Treball final de grau

GRAU DE MATEMÀTIQUES

Facultat de Matemàtiques
Universitat de Barcelona

The Lorentz Group

Autor: Guillem Cobos

Director:	**Dr. Neil Strickland**
Realitzat a:	**Department of Pure Mathematics**

Barcelona, June 29, 2015

Abstract

In this project we aim to understand the structure and some properties of the Lorentz group, a mathematical object arising from the theory of special relativity which is said to express some of the main symmetries in the laws of Physics. The first section of the project is concerned about providing physical intuition about the concepts involved in the definition of the Lorentz group. The main goal of the project is to prove an isomorphism between the restricted Lorentz group and the projective linear group $\mathrm{PSL}_2(\mathbb{C})$. Once achieved this result, we use it to build a scheme that will let us study the conjugacy classes of the restricted Lorentz group.

Contents

1 Introduction to Special Relativity

Einstein, in the special theory of relativity, proved that different observers, in different states of motion, see different realities.

Leonard Susskind

In this chapter we aim to present the frame in which we will work during the rest of the project. As we will see further on, the Lorentz group is an isometry group of transformations of a four dimensional vector space, equipped with a quite special “norm”. This is in fact what we call *Minkowski space*, and it is the basic frame for the work in special relativity. More specifically, the Minkowski space is a four dimensional real vector space, in which there is defined a bilinear form η which captures the idea of “distance”, or “separation” of events in spacetime. To motivate the definition of this bilinear form η, sometimes called *Minkowski inner product*, we have to give an overview of special relativity.

Special Relativity (SR) is a physical theory that tells us about the nature of space and time, and how they are interwoven forming an unsplittable continuum. The theory has its foundations in the so called *principle of relativity*, a concept which is worth giving some thought, and goes like this. Suppose we have two observers which are moving at constant speed relative to each other, then there is no physical experiment they can do in order to determine at what velocity they are moving apart from each other. In other words, the result of any experiment performed by an observer does not depend on his speed relative to other observers (who are not involved in the experiment).

And, roughly speaking, this is saying that in the universe there is no privileged point of view which is still in space (and therefore, every other thing moves towards it or away from it). Hence staying still in space is something which simply doesn't make sense in the construction of the universe used in special relativity, since there could always be another observer for which what we thought was still in space is now moving. Thus the speed of a moving body is not a good property of “the body”, nor are coordinates, since they all depend on the observer from which we are taking the measures. It is reasonable to ask if there is any property of moving bodies in our universe which is not dependent on the observer used to measure its coordinates. This question is answered by considering the so called *spacetime interval*, a key concept that we will introduce in the next section, and that is in the very heart of the theory of special relativity. The spacetime interval will arise from a quadratic form defined on $\mathbb{R}^4$, playing an analog to the Euclidean distance defined in three-dimensional space.

1.1 Starting Assumptions for SR Model

We are willing to find a mathematical model to explain physical phenomena been observed by experimentation. The starting point of special relativity (SR) is the identification of our universe with an abstract set Ω whose elements we are going to call *events*. The set Ω is more commonly known as *spacetime*. Intuitively, an event is something that happens at a certain location in space and at a given time, for example Neil Armstrong stepping on the moon or Usain Bolt beating the 100 meter world record. We wish to give some affine structure to this set, since physical experimentation seems to suggest so.

(1) There is a procedure using rods and clocks which allows an observer $\mathcal{O}$ to build a bijective map $f_{\mathcal{O}} : \Omega \longrightarrow \mathbb{R}^4$. Roughly speaking, this tells us that there is a standard way for any observer to give coordinates to events. This implies that any observer must be able to construct three rods of unit length and display them orthogonally with a right hand orientation. Moreover, it must be able to construct a clock. Both rods and clocks are assumed to rise from physical experimentation. To make things a bit more precise, let us highlight that given an event $u \in \Omega$ the vector $f_{\mathcal{O}}(u)$ will consist of the time coordinate of u measured by $\mathcal{O}$ in the first position, and the space coordinates of u measured by $\mathcal{O}$ in the three last positions.

    Definition 1.1. Any map that arises as $f_{\mathcal{O}}$ for some $\mathcal{O}$ is called an Intertial Reference Frame (or IRF). From the physical point of view, we want to consider an IRF as an unaccelerated *observer* moving through spacetime in a straight line which is able to measure the space and time coordinates of all other events in the universe. That is why we sometimes refer to an IRF as an *inertial observer*, when we want to emphasise the physical picture.

    Notice that when we set a particular IRF on the set of events Ω, we are deciding what event in Ω will correspond to the origin (i.e. the only point with zero space and time coordinates), and also we are deciding what are we going to consider as right and left (which determines completely the directions of front and back, and up and down). We are also deciding what ahead in time is going to be while choosing a specific IRF.
    For further purposes, we should specify what do we mean when we say that two IRFs *share the same origin*. Two IRFs $f_{\mathcal{O}}$ and $f_{\mathcal{O}'}$ are said to share origin if $f_{\mathcal{O}}^{-1}(0) = f_{\mathcal{O}'}^{-1}(0)$.

(2) Experimental fact: if two observers $\mathcal{O}$ and $\mathcal{O}'$ follow the above procedure, then there is an affine map $g : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ such that $f_{\mathcal{O}'} = g \circ f_{\mathcal{O}}$, i.e. $g(x) = Ax + b$ for some $A \in M_{4\times4}(\mathbb{R})$ and $b \in \mathbb{R}^4$. This is actually an idealisation, since for example it has been proven to be false in presence of strong gravity.

(3) We define *speed* of a moving particle as measured by $\mathcal{O}$. If $\mathcal{O}$ sees a particle at events $\xi, \xi' \in \Omega$ with coordinates $f_{\mathcal{O}}(\xi) = (t, x, y, z)$ and $f_{\mathcal{O}}(\xi') = (t', x', y', z')$ then the speed is defined to be

$$s_{\mathcal{O}}(\xi, \xi') = \frac{\sqrt{(x'-x)^2 + (y'-y)^2 + (z'-z)^2}}{t'-t}.$$

The speed has some good and not so brilliant properties. On the one hand, it is an intuitive concept, this is to say, speed is a property of moving particles we seem to understand quite well. But then, it is obvious that in the general case, the speed of a moving particle at events ξ and ξ' will depend on the choice of observer $f_{\mathcal{O}} : \Omega \longrightarrow \mathbb{R}^4$ we make to record the coordinates of events ξ and ξ'.

We are going to introduce a magnitude called *spacetime interval* between two events ξ and ξ' which will actually turn out to be well defined on $\Omega \times \Omega$. In other words, the spacetime interval is going to be a map $\Delta s^2 : \Omega \times \Omega \longrightarrow \mathbb{R}$, such that $\Delta s^2(\xi, \xi')$ will not depend on the choice of observer we make to measure the coordinates of the pair of events (ξ, ξ'). For now we will define the spacetime interval between two events in a particular reference frame $\mathcal{O}$, and later on we shall see that, as a consequence of (4), it is invariant under change of coordinates between inertial reference frames. For an inertial observer $\mathcal{O}$ who gives coordinates $f_{\mathcal{O}}(\xi) = (t, x, y, z)$ and $f_{\mathcal{O}}(\xi') = (t', x', y', z')$ to events ξ and ξ' in Ω respectively, we define the quantity $\Delta s^2_{\mathcal{O}}(\xi, \xi')$ as

$$\Delta s^2_{\mathcal{O}}(\xi, \xi') = (t'-t)^2 - (x'-x)^2 - (y'-y)^2 - (z'-z)^2.$$

Note that if we call $\mathsf{q} : \mathbb{R}^4 \longrightarrow \mathbb{R}$ the quadratic form defined by $\mathsf{q}(t, x, y, z) = t^2 - x^2 - y^2 - z^2$, then $\Delta s^2_{\mathcal{O}}(\xi, \xi')$ can be put as

$$\Delta s^2_{\mathcal{O}}(\xi, \xi') = \mathsf{q}(f_{\mathcal{O}}(\xi') - f_{\mathcal{O}}(\xi)).$$

(4) Experimental fact: assume photons are at events ξ and ξ', then $s_{\mathcal{O}}(\xi, \xi') = s_{\mathcal{O}'}(\xi, \xi')$ for any observers $\mathcal{O}$, $\mathcal{O}'$. We can make that speed be equal to 1 by specifying a correct procedure in (1). This law is commonly known in physics as *constancy of speed of light*, and it can be paraphrased as: "Light is always propagated in empty space with a definite speed c which is independent of the state of motion of the emitting body". That is to say, the speed of light is the same when measured by any inertial observer. Note that in this case, the spacetime interval between the events ξ and ξ' is 0. This experimental fact shows that when two events ξ and ξ' are separated by a light beam, then $\Delta s^2_{\mathcal{O}}(\xi, \xi') = \Delta s^2_{\mathcal{O}'}(\xi, \xi') = 0$, for any choice of observers $\mathcal{O}$ and $\mathcal{O}'$.

(5) Experimental fact: if a rod lies perpendicular to the relative velocity of two observers, then the two observers agree on its length. This fact becomes useful when we try to give a mathematical description of the transformation of coordinates between inertial reference frames.

If we accept these facts, then roughly speaking the way one observer sees the other observer's measures of space and time must suffer transformations, in order to make the speed of light equal for any inertial observer. We are actually able to work out exactly how do the coordinates of an event recorded by two different IRFs should relate.

1.2 Consequences to the Assumptions

In this section we aim to give a motivation for the definition we will later on give of the Lorentz group. The main result we wish to see is known as the *invariance of the spacetime interval*, and it is summed up in the following theorem.

Theorem 1.2. Assume we have two different inertial reference frames $f_{\mathcal{O}}$ and $f_{\mathcal{O}'}$. Then given any two events ξ and ξ' in Ω, the spacetime interval between them measured by $\mathcal{O}$ and $\mathcal{O}'$ is the same. That is to say,

$$\Delta s^2_{\mathcal{O}}(\xi, \xi') = \Delta s^2_{\mathcal{O}'}(\xi, \xi').$$

A proof of this result can be found in Chapter 1.6 of [1]. It is a consequence of assumptions (4) and (5) in the list of last section. The following corollary is important to motivate the technical definition of the Lorentz group, as the group of isometries of spacetime which leave the origin fixed.

Corollary 1.3. If observers $\mathcal{O}$ and $\mathcal{O}'$ share the same origin, then the map $g : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ from (2) is a homogeneous linear transformation and it preserves the quadratic form q.

Proof. The map g must satisfy $f_{\mathcal{O}'} = g \circ f_{\mathcal{O}}$. Call $\omega = f^{-1}_{\mathcal{O}}(0) = f^{-1}_{\mathcal{O}'}(0)$ the event we refer to as origin. Then imposing $f_{\mathcal{O}'}(\omega) = (g \circ f_{\mathcal{O}})(\omega)$ means that $g(0) = 0$. So the affine map g is indeed homogeneous, as we wanted to show.
Moreover, the invariance of the spacetime interval says that $\Delta s^2_{\mathcal{O}}(\xi, \omega) = \Delta s^2_{\mathcal{O}'}(\xi, \omega')$, for any $\xi \in \Omega$. In other words,

$$\mathsf{q}(f_{\mathcal{O}'}(\xi) - f_{\mathcal{O}'}(\omega)) = \mathsf{q}(f_{\mathcal{O}}(\xi) - f_{\mathcal{O}}(\omega)).$$

Hence,

$$\mathsf{q}(g \circ f_{\mathcal{O}}(\xi)) = \mathsf{q}(f_{\mathcal{O}}(\xi)).$$

Because (1) assumes that $f_{\mathcal{O}}$ is a bijection between Ω and $\mathbb{R}^4$, we get that $\mathsf{q}(g(x)) = \mathsf{q}(x)$, for all $x \in \mathbb{R}^4$. Therefore the transformation of coordinates g preserves the quadratic form q. $\square$

Because we will neglect the case where IRFs do not share the origin (and it is reasonable to do so because a simple translation in spacetime will make two IRFs have the same origin) we can say that in this context, the quadratic form is preserved by changes of coordinates between IRFs. From now on, we will refer as *Minkowski rotations* to change of coordinates between IRFs with a shared origin. Thus the last corollary is saying that it is a necessary condition for Minkowski rotations to preserve the quadratic form,

$$\mathsf{q} : (t, x, y, z) \mapsto t^2 - x^2 - y^2 - z^2.$$

Also, Minkowski rotations are assumed to keep the orientation of space and the direction of time. So in other words, when dealing with IRFs with same origin

$$\{\text{Minkowski rotations}\} \subseteq \left\{ \begin{array}{c} f : \mathbb{R}^4 \to \mathbb{R}^4 \text{ linear}, \ \mathsf{q}(x) = \mathsf{q}(f(x)), \ x \in \mathbb{R}^4 \\ \det f > 0 \text{ and } e_1 \cdot f(e_1) > 0 \end{array} \right\}.$$

This observation gives a good starting point for working algebraically with the objects we are interested in: transformations of coordinates between IRFs who share a common origin. And this is of course because it gives a mathematical characterisation of these so called Minkowski rotations. In fact, the right hand set, equipped with the composition of linear maps, possesses a group structure and is what we call *restricted Lorentz group.* The idea is that later on we will see that the inclusion of sets we had before is indeed an equality. In other words, we will prove that elements in the restricted Lorentz group are indeed Minkowski rotations between some IRFs with common origin (and this is exactly what we want the restricted Lorentz group to be). It is worth mentioning that the restricted Lorentz group is contained in what is called the Poincaré group. In the latter, elements are transformations of coordinates between IRFs which may or may not have a shared origin. So the Poincaré group deals with the more general picture.

