets the Lorentz group apart from the rotation group.

**Definition 6.6.** A matrix Lie group G is said to be **compact** if it holds that

- If  $A_m$  is a sequence of matrices in G that converges to a matrix  $A \in M_n(\mathbb{C})$ , then  $A \in G$ .
- There exists a positive real number C such that for all  $A \in G$ ,  $|A_{ij}| < C$  for all  $1 \le i, j \le n$ .

The above definition is essentially a statement of the Heine-Borel theorem, as typically seen in a real analysis course, which states that a subset of  $\mathbb{R}^n$  is compact if and only if it is closed and bounded.

There is one thing left to define in this section, and that is a *Lie algebra*. Intuitively, the Lie algebra of a Lie group generates the group by exponentiation, a sentence that will be made clear shortly. Firstly, we define exponentiation of a matrix just like the usual power series for the exponential function. For any  $n \times n$  matrix X, we make the definition

$$e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k \tag{6.7}$$

It is a fact [9] that such a power series converges for all such matrices X. We can now define a Lie algebra, a tool that will help us study the corresponding groups.

**Definition 6.7.** Let G be a matrix Lie group. Then, the Lie algebra of G, written  $\mathfrak{g}$ , is the set of all matrices X such that  $e^{itX} \in G$  for all real numbers t.

In general, an *algebra* in mathematics is a vector space V equipped with a function  $[\cdot, \cdot] : V \times V \rightarrow V$  called a *commutator*, satisfying a few properties. Namely, for all  $u, v, w \in V$ , we require linearity in both arguments, antisymmetry (i.e. [u, v] = -[v, u]), and that the Jacobi identity holds, which is to say that

$$[u, [v, w]] + [w, [u, v]] + [v, [w, u]] = 0$$
(6.8)

It should be noted that the Jacobi identity implies non-associativity. Knowing this, the name Lie *algebra* is fitting because the commutator can, up to a constant factor, be taken to be the usual matrix commutator, so that for all  $X, Y \in \mathfrak{g}$ , we define

$$[X,Y] = -i(XY - YX) \tag{6.9}$$

It can be shown [9] that for all  $X, Y \in \mathfrak{g}$ , one has  $[X, Y] \in \mathfrak{g}$ , together with the other conditions needed for the commutator. Therefore, a Lie algebra is indeed an algebra.

For many purposes, the Lie algebra of a Lie group is easier to study than the group itself. One of the reasons has to do with linearity: for all  $X, Y \in \mathfrak{g}$ , one has  $aX + bY \in \mathfrak{g}$  for all real a, b. Moreover, since  $\mathfrak{g}$  is a vector space, it is often sufficient to investigate one of its bases. On the contrary, the Lie group may have a more complicated group operation, and there is in general no notion of a basis for a Lie group.