1.3 Picturing the Elements of the Restricted Lorentz Group

It is good to start picturing what are the elements in the restricted Lorentz group are going to look like, so let us work a bit more on our physical intuition to grasp the nature of these transformations between IRFs. We ask ourselves what kind of differences could there be between two IRFs in the most general case (so this means that now we shall not assume that IRFs share the same origin). It turns out that there are only three main transformations of $\mathbb{R}^4$ which make two IRFs not be the same. In other words, given two IRFs, with only three distinctive affine transformations of $\mathbb{R}^4$ we can send one to the other. For this section, imagine we have two different observers $\mathcal{O}$ and $\mathcal{O}'$ in spacetime (i.e. two reference frames $f_{\mathcal{O}}, f_{\mathcal{O}'} : \Omega \longrightarrow \mathbb{R}^4$).

Translations. These are transformations involving space and time. By applying a translation to spacetime we get to connect the two different origins of both IRFs, and therefore it makes it possible to treat the two IRFs as if they initially coincide in spacetime. In other words, there exists a unique $w \in \mathbb{R}^4$ such that

$$0_{\mathcal{O}} = 0_{\mathcal{O}'} + w,$$

where $0_{\mathcal{O}}, 0_{\mathcal{O}'}$ are the coordinates of the respective origins.

Rotations. These linear maps only involve space. It could perfectly be that what $\mathcal{O}$ understands as right/left, front/back and up/down is not what the second does. By rotating the first frame of reference in space, we get to an agreement of the directions of space by the two observers. In a more formal way, there exists $\mathcal{R} \in \mathrm{SO}(3, \mathbb{R})$ such that

$$\mathcal{R}(f_{\mathcal{O}}(\xi)) = f_{\mathcal{O}'}(\xi), \text{ for all } \xi \in \Omega.$$

Boosts. These are the transformations which intertwine space and time. Let us assume we have made the convenient corrections for the two IRFs (this is, applied spacetime translation and spatial rotation to the first to make it coincide with the second). So we can assume that they agree on an origin and their spatial directions coincide. The only way in which these two IRFs still cannot be the same is that it may happen that the two observers are moving apart from each other at some relative velocity $v \in \mathbb{R}^3$, pointing in a certain direction in space. The boost determined by the velocity v is the connection between the two IRF at this point. This will alter the scaling of the axes of the moving IRF seen by the still IRF, so that if we are considering $\mathcal{O}$ to be still in space while $\mathcal{O}'$ moves away from $\mathcal{O}$ at velocity v, it will shrink the spatial axis with the direction of v of $\mathcal{O}'$ as well as its time axis. This will result in an effect of *time dilatation* when $\mathcal{O}$ looks at the running clock of $\mathcal{O}'$, and a *space contraction* when $\mathcal{O}$ looks at the spatial measures on the axis given by the direction v of $\mathcal{O}'$.

1.4 Derivation of the Formula of the Lorentz Boost

From among these three types of affine transformations, perhaps the one we have less intuition upon is the boost. This is most likely because the notion of boost involves the constancy of the speed of light, a postulate somewhat counterintuitive to our daily experiences. We are going to inspect the standard version of the boost, this is the boost in the x-axis.

Assume we have two IRFs $\mathcal{O}$ and $\mathcal{O}'$ with a shared origin. Also, suppose that originally their axes are collinear. Let us take the perspective of $\mathcal{O}$ first, and then say that the frame $\mathcal{O}'$ is moving away from $\mathcal{O}$, which is still in space, at relative velocity v. We can imagine a light pulse emitted at the origin that travels in a spherical wave at speed c. The main point here is that the speed at which this light pulse travels through space will be the same when measured by any inertial observer. Now, let P be a point in space belonging to the light wavefront, which lays at distance r from $\mathcal{O}$ and distance r' from $\mathcal{O}'$. Let x, y, z be the spatial coordinates that $\mathcal{O}$ records of P, and similarly let t be the time that $\mathcal{O}$ observes it takes for the light pulse emitted in the origin to get to P. A symmetric construction applies for defining x', y', z' and t'. From the very definition of distance in space, we get

$$x^2 + y^2 + z^2 = r^2$$
$$x'^2 + y'^2 + z'^2 = r'^2$$

Then, by the constancy of speed of light postulated by special relativity, it must happen that observer $\mathcal{O}$ and $\mathcal{O}'$ see the light travelling at the same speed. Thus

$$r = ct$$

$$r' = ct'.$$

Let us start writing down the relations between the coordinates of P recorded by $\mathcal{O}$ and the ones taken by $\mathcal{O}'$. Since the relative velocity between the two observers is collinear with the x-axis, we shall assume that the coordinates of P perpendicular to the x-axis remain unchanged. So $y' = y$ and $z' = z$. In other words, the directions perpendicular to the relative velocity between the two observers don't notice the effect of the boost. We are assuming that the transformation is linear, so of the form

$$x' = \alpha x + \beta t, \tag{1}$$

for some $\alpha, \beta \in \mathbb{R}$. We know the situation for $x' = 0$, which is $x = vt$. Substituting in the equation before we get that $0 = \alpha vt + \beta t$, therefore $\beta = -\alpha v$. And so the transformation now looks like this

$$x' = \alpha(x - vt). \tag{2}$$

Now consider the inverse transformation of coordinates. This is, make observer $\mathcal{O}'$ still in space and let $\mathcal{O}$ be moving away from $\mathcal{O}'$ at relative velocity $-v$. The whole procedure is analogous, the bit where it differs is when dealing with v in the transformations, we now have to swap v for $-v$. As a result, we get

$$x = \alpha(x' + vt).$$

Combining equations (1) and (2) we get a linear constraint for t', t and x of the form

$$t' = \alpha t + \frac{1 - \alpha^2}{\alpha v} x.$$

But this is not it, since there is one more constraint we have to keep in mind. Recall $r^2 - c^2t^2 = r'^2 - c^2t'^2 = 0$. In a more detailed fashion,

$$x^2 - ct^2 = x'^2 - c^2t'^2.$$

Substituting in the values found for x' and t' and setting equal the coefficients of x^2 in both sides of the equation leads to

$$1 = \alpha^2 - c^2 \left(\frac{1-\alpha^2}{\alpha v} \right)^2 .$$

Solving this quadratic equation for α we get

$$\alpha = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}},$$

which is often called the Lorentz factor. So now we have completely determined the transformation of coordinates of an event P in the light cone between the frames $\mathcal{O}$ and $\mathcal{O}'$:

$$\begin{cases} t' = \alpha t + \dfrac{1-\alpha^2}{\alpha v} x \\ x' = \alpha(x - vt) \\ y' = y \\ z' = z \end{cases} \tag{3}$$

where α is the Lorentz factor. Now notice that (3) only applies in principle to coordinates of events in the light cone. But we can extend linearly to get an endomorphism of $\mathbb{R}^4$. This is because we can find a basis of $\mathbb{R}^4$ sitting inside the light cone, for example $\{(1,1,0,0),(1,-1,0,0),(1,0,1,0),(1,0,0,1)\}$. If the linear map is determined on these vectors, then we get a unique linear map $\mathbb{R}^4 \to \mathbb{R}^4$ by extending linearly.

In the same way we defined the boost in the x-axis given a certain relative speed $v \in \mathbb{R}$, we can define a boost in any vector in space w. The vector w will indeed represent the relative velocity between the two inertial observers. The boost will once more tell us the relation between the coordinates of events recorded by the two inertial observers falling apart at velocity w. The way to obtain an expression for this transformation is pretty straightforward. Let $\mathcal{R} \in \mathrm{SO}(3,\mathbb{R})$ be a rotation mapping $\mathcal{R}(w) = e_1$. Then if we call $\mathcal{B}_x$ the boost in the direction $x \in \mathbb{R}^3$ we have

$$\mathcal{B}_w = \mathcal{R}^{-1} \circ \mathcal{B}_{e_1} \circ \mathcal{R}.$$

2 The Lorentz Group

We are essentially treating how do IRFs with a common origin relate, this means how do coordinates of events recorded by two IRF with the same origin transform. In order to do that efficiently, the starting assumption in that the set of *events* is identified with $\mathbb{R}^4$ with the standard basis, which can be viewed as a *preferred* or *privileged* inertial reference frame.

2.1 Definition and First Consequences

Consider the vector space $\mathbb{R}^4$ with the standard basis $\{e_i\}_{1\leq i\leq 4}$. In order to make notation a bit more simple, we are not going to distinguish between a vector in $\mathbb{R}^4$ and its 4×1 real matrix representation with respect to the standard basis. This is due to the fact that we will always be working with the standard basis, so there should not be any room for ambiguity. In a similar way, we will generally notate indistinguishably an endomorphism of $\mathbb{R}^4$ and its 4×4 real matrix representation in the standard basis. This is, we will refer to linear transformations of $\mathbb{R}^4$ as 4×4 real matrices.

In this context, we define a bilinear form $\eta : \mathbb{R}^4 \times \mathbb{R}^4 \longrightarrow \mathbb{R}$ as follows. Let $v, w \in \mathbb{R}^4$, then we set:

$$\eta\,(v, w) = v^T J w,$$

where

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \in M_{4\times 4}(\mathbb{R}).$$

From now on, when referring to the bilinear form η applied to two vectors $v, w \in \mathbb{R}^4$ we will write simply (v, w) instead of $\eta\,(v, w)$. It is convenient to point out how to compute $(\mathbf{x}, \mathbf{y})$ given coordinates of vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^4$. If we write $\mathbf{x} = (t, x)^T$ with $t \in \mathbb{R}$ and $x \in \mathbb{R}^3$; and $\mathbf{y} = (s, y)^T$ with $s \in \mathbb{R}$ and $y \in \mathbb{R}^3$ we can calculate:

$$(\mathbf{x}, \mathbf{y}) = ts - x \cdot y,$$

where we are using the notation $\cdot$ to refer to the the dot product in $\mathbb{R}^3$. We must warn that later on the same notation might be used as well to talk about the dot product in $\mathbb{R}^4$.

Note that η is *symmetric*, i.e. for all $v, w \in \mathbb{R}^4$ we have $(v, w) = (w, v)$. Moreover, it is *non-degenerate* which means that if $(v, w) = 0$ for all $w \in \mathbb{R}^4$, then $v = 0$. Note that

(v, w) need not be positive (e.g. take $v = w = e_3$ and notice $(v, w) = -1$). That is why we cannot say η is an inner product. We shall give a name to the induced quadratic form by η,

$$\mathtt{q}(\mathbf{x}) = (\mathbf{x}, \mathbf{x}) = t^2 - \|x\|^2.$$

The *Minkowski norm* of a vector $v \in \mathbb{R}^4$ is defined to be

$$\|v\|_{\mathcal{M}} = \sqrt{|\mathtt{q}(v)|}.$$

We should remark that η does not induce a norm, since $\|.\|_{\mathcal{M}}$ fails to be subadditive. Even though, the bilinear form η gives a useful generalisation of the concept of *length* in the Minkowski space. We should carry on with a definition that will be useful in the future. Together with the algebraic statement, we build up with physical intuition of the concepts defined.

Definition 2.1. Let $v \in \mathbb{R}^4$. We say that v is:

1. *Time-like* if $\eta(v, v) > 0$. This happens if and only if there exists an IRF for which v is a pure time direction.

2. *Space-like* if $\eta(v, v) < 0$. This happens if and only if there exists an IRF for which v is a pure space direction.

3. *Light-like* if $\eta(v, v) = 0$. For every IRF we choose v will be an event on a light ray.

We want to consider those endomorphisms which preserve the bilinear form defined above. Knowing that every endomorphism can be represented as an element in $M_{4\times4}(\mathbb{R})$, we see how to formalise the bilinear form preserving property. An endomorphism f represented by the matrix $A \in M_{4\times4}(\mathbb{R})$ is said to preserve the bilinear form $(\ ,\ )$ if we have:

$$(v, w) = (f(v), f(w)) \qquad \forall\ v, w \in \mathbb{R}^4.$$

If we are thinking in terms of coordinates, this condition translates into:

$$v^T J w = (Av)^T J A w = v^T A^T J A w \qquad \forall\ v, w \in M_{4\times1}(\mathbb{R})$$

Because this property must be satisfied for all $v, w \in \mathbb{R}^4$, we can say that A preserves the bilinear form η if and only if it satisfies the equality $J = A^T J A$. This leads to the definition of the Lorentz group, as the set of endomorphisms of $\mathbb{R}^4$ preserving the bilinear form η.

Definition 2.2. The *Lorentz group* $(\Gamma, \cdot)$ is defined to be the set of all matrices in $M_{4\times 4}(\mathbb{R})$ preserving the bilinear form η, with the standard matrix multiplication as the defined operation. Formally,

$$\Gamma = \{A \in M_{4\times 4}(\mathbb{R}) \mid J = A^T JA\}.$$

Remark 2.3. In the first section, when we introduced the restricted Lorentz group from a physical point of view, we said that its underlying set was the set of endomorphisms of $\mathbb{R}^4$ which preserved the quadratic form q, preserved the orientation of space and the direction of time. We want to stress the fact that for a linear map $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$, preserving the bilinear form η and the induced quadratic form q are equivalent things. It is clear that is f preserves η, then it also preserves q.
Conversely, assume $\mathsf{q}(f(v)) = \mathsf{q}(v)$ for all $v \in \mathbb{R}^4$. Then take two vectors $v, w \in \mathbb{R}^4$ and observe

$$\begin{aligned}\mathsf{q}(f(v+w)) &= (f(v+w), f(v+w)) \\ &= (f(v), f(v) + (fw), w)) + 2(f(v), f(w)) \\ &= \mathsf{q}(f(v)) + \mathsf{q}(f(w)) + 2(f(v), f(w)).\end{aligned}$$

On the other hand,

$$\mathsf{q}(v+w) = (v+w, v+w) = \mathsf{q}(v) + \mathsf{q}(w) + 2(v, w).$$

Using the fact that f preserves $\mathsf{q}(v)$ and $\mathsf{q}(w)$, when making equal the expressions of $\mathsf{q}(f(v+w))$ and $\mathsf{q}(v+w)$, we realise that it must happen

$$(v, w) = (f(v), f(w)),$$

which is exactly what we wanted to prove.

The first thing that should be checked is that Γ has a group structure with the standard matrix multiplication. In the first place, we ask if Γ is closed under standard matrix multiplication. This can be checked easily by taking two elements $A, B \in \Gamma$, and seeing:

$$(AB)^T JAB = B^T A^T JAB = B^T JB = J.$$

So we just saw $AB \in \Gamma$. Clearly, the operation is associative (property inherited by the standard matrix multiplication in $M_{4\times 4}(\mathbb{R})$). Moreover, the identity matrix $\mathrm{I} \in M_{4\times 4}(\mathbb{R})$ lies in Γ. We are only left to show that every element has an inverse, but for that we will need the following proposition.

Proposition 2.4. If $A \in \Gamma$, then $|\det A| = 1$.

Proof. Because A is in Γ, we get that $J = A^T JA$ and so

$$\det J = \det(A^T JA) = \det A^T \det J \det A = (\det A)^2 \det J.$$

But $\det J = -1$, so we get $(\det A)^2 = 1$ iff $|\det A| = 1$. □

Using this property we deduce that every matrix in Γ is invertible, so indeed $(\Gamma, \cdot)$ satisfies the inverse element existence property. Moreover, the inverse of any matrix $A \in \Gamma$ also lies in Γ, we just need to see that it preserves the form η:

$$A^T JA = J \implies JA = (A^T)^{-1} J \implies J = (A^{-1})^T JA^{-1}.$$

We are about to focus in the study of a subgroup of the Lorentz group called the restricted Lorentz group, usually denoted as $\Gamma_+^\uparrow$.

Definition 2.5. We call the *restricted Lorentz group* the set of those linear transformations $A = (a_{ij}) \in \Gamma$ preserving the orientation of space and whose first entry, namely a_{11}, is positive. Equivalently,

$$\Gamma_+^\uparrow = \{A = (a_{ij}) \in \Gamma \mid \det A = 1, \quad a_{11} > 0\}.$$

We want to check that $(\Gamma_+^\uparrow, \cdot)$ is a subgroup of $(\Gamma, \cdot)$. In order to do that, we'll have to work first on some definitions.

Definition 2.6. We call the null cone $C \subset \mathbb{R}^4$ the set of vectors $v \in \mathbb{R}^4$ such that $\eta(v, v) = 0$. The subset C_+ consisting of those $v \in C$ such that $v \cdot e_1 > 0$ is called future null cone. Similarly, we denote the set of vectors $v \in C$ such that $v \cdot e_1 < 0$ by C_-, and we call it the past null cone.

Notice that any transformation in the Lorentz group leaves C invariant. Let $A \in \Gamma$, then take $y = Ax \in A(C)$, and check that $(y, y) = (Ax, Ax) = (x, x) = 0$ since $x \in C$. Therefore $AC \subset C$. It is also worth noticing that C can be put as the disjoint union of C_+ and C_-. All these properties are really useful and will be applied in the following propositions.

Lemma 2.7. If $A \in \Gamma$, then $A^T \in \Gamma$.

Proof. From the matrix equation

$$A^T JA = J$$

We take the inverse both sides

$$A^{-1} J (A^T)^{-1} = J$$

And so we get

$$J = AJA^T$$

□

The following proposition is going to be useful to characterise the restricted Lorentz group in a more intuitive way.

Proposition 2.8. Let $A = (a_{ij}) \in \Gamma$. We have one of the following situations:

1. $a_{11} > 0$ iff $AC_+ \subset C_+$ iff $AC_- \subset C_-$.
2. $a_{11} < 0$ iff $AC_+ \subset C_-$ iff $AC_- \subset C_+$.

Proof. Suppose $a_{11} > 0$. Then, take $x \in C$. It is clear that Ax will lie in C, since $(Ax, Ax) = (x, x) = 0$. We just have to see that $(Ax)_1 = (Ax) \cdot e_1 > 0$. We write the matrix A in blocks as follows:

$$A = \begin{pmatrix} a_{11} & u^T \\ v & B \end{pmatrix}$$

where $u \in \mathbb{R}^3$, $v \in \mathbb{R}^3$ and $B \in M_{3\times 3}(\mathbb{R})$. With this construction, observe that

$$a_{11}^2 - \|u\|^2 = (Ae_1, Ae_1) = (e_1, e_1) = 1,$$

and $a_{11} > 0$, so $a_{11} > \|u\|$. Now an arbitrary element $x \in C_+$ has the form $x = (\|w\|, w)$, for some $w \in \mathbb{R}^3 \setminus \{0\}$, and

$$(Ax)_1 = a_{11}\|w\| + u^T \cdot w \geq a_{11}\|w\| - \|u\|\|w\| = (a_{11} - \|u\|)\|w\| > 0.$$

Conversely, suppose $AC_+ \subset C_+$. Using the same notation as before, we define the vector:

$$m = \begin{pmatrix} 1 \\ w \end{pmatrix}$$

Where $w \in \mathbb{R}^3$ is a unit vector orthogonal to u. Notice that m lies in C_+. By hypothesis, we know that $Am \in C_+$, which implies $(Am)_1 = a_{11} + u^T \cdot w = a_{11} > 0$.

We shall now proceed to prove that $AC_+ \subset C_+$ if and only if $AC_- \subset C_-$, whenever $A \in \Gamma$. Let $x \in C_-$. Notice that $-x \in C_+$, hence $A(-x) \in C_+$. Because $A(-x) = -Ax \in C_+$, we get that Ax must be in C_-. We would prove the converse in a similar way.

In order to prove 2, we will start with a matrix $A \in \Gamma$ such that $A_{11} < 0$. In that case, we notice that $-JA$ satisfies $(-JA)_{11} = -A_{11} > 0$. Therefore, it is equivalent to have $A_{11} < 0$ or $(-JA)_{11} > 0$. By applying what we just proved, we can see that $(-JA)C_+ \subset C_+$. By multiplying by minus the identity both sides, we get that the last inclusion is equivalent to $JAC_+ \subset C_-$. Now, remember that J leaves the past null cone invariant, hence last inclusion is same as $AC_+ \subset JC_- \subset C_-$. Hence A sends C_+ to C_-. On the other hand, if we start working from the inclusion $(-JA)C_- \subset C_-$ provided by the first part of the proposition, and we follow similar steps as before we get that A must send C_- to C_+.

□

2.2 Topological Properties of the Lorentz Group

The Lorentz group, as a subset of $M_{4\times 4}(\mathbb{R})$ inherits the topology of $\mathbb{R}^{16}$. To make this a little bit more precise, consider the composition

$$\Gamma \xrightarrow{\ i\ } M_{4\times 4}(\mathbb{R}) \xrightarrow{\ j\ } \mathbb{R}^{16},$$

where i is the natural inclusion of the subset Γ in $M_{4\times 4}(\mathbb{R})$, and j is the isomorphism of vector spaces $M_{4\times 4}(\mathbb{R})$ and $\mathbb{R}^{16}$. So in this context, a subset $U \in \Gamma$ will be open iff there is an open set $V \subset \mathbb{R}^{16}$ with the Euclidean topology such that $U = (j \circ i)^{-1}(V)$. Notice that this is nothing more than viewing Γ as a subset of $\mathbb{R}^{16}$, and giving it the subspace topology.

Because we just made Γ a topological space, we can start considering its topological properties. We will see that it is indeed disconnected, and has four connected components. Using the topological structure of Γ, we will be able to show that $\Gamma_+^\uparrow$ is normal subgroup of index 4 in Γ. Informally, this will mean that the Lorentz group will consist of 4 disjoint copies of $\Gamma_+^\uparrow$.

Proposition 2.9. $\Gamma_+^\uparrow$ is a subgroup of the Lorentz group of index 4 in Γ. Moreover, it is a normal subgroup, and $\Gamma/\Gamma_+^\uparrow$ is isomorphic to $C_2^2 = (\mathbb{Z}/2 \times \mathbb{Z}/2, \cdot)$.

Proof. We shall proceed in the following way. First of all, notice that $\{\pm I, \pm J\}$ is a subgroup of Γ. We want to build a continuous homomorphism from Γ to the multiplicative Klein 4-group, which we will call C_2^2. Consider the continuous maps $\det : \Gamma \longrightarrow \{-1, 1\}$ and $\operatorname{sgn}(-)_{11} : \Gamma \longrightarrow \{-1, 1\}$ defined as $\det(A) = \det A$ and $\operatorname{sgn}(-)_{11}(A) = \operatorname{sgn} a_{11}$, for all $A = (a_{ij}) \in \Gamma$. We should highlight the fact that $\operatorname{sgn}(-)_{11}$ is well-defined as a map thanks to Proposition 2.9, which tell us that $a_{11} \neq 0$.

Then we can define a continuous map

$$\begin{array}{rccc} \theta = (\det, \operatorname{sgn}(-)_{11}) : & \Gamma & \longrightarrow & C_2^2 \\ & A & \longmapsto & (\det A, \operatorname{sgn} a_{11}). \end{array}$$

We shall proceed to prove that θ is a group homomorphism. We already know that $\det : \Gamma \longrightarrow C_2$ is a group homomorphism. So it suffices to show that $\operatorname{sgn}(-)_{11} : \Gamma \longrightarrow C_2$ satisfies the axioms of group homomorphism. It is clear that $\operatorname{sgn}(-)_{11}$ preserves the identity element. So we only need to see it preserves inverses and multiplication.

Start by showing $\operatorname{sgn}(-)_{11}$ preserves inversion. Let $A \in \Gamma$, suppose $(A)_{11} > 0$. By the proposition above we can tell that $AC_+ \subset C_+$ and $AC_- \subset C_-$. We argue by contradiction. Suppose $A^{-1}C_+ \not\subset C_+$, this means there exists an $x \in C_+$ such that $A^{-1}x \in C_-$, since any A maps C into itself. Now it follows that $A(A^{-1}x)$ must be in C_-, because A maps

C_- into itself. But this contradicts our choice of $x \in C_+$. Therefore $A^{-1}C_+ \subset C_+$, and by the preceding proposition, we get that $(A^{-1})_{11} > 0$, thus $\operatorname{sgn}(A^{-1}) = \operatorname{sgn} A_{11} = 1$. Using the last proposition, we would similarly get to see that, if $A_{11} < 0$, then $\operatorname{sgn}(A^{-1}) = -1$.

We now have to show the multiplicative property of $\operatorname{sgn}(-)_{11}$. Take $A, B \in \Gamma$. Suppose $(AB)_{11} > 0$, this means AB maps C_+ into itself. There are only two possible cases. In the first place, it could happen that both A and B map C_+ into itself. In that case it easy to see $\operatorname{sgn}(AB)_{11} = \operatorname{sgn} A_{11} \cdot \operatorname{sgn} B_{11} = 1$. Also, it could happen that B maps C_+ into C_- and A maps C_- into C_+. Then, by the last proposition, $\operatorname{sgn} A_{11} = \operatorname{sgn} B_{11} = -1$, and $\operatorname{sgn}(AB)_{11} = \operatorname{sgn} A_{11} \cdot \operatorname{sgn} B_{11}$ still holds.

Now, if $(AB)_{11} < 0$, the future null cone C_+ is mapped to C_- by AB. Reasoning by cases, we deduce that either B maps C_+ into C_- and A leaves C_- still, or B leaves C_+ still and A maps C_+ into C_-. In the former case, $\operatorname{sgn} B_{11} = -1$ and $\operatorname{sgn} A_{11} = 1$, while in the latter it is the other way around. In both cases though it happens that $\operatorname{sgn}(AB)_{11} = \operatorname{sgn} A_{11} \cdot \operatorname{sgn} B_{11} = -1$.

So we have seen there is a continuous homomorphism θ between Γ and C_2^2. Notice that the restriction of θ to the subgroup $\{\pm I, \pm J\}$ yields to an isomorphism with C_2^2. Thus θ is surjective. Define $\Gamma_+^\uparrow = \operatorname{Ker}\theta$, and because we put $\Gamma_+^\uparrow$ as the kernel of a group homomorphism, we automatically get that it is a normal subgroup of Γ. Moreover, by the first isomorphism theorem we get that $\Gamma/\Gamma_+^\uparrow \cong C_2^2$, and by the order of the group C_2^2 we can guess that the index of $\Gamma_+^\uparrow$ will be 4.

It is worth seeing that, since $\Gamma_+^\uparrow$ is a normal subgroup of Γ, the cosets of $\Gamma_+^\uparrow$ in Γ form a partition of Γ. This is indeed a general result from group theory, it falls directly from the fact that $\pi : \Gamma \longrightarrow \Gamma/\operatorname{Ker}\theta$ is well defined and surjective. The cosets of $\Gamma_+^\uparrow$ are disjoint since π is a well defined map. Moreover, by the surjectiveness of π we can ensure that the reunion of all the cosets give the whole group. So, we can write

$$\Gamma = \Gamma_+^\uparrow \sqcup (-I)\Gamma_+^\uparrow \sqcup J\Gamma_+^\uparrow \sqcup (-J)\Gamma_+^\uparrow.$$

□

Observe that having $\Gamma_+^\uparrow$ as the preimage of a continuous map allows us to say it is both closed and open. Recall $\Gamma_+^\uparrow = \theta^{-1}(\{1\}, \{1\})$. The codomain of θ, namely C_2^2 gets the subspace topology from $\mathbb{R}^2$, which is no other than the discrete topology on C_2^2. Therefore $\{(1, 1)\} \subset C_2^2$ is an open and closed subset, hence $\Gamma_+^\uparrow = \theta^{-1}\{(1, 1)\}$ must be open and closed. This fact is going to be of vital importance when trying to show that the restricted Lorentz group is the connected component of the identity element.

Proposition 2.10. The Lorentz group is disconnected. The restricted Lorentz group is connected, and it is indeed the connected component of the identity element.

Proof. Consider the continuous homomorphism $\theta : \Gamma \longrightarrow C_2^2$ defined above. Recall that if $A \in \Gamma$ then $|(A)_{11}| = |a_{11}| \geq 1$. Define $j_1 = -1$ and $j_2 = 1$ elements in $\mathbb{R}$. Observe that $\theta^{-1}(\{i\}, \{j_k\})$ forms a closed disjoint cover of Γ for $i = 1, -1$ and $k = 1, 2$. Therefore Γ cannot be connected.

Let us see now that $\Gamma_+^\uparrow$ is connected. Define

$$P = \{u \in \mathbb{R}^4 \mid u_1 > 0,\ (u, u) = 1\}.$$

There is a homeomorphism $\sigma : \mathbb{R}^3 \longrightarrow P$ given by $\sigma(x) = (\sqrt{1 + \|x\|^2}, x)$, so P is path connected. Also, using σ and the Cauchy-Schwartz inequality we can check that $(u, v) > 0$ for all $u, v \in P$. In fact, if we write $u = (u_1, x)$ and $v = (v_1, y)$ we can deduce from the fact that they both belong to P, that $\|x\| = \sqrt{u_1^2 - 1}$ and $\|y\| = \sqrt{v_1^2 - 1}$. By using the Cauchy-Schwartz inequality we get that

$$|\langle x, y\rangle| \leq \|x\|\ \|y\| = \sqrt{u_1^2 - 1}\ \sqrt{v_1^2 - 1} < u_1 v_1.$$

Hence the product $(u, v) = u_1 v_1 - \langle x, y\rangle$ will be strictly positive.

Now if we fix $u, v \in P$ we can define a linear map $\phi_{uv} : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ by

$$\phi_{uv}(x) = x + 2(x, u)v - \frac{(x, u + v)}{1 + (u, v)}(u + v).$$

We must show two properties concerning ϕ_{uv}. First, we want to see that $\phi_{uv} \in \Gamma^+$, for all $u, v \in P$. Later on, we will prove that $\phi_{uv}(u) = v$, for any given $u, v \in P$.

We start off by seeing $\phi_{uv} \in \Gamma$, with $u, v \in P$. It suffices to show $(\phi_{uv}(x), \phi_{uv}(x)) = (x, x)$, for all $x \in \mathbb{R}^4$. Later on we will argue how this implies that ϕ_{uv} preserves the bilinear form η, and therefore is an element of Γ. So take $x \in \mathbb{R}^4$, and expand

$$\begin{aligned}(\phi_{uv}(x), \phi_{uv}(x)) &= \left(x + 2(x,u)v - \frac{(x,u+v)}{1+(u,v)}(u+v), x + 2(x,u)v - \frac{(x,u+v)}{1+(u,v)}(u+v) \right) \\ &= (x,x) + 2(x,u)(x,v) - \frac{(x,u+v)}{1+(u,v)}(x,u+v) + 2(x,u)(v,x) + \\ &\quad + 4(x,u)^2(v,v) - 2(x,u)\frac{x,u+v}{1+(u,v)}(v,u+v) - \frac{(x,u+v)}{1+(u,v)}(u+v,x) + \\ &\quad - \frac{2(x,u)(x,u+v)}{1+(u,v)}(u+v,v) + \frac{(x,u+v)^2}{[1+(u,v)]^2}(u+v,u+v).\end{aligned}$$

So far we have just bilinearity and symmetry of the form $(\,,)$. We now use the fact that $(u,u) = (v,v) = 1$ to obtain simplifications such as $(v,u+v) = 1 + (u,v)$ and $(u+v,u+v) = 2 + 2(u,v)$. Overall, we get a much more simplified expression

$$\begin{aligned}(\phi_{uv}(x), \phi_{uv}(x)) &= (x,x) + 4(x,u)(x,v) + 4(x,u)^2 - 4(x,u)(x,u+v) \\ &\quad - 2\frac{(x,u+v)^2}{1+(u,v)} + 2\frac{(x,u+v)^2}{1+(u,v)} = \\ &= (x,x) + 4(x,u)(x,v) + 4(x,u)^2 - 4(x,u)^2 - 4(x,u)(x,v) = \\ &= (x,x).\end{aligned}$$

Once we have this result, we can argue by polarisation to prove that ϕ_{uv} preserves the bilinear form η. This is exactly the same proof that we did on Remark 2.3, which showed that any linear map preserving the form $\mathtt{q}$ will actually preserve η. Because the linear map ϕ_{uv} preserves η we get that it must be an element of the Lorentz group.

We still have to show that $\phi_{uv} \in \Gamma_+^\uparrow$. The first thing we check is that $(\phi_{uv})_{11}$ is strictly positive. Directly form the general expression of ϕ_{uv} we check

$$(\phi_{uv})_{11} = (e_1, \phi_{uv}(e_1)) = 1 + u_1v_1 + (u_1^2 - 1)v_1^2 + (v_1^2 - 1)u_1^2 > 0.$$

We have used the fact that $u, v \in P$ when taking into account $u_1, v_1 \geq 1$.
Now we aim to demonstrate $\det \phi_{uv} = 1$. Observe that the orthogonal subspace to $\langle u, v\rangle$ is a 2 dimensional real subspace of $\mathbb{R}^4$.

$$\langle u, v\rangle^0 = \{x \in \mathbb{R}^4 \mid (x,u) = (x,v) = 0\}.$$

The subspace $\langle u, v\rangle^0$ is given by two linearly independent equations, namely $(x,u) = 0$ and $(x,v) = 0$, therefore $\dim\langle u, v\rangle^0 = 2$. We can choose $p, q \in \langle u, v\rangle^0$ linearly independent vectors, and we obtain a basis of $\mathbb{R}^4$ by putting $\{u, v, p, q\}$. Observe that, from the very definition of ϕ_{uv} we get that $\phi_{uv}(p) = p$ and $\phi_{uv}(q) = q$. We wish to give the matrix of the endomorphism ϕ_{uv} in the basis specified above. Notice that $\phi_{uv}(u) = v$ and $\phi_{uv}(v) = -u - 2(u, v)v$. Hence

$$\phi_{uv} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & -2(u,v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and $\det \phi_{uv} = 1$.

Finally put $G = \{g \in \Gamma_+^\uparrow \mid ge_1 = e_1\}$. This is just the group of rotations of the space like basis vectors e_2, e_3 and e_4, therefore isomorphic to $\mathrm{SO}(3, \mathbb{R})$ and so connected. Define $\mu : P \times G \longrightarrow \Gamma_+^\uparrow$ by $\mu(v, g) = \phi_{e_1 v} \circ g$. We check that μ is a homeomorphism. If we achieve this, we will have shown $\Gamma_+^\uparrow$ is connected, and we are using that connectedness is a topologic invariant.

Let us define $\mu^{-1} : \Gamma_+^\uparrow \longrightarrow P \times G$ by setting $\mu^{-1}(h) = (h(e_1), \phi_{e_1 h(e_1)}^{-1} \circ h)$. Check that μ^{-1} is indeed the inverse of μ:

$$(\mu \circ \mu^{-1})(h) = \phi_{e_1 h(e_1)} \circ \phi_{e_1 h(e_1)}^{-1} \circ h = h,$$

for all $h \in \Gamma_+^\uparrow$. To show that the composition $\mu^{-1} \circ \mu$ is the identity over $P \times G$, we will need to recall the property $\phi_{uv}(u) = v$, for all $u, v \in P$. Also, to make things a bit more clear, we will observere beforehand that $\mu(v, g)(e_1) = (\phi_{e_1 v} \circ g)(e_1) = \phi_{e_1 v}(e_1) = v$. Then,

$$(\mu^{-1} \circ \mu)(v, g) = (\mu(v, g)(e_1), \phi_{e_1 \mu(v,g)(e_1)}^{-1} \circ \mu(v, g)) = (v, \phi_{e_1 v}^{-1} \circ \phi_{e_1 v} \circ g) = (v, g),$$

for all $v \in P$ and $g \in G$. Therefore, μ and μ^{-1} are inverse continuous maps, hence μ is a homeomorphism.

Notice that any connected clopen set in a topological space different from the empty set and the whole space must be a connected component. Now the restricted Lorentz group is clopen, and it contains the identity, so it must be equal to the connected component of the identity. □

From the construction we made for this proof we can get some other results with a rather physical flavour. We remind that from the point of view of special relativity, it is important to determine how coordinates of events are transformed from one inertial reference frame to an other. Restricting our attention to only transformations of coordinates of events between inertial reference frames with a shared origin in spacetime, it is not hard to see that these have a group structure under composition, and we claim that this group is indeed $\Gamma_+^\uparrow$. Note that this is not completely obvious. On the one hand it is clear that any transformation of coordinates between two IRFs with the same origin is going to preserve the quadratic form q, the direction of time and the orientation of space. But conversely it is not straightforward that any linear map $\mathbb{R}^4 \to \mathbb{R}^4$ preserving the quadratic form q, the direction of time and the orientation of space will correspond to the transformation of coordinates between two IRFs with a common origin. Take a look at the following proposition, which pretends to make this result a little bit more clear.

Proposition 2.11. Any element in $\Gamma_+^\uparrow$ can be put as the composite of a boost in a space-like direction and a rotation of space.

Proof. Consider the isomorphism $\mu : P \times G \longrightarrow \Gamma_+^\uparrow$ defined in the proof of last proposition. An element in the domain $P \times G$ consists of a pair (v, g) where v is a time-like vector, and g is a rotation of the space-like basis formed by e_2, e_3 and e_4. Let x be a linear combination of the space-like basis e_2, e_3, e_4. Then if we set $v = (\sqrt{1 + \|x\|^2}, x) \in \mathbb{R}^4$, the map $\phi_{e_1,v} : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ is a boost in the space-like direction of x.
The map μ is telling us that a vector in space and a rotation of space combined form an element of $\Gamma_+^\uparrow$, as we already knew. The fact that μ is an isomorphism (so in particular μ is surjective) gives us the desired result.

□

Not only Γ is a topological space, but in fact it is a topological group. This is given by the fact that multiplication is continuous, since the entries of AB are no other than polynomials in the entries of A and B, if $A, B \in \Gamma$. Also, inversion is continuous by Cramer's rule. In fact, multiplication and inversion are both smooth operations, so when we give a smooth structure on Γ we will already be able to say that the Lorentz group is a Lie group.

Proposition 2.12. The Lorentz group is a smooth submanifold of $M_{4\times4}(\mathbb{R})$ of dimension 6.

Proof. Call M the vector space of all 4×4 real matrices. Consider the subspaces

$$U = \{B \in M \mid JBJ = -B^T\}$$

$$V = \{B \in M \mid JBJ = B^T\}.$$

It is actually an easy check to see that, indeed, U and V are vector subspaces of M. We aim to show that, as vector subspaces, U and V have dimension 6 and 10, respectively.

In order to do that, we will write down a generic matrix in $M = M_{4\times 4}(\mathbb{R})$, and we will observe how the matrix equations defining the subspaces U and V translate into conditions to the entries of our general matrix.

If we put $B = (b_i^j) \in M$ we realise that $B \in U$ if and only if B is of the form

$$B = \begin{pmatrix} 0 & b_2^1 & b_3^1 & b_4^1 \\ b_2^1 & 0 & b_3^2 & b_4^2 \\ b_3^1 & -b_3^2 & 0 & b_4^3 \\ b_4^1 & -b_4^2 & -b_4^3 & 0 \end{pmatrix},$$

with $b_i^j \in \mathbb{R}$ for all $1 \leq j < 4$ and $j < i \leq 4$. Similarly, a matrix $A = (a_i^j) \in M$ belongs to V if and only if it is one of the form

$$A = \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 & a_4^1 \\ -a_2^1 & a_2^2 & a_3^2 & a_4^2 \\ -a_3^1 & a_3^2 & a_3^3 & a_4^3 \\ -a_4^1 & a_4^2 & a_4^3 & a_4^4 \end{pmatrix},$$

with $a_i^j \in \mathbb{R}$ for all $1 \leq j \leq 4$ and $j \leq i \leq 4$. Once we have seen this, it is quite straight forward to see that $\dim U = 6$ and $\dim V = 10$. Also, since $U \cap V = \{0\}$, we can deduce that $M = U \oplus V$. We are now going to construct a smooth map $g : M \longrightarrow V$ with the intention of applying the regular value theorem. Put

$$g(A) = JA^T JA - \mathrm{I},$$

we must show g is well defined, i.e. for all $A \in M$ we have $g(A) \in V$. This is kind of a straight forward proof, just need to check $Jg(A)J = g(A)^T$, for all $A \in M$. Indeed,

$$J(JA^T JA - \mathrm{I})J = A^T JAJ - J^2 = A^T JAJ - \mathrm{I} = (JA^T JA - \mathrm{I})^T.$$

Observe that $\Gamma = g^{-1}(\{0\})$. We claim that $0 \in V$ is a regular value of g, i.e. $d_A g : T_A M \longrightarrow T_0 V$ is surjective for all $A \in g^{-1}(\{0\})$. Because V is diffeomorphic to $\mathbb{R}^{10}$, we can just say $T_0 V \cong V$ naturally. So what we are about to show now is the following statement. If we fix any $C \in V \cong T_0 V$, then we can find a smooth path $\gamma : I \longrightarrow M$ such that $\gamma(0) = A$ and $(d_A g)(\gamma'(0)) = C$, for all $A \in \Gamma$. That would indeed prove $d_A g$ is surjective.

Given $A \in \Gamma$ and $C \in V$, define $\gamma(t) = A + \frac{t}{2} AC$, a smooth path in M passing through A at $t = 0$. Now

$$(d_A g)(\gamma'(0)) = \frac{d}{dt}(g \circ \gamma)\bigg|_{t=0}.$$

It would be really convenient to work out an expression for $(g \circ \gamma)(t)$, a little bit of work shows that

$$(g \circ \gamma)(t) = J\left(A + \frac{t}{2}AC\right)^T J\left(A + \frac{t}{2}AC\right) - \mathrm{I} =$$

$$JA^T JA - \mathrm{I} + \frac{t}{2}(JA^T JAC + JC^T A^T JA) + \mathcal{O}(t^2).$$

Recall $A \in \Gamma$, so $JA^T JA - \mathrm{I}$ in the previous expression vanishes. Let us call $\mathcal{O}(t^2)$ the quadratic terms in the expression above. Taking into account the facts that $JA^T JA = \mathrm{I}$ and $JC^T = CJ$, falling from $A \in \Gamma$ and $C \in V$ respectively, we can make the following simplifications:

$$(g \circ \gamma)(t) = tC + \mathcal{O}(t^2).$$

Hence we get that $(d_A g)(\frac{1}{2}AC) = C$, for all $C \in V$ and $A \in \Gamma$. Now it is clear that $d_A g$ is surjective.

We can apply the regular value theorem, which assures us that $\Gamma = g^{-1}(\{0\})$ is a smooth submanifold of $M = M_{4\times4}(\mathbb{R})$ of dimension $\dim M - \dim V = 16 - 10 = 6$.

$\square$

3 Structure of the Restricted Lorentz Group

We are interested in the group structure of $\Gamma_+^\uparrow$. The main theorem in this section will take some work to prove, but will certainly help in understanding the structure of the restricted Lorentz group. Before stating it we need some definitions.

Definition 3.1. We will refer to the group of 2×2 invertible complex matrices as $\mathrm{GL}_2(\mathbb{C})$. This usually called the *general linear group* of degree 2 over $\mathbb{C}$. It is useful to introduce the notation for the subgroup of $\mathrm{GL}_2(\mathbb{C})$ consisting of the non-zero complex multiples of the identity matrix. So we put $Z = \{\lambda\,\mathrm{I} \mid \lambda \in \mathbb{C}^\times\}$.

Definition 3.2. We will denote the group of 2×2 complex matrices with determinant equal to 1 as $\mathrm{SL}_2(\mathbb{C})$. It is widely known as the *special linear group* of degree 2 over $\mathbb{C}$. Let SZ be the normal subgroup of $\mathrm{SL}_2(\mathbb{C})$ generated by $-\mathrm{I} \in \mathrm{SL}_2(\mathbb{C})$, this is $SZ = \{\mathrm{I}, -\mathrm{I}\}$. Observe that $SZ = \mathrm{SL}_2(\mathbb{C}) \cap Z$.

Definition 3.3. We can define $\mathrm{PSL}_2(\mathbb{C})$ to be the quotient group $\mathrm{SL}_2(\mathbb{C})/SZ$. This group is called the *projective linear group* of degree 2 over $\mathbb{C}$.

Notice that we can also say $\mathrm{PSL}_2(\mathbb{C}) = \mathrm{GL}_2(\mathbb{C})/Z$. Let us demonstrate why can we say this. The proof is basically supported on the second isomorphism theorem for groups. Observe that $\mathrm{GL}_2(\mathbb{C}) = \mathrm{SL}_2(\mathbb{C})Z$, since any matrix $A \in \mathrm{GL}_2(\mathbb{C})$ can be written as $A = (\det A^{-\frac{1}{2}}\ A)\cdot(\det A^{\frac{1}{2}}\ \mathrm{I})$. First factor clearly belongs to $\mathrm{SL}_2(\mathbb{C})$ since has determinant 1. Since $\det A \in \mathbb{C}^\times$, the second factor lies in Z. By the second isomorphism theorem,

$$\mathrm{SL}_2(\mathbb{C})/(\mathrm{SL}_2(\mathbb{C}) \cap Z) \cong \mathrm{SL}_2(\mathbb{C})Z/Z.$$

Hence we get the desired result $\mathrm{SL}_2(\mathbb{C})/SZ \cong \mathrm{GL}_2(\mathbb{C})/Z$.

Now that we have introduced the notation for the most commonly used groups in the following section, we are ready to state what is going to be one of the big goals in this project.

Theorem 3.4. The restricted Lorentz group $\Gamma_+^\uparrow$ is isomorphic to $\mathrm{PSL}_2(\mathbb{C})$.

In order to build such isomorphism of groups, we will need to introduce new spaces. The idea underlying all the technical process that is about to come is the following. We will take a four dimensional vector space V over $\mathbb{R}$, and we will see that $\mathrm{PSL}_2(\mathbb{C})$ acts on it. This means, for every matrix in $\mathrm{PSL}_2(\mathbb{C})$ the action will give us an automorphism of V. Because, as we will prove later, the space V is isomorphic to $\mathbb{R}^4$, automorphisms of V are in one to one correspondence with automorphisms of $\mathbb{R}^4$, i.e. non-singular matrices. We will see that the automorphisms of $\mathbb{R}^4$ that arise from considering the group action of $\mathrm{PSL}_2(\mathbb{C})$ in the set V preserve the bilinear form η, hence are elements of Γ.

3.1 The Action of $\mathrm{PSL}_2(\mathbb{C})$ on the Space of Hermitian Matrices

The space V introduced in the last paragraph of the previous section is indeed the four dimensional vector space over the real numbers of the order 2 Hermitian matrices. It is made more precise in the following definition.

Definition 3.5. The set of 2×2 Hermitian matrices V is defined to be

$$V = \{A \in M_{2\times 2}(\mathbb{C}) \mid A^\dagger = A\},$$

where we are using the notation $\dagger$ with the following meaning: for any $A = (a_{ij}) \in M_{2\times 2}(\mathbb{C})$, we set $A^\dagger$ to mean the complex conjugate transpose, i.e. $A^\dagger = \overline{A^T} = (\overline{a}_{ji})$.

It is an easy check that Hermitian complex 2×2 matrices have structure of $\mathbb{R}$-vector space of dimension 4. Indeed the following linear map describes an isomorphism between $\mathbb{R}^4$ and V. Consider

$$p : \mathbb{R}^4 \longrightarrow V,$$

defined as

$$p\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t+x & y+iz \\ y-iz & t-x \end{pmatrix}, \quad \text{for all } \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \in \mathbb{R}^4.$$

The map p is clearly linear. In order to prove that it is in fact an isomorphism of $\mathbb{R}$-vector spaces, we provide of an inverse, i.e. a linear map $q : V \longrightarrow \mathbb{R}^4$ such that $p \circ q = \mathrm{id}_V$ and $q \circ p = \mathrm{id}_{\mathbb{R}^4}$. Let q be defined by

$$q\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{2}\begin{pmatrix} a+d \\ a-d \\ b+c \\ i(c-b) \end{pmatrix}, \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V.$$

Thus V and $\mathbb{R}^4$ are isomorphic as $\mathbb{R}$-vector spaces.

We are now ready to see in which way we are going to obtain an action from the group $\mathrm{PSL}_2(\mathbb{C})$ on V. Let $U \in \mathrm{GL}_2(\mathbb{C})$ and $A \in V$. Let us define an operation $* : \mathrm{GL}_2(\mathbb{C}) \times V \longrightarrow V$ by setting

$$U * A = \frac{1}{|\det U|} \cdot UAU^\dagger.$$

We check it is in fact well defined, i.e. for all $U \in \mathrm{GL}_2(\mathbb{C})$ and $A \in V$, we must see that $U * A \in V$.

$$(U * A)^\dagger = \left(\frac{1}{|\det U|} \cdot UAU^\dagger\right)^\dagger = \frac{1}{|\det U|} \cdot U^{\dagger\dagger}A^\dagger U^\dagger = \frac{1}{|\det U|} \cdot UAU^\dagger = U * A,$$

where we have used the fact that $U^{\dagger\dagger} = U$ for any $U \in M_{2\times 2}(\mathbb{C})$ and $A^\dagger = A$ by assumption.

Proposition 3.6. The map $* : \mathrm{GL}_2(\mathbb{C}) \times V \longrightarrow V$ is a group action of the matrix group $\mathrm{GL}_2(\mathbb{C})$ on the space of 2×2 complex Hermitian matrices. This action factors through $\mathrm{PSL}_2(\mathbb{C})$, and in fact gives an injection from $\mathrm{PSL}_2(\mathbb{C})$ to $\mathrm{Aut}(V)$.

Proof. We check that the two conditions we need for $*$ to be a group action are satisfied. In the first place, notice that $\mathrm{I} * A = IAI^\dagger = A$, for all $A \in V$. The second axiom that need to be checked is the one that follows. For all $U, W \in \mathrm{GL}_2(\mathbb{C})$ we have that $(UW) * A = U * (W * A)$. Indeed, we have

$$(UW) * A = \frac{1}{|\det UW|} \cdot (UW)A(UW)^\dagger = \frac{1}{|\det U||\det W|} \cdot U(WAW^\dagger)U^\dagger =$$

$$= \frac{1}{|\det U|} \cdot U(W * A)U^\dagger = U * (W * A).$$

So we get the desired result, this is that $*$ is a group action of $\mathrm{GL}_2(\mathbb{C})$ on V. It is important to remark that, and this is a general result for group actions on sets, for every $U \in \mathrm{GL}_2(\mathbb{C})$ we get an isomorphism ϕ_U from V to V (i.e. an automorphism) sending A to $\phi_U(A) = U * A$, for every $A \in V$. Because of this result, we will think of the action $*$ as a homomorphism $\mathfrak{A} : \mathrm{GL}_2(\mathbb{C}) \longrightarrow \mathrm{Aut}(V)$. Notice that smoothness is guaranteed since the domain of $\mathfrak{A}$ is the invertible 2×2 complex matrices, therefore the complex modulus of the determinant never vanishes.

Observe that, Z acts trivially on V. This means, if $X \in Z$ we have $\mathfrak{A}(X)$ is the identity on V. We check this by letting $X = \lambda\,\mathrm{I}$, with $\lambda \in \mathbb{C}^\times$, and $A \in V$

$$\mathfrak{A}(X)(A) = X * A = \frac{1}{|\det X|} \cdot XAX^\dagger = \frac{1}{|\lambda^2|} \cdot \lambda\,\overline{\lambda}\,A = A.$$

So the action factors through $\mathrm{PSL}_2(\mathbb{C})$, i.e. we have group homomorphisms

$$\begin{array}{rcccccc} \mathfrak{A} : & \mathrm{GL}_2(\mathbb{C}) & \xrightarrow{\mathfrak{P}} & \mathrm{PSL}_2(\mathbb{C}) & \xrightarrow{\tilde{\mathfrak{A}}} & \mathrm{Aut}(V) \\ & U & \longmapsto & [U] & \longmapsto & \phi_U, \end{array}$$

where $\mathfrak{P}$ is the canonical projection, and $\tilde{\mathfrak{A}}$ is the well defined homomorphism induced by the action of PSL_2 on V, such that $\mathfrak{A} = \tilde{\mathfrak{A}} \circ \mathfrak{P}$.

In fact, we show now a stronger statement. We want to prove that the obtained homomorphism $\tilde{\mathfrak{A}}$ is injective. For that we claim $\mathrm{Ker}\,\mathfrak{A} = Z$. In order to show this, we want to consider the set of $U \in \mathrm{GL}_2(\mathbb{C})$ such that $\mathfrak{A}(U) = \phi_U = \mathrm{id}_V$, i.e. $U * A = A$, for all

$A \in V$. We choose a basis $(A_i)_{1\leq i\leq 4}$ of V and discard those matrix in $U \in \mathrm{GL}_2(\mathbb{C})$ that do not satisfy $U * A_i = A_i$ for some i. Conversely, those matrix $U \in \mathrm{SL}_2(\mathbb{C})$ that satisfy $U * A_i = A_i$ for all i, will indeed be elements in $\operatorname{Ker}\mathfrak{A}$. The basis we choose for V is $A_i = p(e_i)$. More explicitly,

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The condition $U * A_1 = A_1$ is equivalent to imposing $UU^\dagger = |\det U|\,\mathrm{I}$. Notice that this can be equivalently reformulated as $U^\dagger = |\det U|U^{-1}$. Observe for a moment that this implies that for any $A \in V$, $U * A = A$ and $UA = AU$ are equivalent. Let us check

$$U * A = \frac{1}{|\det U|} \cdot UAU^\dagger = \frac{1}{|\det U|} \cdot UA\;|\det U|\;U^{-1} = UAU^{-1} = A,$$

if and only if $AU = UA$. So, imposing $U * A_4 = A_4$ is the same as saying $UA_4 = A_4U$, and from this condition we see that $U \in \mathrm{GL}_2(\mathbb{C})$ must be of the form

$$U = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix},$$

for some $a, b \in \mathbb{C}^\times$. Moreover, the condition $UA_2 = A_2U$ implies $a = b$. Thefore $U \in Z$, and we just proved that if we take any $U \in Z$ it happens that $U * A_i = A_i$, for all i. Because A_i forms a basis of V, we get $U * A = A$ for all $A \in V$. Hence we just proved $\operatorname{Ker}\mathfrak{A} \subset Z$, therefore $\operatorname{Ker}\mathfrak{A} = Z$. Applying the first isomorphism theorem we get the desired result. $\square$

Now, because we are ultimately interested in describing an isomorphism between $\mathrm{PSL}_2(\mathbb{C})$ and the restricted Lorentz group (which is no other than a subgroup of the linear group $\mathrm{GL}_4(\mathbb{R})$), we will focus on the relation between the automorphisms of V and those of $\mathbb{R}^4$. Indeed, we will show that it is essentially the same to talk about automorphisms of V and automorphisms of $\mathbb{R}^4$, i.e. elements of $\mathrm{GL}_4(\mathbb{R})$. Recall the isomorphism we built before $p : \mathbb{R}^4 \longrightarrow V$, we have the following diagram

$$\begin{tikzcd} V \arrow[r, "\alpha"] & V \arrow[d, "q"] \\ \mathbb{R}^4 \arrow[u, "p"] \arrow[r, "\beta"] & \mathbb{R}^4 \end{tikzcd}$$

where $\beta = q \circ \alpha \circ p$. Remember q was defined to be p^{-1}. Since the composition of isomorphisms is an isomorphism, we get that, for every $\alpha \in \mathrm{Aut}(V)$, $q \circ \alpha \circ p$ is an element of $\mathrm{GL}_4(\mathbb{R})$. So we get a linear map $q_* : \mathrm{Aut}(V) \longrightarrow \mathrm{GL}_4(\mathbb{R})$. We can also define $p_* : \mathrm{GL}_4(\mathbb{R}) \longrightarrow \mathrm{Aut}(V)$ linear map by setting $p_*(\beta) = p \circ \beta \circ q$, for every $\beta \in \mathrm{GL}_4(\mathbb{R})$.

We can easily check that $q_* \circ p_*$ is the identity on $\mathrm{GL}_4(\mathbb{R})$, and $p_* \circ q_*$ is the identity on $\mathrm{Aut}(V)$. So q_* provides an isomorphism of groups between $\mathrm{Aut}(V)$ and $\mathrm{GL}_4(\mathbb{R})$. Moreover, q_* and p_* are linear maps, therefore smooth maps between manifolds. Therefore q_* is a diffeomorphism between $\mathrm{Aut}(V)$ and $\mathrm{GL}_4(\mathbb{R})$.

Now that we know that $\mathrm{GL}_4(\mathbb{R})$ and $\mathrm{Aut}(V)$ are essentially the same (i.e. we have isomorphisms q_* and p_* going from one to another) it is reasonable to ask for a characterisation of $p_*(\Gamma)$, since Γ is contained in $\mathrm{GL}_4(\mathbb{R})$. We fix the notation putting $\Gamma' := p_*(\Gamma)$. We start with the answer, and then we shall prove that it is indeed correct. So our claim is that $\Gamma' = \{\alpha \in \mathrm{Aut}(V) \mid \det \alpha(A) = \det A, \text{ for all } A \in V\}$. Observe that, if $\mathbf{x} = (t, x, y, z)^T \in \mathbb{R}^4$, then

$$\det \begin{pmatrix} t+x & y+iz \\ y-iz & t-x \end{pmatrix} = t^2 - x^2 - y^2 - z^2.$$

In other words, $\det(p(\mathbf{x})) = \mathtt{q}(\mathbf{x})$. Roughly speaking, the isomorphism p translates the quadratic form $\mathtt{q}$ in $\mathbb{R}^4$ into the determinant in V. So, starting with the definition of Γ we see that

$$\Gamma' = \{p_*(M) \text{ where } M \in \mathrm{GL}_4(\mathbb{R}) \mid \mathtt{q}(M\,\mathbf{x}) = \mathtt{q}(\mathbf{x}), \quad \forall\, \mathbf{x} \in \mathbb{R}^4\}.$$

But now this is the same as

$$\{p_*(M) \text{ where } M \in \mathrm{GL}_4(\mathbb{R}) \mid \det(p \circ M)(\mathbf{x}) = \det p(\mathbf{x}), \quad \forall\, \mathbf{x} \in \mathbb{R}^4\}.$$

Recall that $p \circ M = p_*(M) \circ p$. Putting $\alpha = p_*(M)$ and $V = p(\mathbf{x})$ we get

$$\Gamma' = \{\alpha \in \mathrm{Aut}(V) \mid \det \alpha A = \det A, \quad \forall A \in V\}.$$

For the making the work in the future more clear, we shall fix here a notation. We call $\Gamma'^{\uparrow}_{+}$ the image of $\Gamma^{\uparrow}_{+}$ under the isomorphism p_*, i.e. $\Gamma'^{\uparrow}_{+} := p_*(\Gamma^{\uparrow}_{+})$.

Note 3.7. We will often be characterising *spacetime* (usually treated as $\mathbb{R}^4$) as the space of 2×2 Hermitian matrices, V. And the jump from one to the other is simply provided by the isomorphism $p : \mathbb{R}^4 \longrightarrow V$ defined previously. A clear advantage of dealing with spacetime as the space V is that the quadratic form of an element in V is translated into its determinant, which is a well understood and manageable tool. Similarly, many manipulations concerning the Lorentz group are made much more easy if we think of it as a subset of $\mathrm{Aut}(V)$ instead of viewing it in the traditional way; this is, as a subset of $\mathrm{Aut}(\mathbb{R}^4)$. And there is no problem in going from one point of view to the other since we have an isomorphism $p_* : \mathrm{Aut}(\mathbb{R}^4) \longrightarrow \mathrm{Aut}(V)$ which allows us to do so. During the rest of the text, we have tried to make it clear when we are dealing with subsets of $\mathbb{R}^4$ or their corresponding subsets of V. The notation used is always the same. If we let $\mathcal{A} \subset \mathbb{R}^4$ then we call $\mathcal{A}' = p(\mathcal{A}) \subset V$. On the other hand, if we have $\mathcal{B} \subset \mathrm{Aut}(\mathbb{R}^4)$ then we call $\mathcal{B}' = p_*(\mathcal{B}) \subset \mathrm{Aut}(V)$. Hopefully with this note we will prevent confusion in further reading.

The preceding work becomes now useful, when we intend to prove that there is a group homomorphism from $\mathrm{PSL}_2(\mathbb{C})$ to Γ. We formalise this result in the following proposition.

Proposition 3.8. The image of the group homomorphism $\mathfrak{A} : \mathrm{GL}_2(\mathbb{C}) \longrightarrow \mathrm{Aut}(V)$, is contained in Γ'.

Proof. We have to show that for all $U \in \mathrm{GL}_2(\mathbb{C})$ and $A \in V$, $\mathfrak{A}(U)(A)$ is an Hermitian matrix and $\det \mathfrak{A}(U)(A) = \det A$. Firstly, because we already know $U * A \in V$ and V is closed under real multiplication, we easily get that $\mathfrak{A}(U)(A) \in V$. Let us see that $\mathfrak{A}(U) \in \mathrm{End}(V)$ preserves the determinant:

$$\det \mathfrak{A}(U)(A) = \frac{\det(UAU^\dagger)}{|\det U|^2} = \frac{\det U \det A \det U^\dagger}{|\det U|^2} = \frac{\det U \det A\ \overline{\det U}}{|\det U|^2} = \det A.$$

Note that that we have taken into account that $\mathfrak{A}(U)(A)$ is a 2×2 matrix, therefore constants jump out of the determinant with a square. We have also used the fact that $\det U^\dagger = \overline{\det U}$, for all U size 2 square complex matrix. $\square$

Proposition 3.9. We can define an injective group homomorphism from $\mathrm{PSL}_2(\mathbb{C})$ to Γ.

Proof. From the proposition above $\mathfrak{A}(U) = \phi_U \in \Gamma'$, where $U \in \mathrm{GL}_2(\mathbb{C})$. Then, when composing with q_* we get that $(q_* \circ \mathfrak{A})(U) \in \Gamma$ for all $U \in \mathrm{GL}_2(\mathbb{C})$. And so, we get a group homomorphism from $\mathrm{GL}_2(\mathbb{C})$ to Γ by composing $\mathfrak{A}$ with q_*,

$$\mathrm{GL}_2(\mathbb{C}) \quad \xrightarrow{\mathfrak{A}} \quad \Gamma' \quad \xrightarrow{q_*} \quad \Gamma.$$

Moreover, by the last proposition, we have an injective homomorphism from $\mathrm{PSL}_2(\mathbb{C})$ to Γ

$$\mathrm{PSL}_2(\mathbb{C}) \quad \xrightarrow{\tilde{\mathfrak{A}}} \quad \Gamma' \quad \xrightarrow{q_*} \quad \Gamma.$$

$\square$

3.2 Using Smooth Structure of Groups to Show the Result

In order to prove that $\Gamma_+^\uparrow$ is isomorphic to $\mathrm{PSL}_2(\mathbb{C})$ we will need to work out some of their properties as smooth manifolds.

Lemma 3.10. $\mathrm{PSL}_2(\mathbb{C})$ and Γ are smooth manifolds of same dimension.

Proof. We have already shown that the dimension of Γ is 6. Now in this proof we aim to show that $\mathrm{PSL}_2(\mathbb{C})$ has also dimension 6. Recall the definition $\mathrm{PSL}_2\,\mathbb{C} = \mathrm{SL}_2\,\mathbb{C}/\langle -\mathrm{I}\rangle$. We are going to show that $\dim \mathrm{PSL}_2(\mathbb{C}) = \dim \mathrm{SL}_2(\mathbb{C})$.Take any U open neighbourhood of $\mathrm{I} \in \mathrm{SL}_2(\mathbb{C})$ with $U \cap (-U) = \emptyset$, for example

$$U = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{C}) \mid \Re(a) > 0 \right\}.$$

Then the quotient map $\mathfrak{p} : \mathrm{SL}_2(\mathbb{C}) \longrightarrow \mathrm{PSL}_2(\mathbb{C})$ restricts to a diffeomorphism $\mathfrak{p}|_U : U \longrightarrow \mathfrak{p}(U)$. Notice that $\mathfrak{p}(U)$ is an open neighbourhood of $\mathrm{I} \in \mathrm{PSL}_2(\mathbb{C})$, which is at the same time a 6-manifold since it comes from the image under a diffeomorphism of an open subset of $\mathrm{SL}_2(\mathbb{C})$. Because left multiplication in a Lie group is a diffeomorphism, we get that any $A \in \mathrm{PSL}_2(\mathbb{C})$ has a neighbourhood homeomorphic to $\mathbb{R}^6$, since $A \cdot \mathfrak{p}(U)$ is a 6-manifold containing A. Hence, $\dim \mathrm{PSL}_2(\mathbb{C}) = 6$.

So it suffices to show that $\dim \mathrm{SL}_2(\mathbb{C}) = 6$. Consider the smooth map $f = \det : M_{2\times 2}(\mathbb{C}) \longrightarrow \mathbb{C}$. We compute the derivative of f. Let $A \in M_{2\times 2}(\mathbb{C})$, if we recall that $T_A M_{2\times 2}(\mathbb{C}) \cong M_{2\times 2}(\mathbb{C})$ for all $A \in M_{2\times 2}(\mathbb{C})$ then $d_A f : M_{2\times 2}(\mathbb{C}) \longrightarrow \mathbb{C}$. This can be defined using curves through A. Define $\gamma_X : I \longrightarrow M_{2\times 2}(\mathbb{C})$ as $\gamma(t) = A + tX$, where $X \in M_{2\times 2}(\mathbb{C})$. Then,

$$(d_A f)(X) = \frac{d}{dt}(f \circ \gamma_X)\bigg|_{t=0}.$$

Now using Jacobi's formula for the derivative of the determinant, we get

$$(d_A f)(X) = \operatorname{tr}\left(\operatorname{Ad}\gamma_X \cdot \frac{d\gamma_X}{dt}\right)\bigg|_{t=0} = \operatorname{tr}\left(\operatorname{Ad} A \cdot X\right).$$

Notice that $1 \in \mathbb{C}$ is a regular value of f. This means, that for all $A \in f^{-1}(1) = \mathrm{SL}_2(\mathbb{C})$ the map $d_A f$ is surjective. This shows no major inconvenience. Take any $\alpha \in \mathbb{C}$, and define $X = \frac{\alpha}{2}(\mathrm{Ad}A)^{-1}$. Observe that, in fact we get $(d_A f)(X) = \operatorname{tr}(\frac{\alpha}{2} \cdot \mathrm{I}) = \alpha$.

So as a result we get that $\mathrm{SL}_2(\mathbb{C})$ is a smooth submanifold of $M_{2\times 2}(\mathbb{C})$ of dimension $\dim M_{2\times 2}(\mathbb{C}) - \dim \mathbb{C} = 6$.

□

Proposition 3.11. Let us denote by $\mathfrak{a}$ the restriction of $\mathfrak{A}$ onto $\mathrm{SL}_2(\mathbb{C})$. The map $q_* \circ \mathfrak{a} : \mathrm{SL}_2(\mathbb{C}) \longrightarrow \Gamma$ is an open homomorphism of Lie groups .

To proceed with the proof of this proposition, we need to use some tools concerning the Lie group structure of $\mathrm{PSL}_2(\mathbb{C})$. We will need the following lemma.

Lemma 3.12. Let $\mathcal{G}, H$ be Lie groups of same dimension, and let $f : \mathcal{G} \longrightarrow \mathcal{H}$ be a smooth homomorphism such that $d_g f : T_g\mathcal{G} \longrightarrow T_{f(g)}\mathcal{H}$ is an isomorphism for all $g \in \mathcal{G}$. Then f is open by the inverse function theorem.

Proof. This is actually a general topology fact. Take any open set U in $\mathcal{G}$. The inverse function theorem tells us that, for any $p \in U$ we can find an open neighbourhood of p, call it V_p, and an open neighbourhood of $f(p)$, say $W_f(p)$ such that the restriction of f to V_p gives a homeomorphism onto $W_{f(p)}$. So this is, we have that

$$f|_{V_p} : V_p \longrightarrow W_{f(p)}$$

is a homeomorphism, for every $p \in U$. Notice that we can decompose U as the union of the open sets $V_p \cap U$, for all $p \in U$. That way we see

$$f(U) = f\left(\bigcup_{p\in U} V_p \cap U\right) = \bigcup_{p\in U} f|_{V_p}(V_p \cap U).$$

Recall each $f|_{V_p}$'s is a homeomorphism, which means $f|_{V_p}(V_p \cap U)$ is going to be open in $W_{f(p)}$, hence open in $\mathcal{H}$. So $f(U)$ must be open in $\mathcal{H}$. □

Now our problem turns into showing that $d_g f$ is an isomorphism for all $g \in \mathcal{G}$. We will make use of the next lemma, which will take into account properties arising from the Lie group structure of $\mathcal{G}$ and $\mathcal{H}$.

Lemma 3.13. Let $f : \mathcal{G} \longrightarrow \mathcal{H}$ be a map between Lie groups. If $d_e f$ is an isomorphism between $T_e\mathcal{G}$ and $T_{f(e)}\mathcal{H}$, where e represents the identity element of the group $\mathcal{G}$, then $d_g f$ is an isomorphism for all $g \in \mathcal{G}$.

Proof. Our starting point is our isomorphism of Lie groups $f : \mathcal{G} \longrightarrow \mathcal{H}$, which satisfies that $d_e f : T_e\mathcal{G} \longrightarrow T_{e'}\mathcal{H}$ is an isomorphism. Note that we are writing e, e' for the identity elements of $\mathcal{G}$ and $\mathcal{H}$ respectively. By the inverse function theorem we get an open neighbourhood V of $e \in \mathcal{G}$, such that the restriction of f on V gives a diffeomorphism from V onto its image. Notice that $f(V)$ will be an open neighbourhood of e', since $f|_V$ is a homeomorphism that maps e to e'. We are aiming to find an open neighbourhood

W_g of any given $g \in \mathcal{G}$, such that the restriction $f|_{W_g}$ is a diffeomorphism onto its image. We hope to do that by expressing $f|_{W_g}$ as the composition of certain differomorphisms.

Recall that given any g in $\mathcal{G}$, we define the left multiplication map $L_g : \mathcal{G} \longrightarrow \mathcal{G}$ by setting $L_g(x) = g \cdot x$, and because of the Lie group structure of $\mathcal{G}$, we automatically get that L_g is smooth. Moreover, we can define its inverse map $L_g^{-1} = L_{g^{-1}}$, which is also smooth, and so L_g is a diffeomorphism.

Now, fix an element $g \in \mathcal{G}$. Define $W_g = L_g(V)$, which is an open neighbourhood of g, since we know L_g is a diffeomorphism. We claim that the restriction of f over W_g can be put as the composition of diffeomorphisms, in the following fashion:

$$f|_{W_g} = L_{f(g)} \circ f|_V \circ L_{g^{-1}}.$$

For any $x \in W_g$, we check that the composition is well defined, and because f preserves the group structure, we get that $f(L_{f(g)} \circ f|_V \circ L_{g^{-1}})(x) = f(x)$ indeed. Therefore, f is a diffeomorphism in a neighbourhood of g, hence $d_g f$ is an isomorphism. $\square$

Let us bring all this theory into our specific frame. We have two Lie groups $\mathrm{SL}_2(\mathbb{C})$ and Γ' with a homomorphism $\mathfrak{a}$ going from the former to the latter. Our ultimate wish is to calculate the derivative of $\mathfrak{a}$ at the identity, and see that it gives an isomorphism of vector spaces between $T_{\mathrm{I}} \mathrm{SL}_2(\mathbb{C})$ and $T_{\mathrm{I}}\Gamma'$. Now we will not work with the derivative of $\mathfrak{a}$ since its domain is a rather specific linear subspace of $\mathrm{GL}_2(\mathbb{C})$, which makes it a bit hard to work with. Instead, our plan is to calculate the derivative of $\mathfrak{A}$, a homomorphism of Lie groups between $\mathrm{GL}_2(\mathbb{C})$ and Γ' that extends by definition $\mathfrak{a}$ to $\mathrm{GL}_2(\mathbb{C})$. Therefore we can sum up the situation in this diagram

$$\begin{tikzcd}
\mathrm{GL}_2(\mathbb{C}) \arrow[r, "\mathfrak{A}"] & \Gamma' \\
\mathrm{SL}_2(\mathbb{C}). \arrow[u, "i"] \arrow[ur, "\mathfrak{a}"'] &
\end{tikzcd}$$

We focus now on finding the derivative of $\mathfrak{A}$ at the identity, but in order to do that, we will have to previously work on the domain of such function. In other words, we will have to characterise $T_{\mathrm{I}} \mathrm{GL}_2(\mathbb{C})$. We claim that $T_{\mathrm{I}} \mathrm{GL}_2(\mathbb{C}) \cong M_{2\times 2}(\mathbb{C})$. If for any $X \in M_{2\times 2}(\mathbb{C})$ we find a path through the identity, fully contained in $\mathrm{GL}_2(\mathbb{C})$, such that its derivative at the identity is X, we will see that X is in fact a tangent vector. This would mean $M_{2\times 2}(\mathbb{C}) \subset T_{\mathrm{I}} \mathrm{GL}_2(\mathbb{C})$, and since these two have the same dimension, we would get the equaltiy.
For any $X \in M_{2\times 2}(\mathbb{C})$, define $\gamma : I \longrightarrow \mathrm{GL}_2(\mathbb{C})$ by $\gamma(t) = \mathrm{I} + tX$. It is well defined, since $\det \gamma(t) = 1 + t \operatorname{tr} Xt + t^2 \det X$ and by a continuity argument, we can find an interval containing 0, call it $J \subset I$, for which $\det \gamma(t) > 0$, for all $t \in J$. Therefore, X is a tangent

vector to $\mathrm{I} \in \mathrm{GL}_2(\mathbb{C})$.
In order to determine the codomain of $d_{\mathrm{I}}\mathfrak{A}$, i.e. the tangent space of Γ' at the identity, we will use the following lemma.

Lemma 3.14. Let E be a finite dimensional vector space and M a submanifold of E. Call $j : M \hookrightarrow E$ the inclusion embedding. Then for all $x \in M$, we have that $T_x M \subseteq E$.

Proof. The lemma doesn't actually require a proof, but rather it needs to be explained in a more detailed fashion. We know $T_x M$ and E are both vector spaces, but we need to make more precise what we mean when we say that the former is contained in the latter. Recall that a submanifold of E is actually a pair consisting of a smooth manifold M and an embedding $j : M \hookrightarrow E$. At any $x \in M$, we have the differential $j_x : T_x M \longrightarrow T_{j(x)} E$, which is an injective linear map. Recall that E is a smooth manifold and a finite dimensional vector space at the same time, therefore diffeomorphic to some $\mathbb{R}^n$, and this allows us to say that the tangent space of E at any given point of E will be isomorphic (as a vector space) to E. Hence we will have $T_{j(x)} E \cong E$, for all $x \in M$. Finally, this justifies the fact that we think of elements in $T_x M$ as vectors of E. $\square$

Now this becomes useful when it comes to picturing who will $T_{\mathrm{I}}\Gamma'$ be. Because Γ' is a submanifold of $\mathrm{End}(V)$, we get that $T_{\mathrm{I}}\Gamma' \subset \mathrm{End}(V)$.

We are now ready to calculate the derivative of $\mathfrak{A}$ at the identity. Using the same notation as before, let γ be a smooth curve defined by $\gamma(t) = \mathrm{I} + tX$, where $X \in M_{2\times 2}(\mathbb{C})$. Notice that $(d_{\mathrm{I}}\mathfrak{A})(X)$ is an endomorphism of V, so we give its description by the action on $A \in V$. We calculate

$$(d_{\mathrm{I}}\mathfrak{A})(X)(A) = \left.\frac{d}{dt}\right|_0 (\mathfrak{A} \circ \gamma)(A).$$

By some hand work we figure out an expression for $(\mathfrak{A} \circ \gamma)(A)$. Since A is fixed and the expression will be t dependent, we write it this way:

$$(\mathfrak{A} \circ \gamma)_A(t) = \frac{1}{|\det(\mathrm{I} + tX)|} \cdot (\mathrm{I} + tX)A(I + tX^\dagger).$$

It is pretty straight forward to check that

$$(\mathfrak{A} \circ \gamma)_A(t) = \gamma(t) * A = \frac{1}{|\det(\mathrm{I} + tX)|} \cdot \left[A + (XA + AX^\dagger)t\right] + \mathcal{O}(t^2).$$

Recall we are only interested in the linear term of $(\mathfrak{A} \circ \gamma)_A(t)$, since this will in fact be the derivative of $\mathfrak{A}$ evaluated in the direction X. Let us focus now on finding the polynomial expansion of $|\det(\mathrm{I} + tX)|^{-1}$. As we had seen before,

$$\det(\mathrm{I}+tX) = 1 + \operatorname{tr} Xt + \det Xt^2.$$

We now proceed to find a polynomial expansion for the complex modulus of $\det(\mathrm{I}+tX)$. Observe that, since we are only interested in the linear term, we can omit the quadratic term of the expression above. We introduce now the notation $\approx$ which means equality of functions up to linear terms. So, what we found before is just $\det(\mathrm{I}+tX) \approx 1 + t \operatorname{tr} X$. Now we are going to use the Maclaurin series of $\sqrt{1+x}$ up to linear terms, which states $\sqrt{1+x} \approx 1 + \frac{x}{2}$, in order to find the polynomial expansion of $|\det(\mathrm{I}+tX)|$.

$$|\det(\mathrm{I}+tX)| \approx \sqrt{(1 + \Re(\operatorname{tr} X)t)^2 + (\Im(\operatorname{tr} X)t)^2}$$

We expand what is sitting inside the square root, and we only consider the expression up to linear terms. Therefore, using Maclaurin's series up to linear terms for $\sqrt{1+x}$ we get

$$|\det(I + tX)| \approx 1 + \Re(\operatorname{tr} X)t.$$

We can imagine that t takes values in a rather small neighbourhood of $0 \in \mathbb{R}$ such that $|\Re(\operatorname{tr} X)t| < 1$, for all t. In that case we can apply the Maclaurin series of $\frac{1}{1-x}$ up to linear terms, which is $1 + x$, to get the desired polynomial expansion,

$$\frac{1}{|\det(I + tX)|} \approx \frac{1}{1 + \Re(\operatorname{tr} X)t} \approx 1 - \Re(\operatorname{tr} X)t.$$

And we are done finding the polynomial expansion of $|\det(I + tX)|^{-1}$ up to linear terms. We now put it all together in order to determine the linear term of the polynomial expansion of $(\mathfrak{A} \circ \gamma)_A$.

$$(\mathfrak{A} \circ \gamma)_A(t) = \frac{(\mathrm{I}+tX)A(I + tX^\dagger)}{|\det(\mathrm{I}+tX)|} \approx (1 - \Re(\operatorname{tr} X)t) \cdot (A + (XA + AX^\dagger)t)$$

Now the differential of $\mathfrak{A}$ at the identity in the X direction will simply be the derivative of the right hand of the equation above evaluated at $t = 0$. And so it is now clear that $(d_{\mathrm{I}}\mathfrak{A})(X)(A) = XA + AX^\dagger - \Re(\operatorname{tr} X)A$.

Once we have got to this point it is worth remembering that our aim was to show that the derivative of $\mathfrak{a}$ (not $\mathfrak{A}$) at the identity gave an isomorphism of the tangent spaces $T_{\mathrm{I}} \mathrm{SL}_2(\mathbb{C})$ and $T_{\mathrm{I}}\Gamma'$. But we notice that $\mathrm{SL}_2(\mathbb{C})$ is a submanifold of $\mathrm{GL}_2(\mathbb{C})$, simply embedded in $\mathrm{GL}_2(\mathbb{C})$ by the inclusion map. In this context, it is easily shown that

$$d_{\mathrm{I}}\, \mathfrak{A}|_{T_{\mathrm{I}}\mathrm{SL}_2(\mathbb{C})} = d_{\mathrm{I}}\mathfrak{a},$$

and so we found our way of calculating $d_{\mathrm{I}}\mathfrak{a}$.

Notice this last equality makes sense, since any tangent vector $X \in T_{\mathrm{I}}\mathrm{SL}_2(\mathbb{C})$ can be expressed as the derivative of some smooth curve γ fully contained in $\mathrm{SL}_2(\mathbb{C})$ passing through the identity at $t = 0$. So,

$$(d_{\mathrm{I}}\mathfrak{A})(X) = \left.\frac{d}{dt}\right|_0 (\mathfrak{A} \circ \gamma) = \left.\frac{d}{dt}\right|_0 (\mathfrak{a} \circ \gamma) = (d_{\mathrm{I}}\mathfrak{a})(X).$$

So, if we are to study the restriction of $d_{\mathrm{I}}\mathfrak{A}$ onto $T_{\mathrm{I}}\,\mathrm{SL}_2(\mathbb{C})$, we need to know something about $T_{\mathrm{I}}\,\mathrm{SL}_2(\mathbb{C})$. We claim that $T_{\mathrm{I}}\,\mathrm{SL}_2(\mathbb{C})$ is the set of square complex matrices of dimension 2 whose trace is null. Now, we consider the smooth map $f = \det : M_{2\times 2}(\mathbb{C}) \longrightarrow \mathbb{C}$. We know that $T_A\,\mathrm{SL}_2(\mathbb{C}) = \operatorname{Ker} d_A f$, for any $A \in \mathrm{SL}_2(\mathbb{C})$. In particular, taking $A = \mathrm{I}$, we get $T_{\mathrm{I}}\,\mathrm{SL}_2(\mathbb{C}) = \operatorname{Ker} d_{\mathrm{I}} f = \{X \in M_{2\times 2}(\mathbb{C}) \mid \operatorname{tr}(X) = 0\}$. For sake of simplicity, we will call $\mathcal{X}$ this subspace of $M_{2\times 2}(\mathbb{C})$. It is a simple but important observation noticing that $\mathcal{X}$ is a 6-dimensional $\mathbb{R}$-vector space.

Recall that what we were trying to prove is that the linear map $d_{\mathrm{I}}\mathfrak{A} : M_{2\times 2}(\mathbb{C}) \longrightarrow T_{\mathrm{I}}\Gamma'$ restricts to an isomorphism from $\mathcal{X}$ to $T_{\mathrm{I}}\Gamma'$. Our plan to get to this result is to show that $M_{2\times 2}(\mathbb{C}) = \mathcal{X} \oplus \operatorname{Ker} d_{\mathrm{I}}\mathfrak{A}$. In this case, by the first isomorphism theorem, we will get an isomorphism from $M_{2\times 2}(\mathbb{C})/\operatorname{Ker} d_{\mathrm{I}}\mathfrak{A}$ to $\operatorname{Im}\, d_{\mathrm{I}}\mathfrak{A}$. Note that $M_{2\times 2}(\mathbb{C})/\operatorname{Ker} d_{\mathrm{I}}\mathfrak{A} \cong \mathcal{X}$, therefore $\mathcal{X} \cong \operatorname{Im}\, d_{\mathrm{I}}\mathfrak{A}$. Because $\mathcal{X}$ has dimension 6, it follows that $\operatorname{Im}\, d_{\mathrm{I}}\mathfrak{A}$ is a 6-dimensional subspace of $T_{\mathrm{I}}\Gamma'$. Hence $\operatorname{Im}\, d_{\mathrm{I}}\mathfrak{A} = T_{\mathrm{I}}\Gamma'$, and we just showed

$$\mathcal{X} = T_{\mathrm{I}}\mathrm{SL}_2\mathbb{C} \cong T_{\mathrm{I}}\Gamma'.$$

So we just need to find $\operatorname{Ker} d_{\mathrm{I}}\mathfrak{A}$ and show $M_{2\times 2}(\mathbb{C}) = \mathcal{X} \oplus \operatorname{Ker} d_{\mathrm{I}}\mathfrak{A}$. We will sum up these results in the following lemma.

Lemma 3.15. The kernel of $d_{\mathrm{I}}\mathfrak{A}$ is a 2-dimensional subspace of $M_{2\times 2}(\mathbb{C})$, given by

$$\operatorname{Ker} d_{\mathrm{I}}\mathfrak{A} = \left\{ \begin{pmatrix} x+yi & 0 \\ 0 & x+yi \end{pmatrix} : x, y \in \mathbb{R} \right\}.$$

Moreover, $M_{2\times 2}(\mathbb{C})$ can be decomposed as the direct sum of the subspaces $\mathcal{X}$ and $\operatorname{Ker} d_{\mathrm{I}}\mathfrak{A}$.

Proof. We are going to start with a general 2×2 complex matrix

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times 2}(\mathbb{C}),$$

and we will suppose that $X \in \operatorname{Ker} d_{\mathrm{I}}\mathfrak{A}$. We take an Hermitian matrix A and impose $(d_{\mathrm{I}}\mathfrak{A})(X)(A) = 0$, equation from which we will extract restrictions for the coefficients of X. By letting

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in V,$$

and imposing $XA + AX^{\dagger} - \Re(\operatorname{tr} X)A = 0$, we get that $c = 0$ and $\Re(a) = \Re(d)$. Similarly, if we fix

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in V,$$

and proceed in an analogous way, we get a new restriction for the matrix entries of X, specifically we get $b = 0$. Once we have got to this point, it is worth assuming that our general matrix X in the kernel of $d_{\mathrm{I}}\mathfrak{A}$ is of the form

$$X = \begin{pmatrix} x + yi & 0 \\ 0 & x + zi \end{pmatrix} : \quad x, y, z \in \mathbb{R}.$$

By letting $(d_{\mathrm{I}}\mathfrak{A})(X)$ act on a general Hermitian matrix $A = (a_i^j) \in V$, we see if we extract any other restriction for the coefficients of X from the equation $(d_{\mathrm{I}}\mathfrak{A})(X)(A) = 0$. After doing some calculations, it can be seen that

$$(d_{\mathrm{I}}\mathfrak{A})(X)(A) = \begin{pmatrix} 0 & a_2^1(y - z) \\ a_1^2(z - y) & 0 \end{pmatrix}.$$

It is clear that $(d_{\mathrm{I}}\mathfrak{A})(X)(A) = 0$ for all $A \in V$ iff $y = z$. That way we just proved that a complex 2×2 matrix is in $\operatorname{Ker} d_{\mathrm{I}}\mathfrak{A}$ iff it is of the form

$$X = \begin{pmatrix} x + yi & 0 \\ 0 & x + yi \end{pmatrix} : \quad x, y \in \mathbb{R}.$$

Now the rest of the lemma is quite obvious once we have got to this point. We only need to observe that the trace of a nonzero matrix in $\operatorname{Ker} d_{\mathrm{I}}\mathfrak{A}$ will never be zero, therefore $\mathcal{X} \cap \operatorname{Ker} d_{\mathrm{I}}\mathfrak{A} = \{0\}$. Arguing by dimensionality of the subspaces we get what we wanted,

$$M_{2\times 2}(\mathbb{C}) = \mathcal{X} \oplus \operatorname{Ker} d_{\mathrm{I}}\mathfrak{A}.$$

$\square$

So we have finally shown $d_{\mathrm{I}}\mathfrak{a} = d_{\mathrm{I}}\mathfrak{A}|_{T_{\mathrm{I}}\mathrm{SL}_2(\mathbb{C})}$ is an isomorphism of vector spaces. Therefore $\mathfrak{a} : \mathrm{SL}_2(\mathbb{C}) \longrightarrow \Gamma'$ is an open homomorphism of Lie groups. By composing with q_* we get that $q_* \circ \mathfrak{a} : \mathrm{SL}_2(\mathbb{C}) \longrightarrow \Gamma$ is an open homomorphism of Lie groups.

We need to put all these things together in order to show the isomorphism between $\mathrm{PSL}_2(\mathbb{C})$ and $\Gamma_+^\uparrow$. So far, we have the following diagram,

$$\begin{tikzcd} \mathrm{SL}_2(\mathbb{C}) \arrow[r, "\mathfrak{a}"] \arrow[d, "\mathfrak{p}"'] & \Gamma' \arrow[r, "q_*"] & \Gamma \\ \mathrm{PSL}_2(\mathbb{C}) \arrow[ur, "\tilde{\mathfrak{A}}"'] & & \end{tikzcd}$$

What do we know about the maps $q_* \circ \mathfrak{a}$ and $q_* \circ \tilde{\mathfrak{A}}$? Well, as we just proved, $q_* \circ \mathfrak{a}$ is an open homomorphism between Lie groups. We wish to extend this to $q_* \circ \tilde{\mathfrak{A}}$, meaning we would like to say that $q_* \circ \tilde{\mathfrak{A}}$ is a open homomorphisms between Lie groups. We already know $q_* \circ \tilde{\mathfrak{A}}$ is a smooth map preserving group structure, so just need to check openness. From the very definition of the *quotient topology* (defined on $\mathrm{PSL}_2(\mathbb{C})$ in our case), we get that, if U is an open set in $\mathrm{PSL}_2(\mathbb{C})$, then $\mathfrak{p}^{-1}(U)$ is open in $\mathrm{SL}_2(\mathbb{C})$. Taking the image of U open set in $\mathrm{PSL}_2(\mathbb{C})$ by $q_* \circ \tilde{\mathfrak{A}}$,

$$(q_* \circ \tilde{\mathfrak{A}})(U) = (q_* \circ \mathfrak{a})(\mathfrak{p}^{-1}(U)),$$

we realise that it must be open, since we just saw $q_* \circ \mathfrak{a}$ is an open map.

We sum up the situation until now: we just got an injective open homomorphism $q_* \circ \tilde{\mathfrak{A}}$ between $\mathrm{PSL}_2(\mathbb{C})$ and Γ. Because this is going to be of vital importance in our work, we give such homomorphism a proper name, say $\Psi = q_* \circ \tilde{\mathfrak{A}}$. It is a quite standard result from general topology that an injective open map between two topological spaces is a topological *embedding*. In other words, if X and Y are topological spaces, and $f : X \longrightarrow Y$ is an injective continuous open map, then the induced map $\mathrm{f} : X \longrightarrow f(X)$ is a homeomorphism. In our case, Ψ is a topological embedding, which means it is a homeomorphism onto its image.

We now use another result from general topology, that states that connected components are preserved under homeomorphisms. To make this proposition a bit more specific, look at it in our context. We just proved the homeomorphism $\mathrm{PSL}_2(\mathbb{C}) \cong \Psi(\mathrm{PSL}_2(\mathbb{C}))$, which means Ψ sends connected components in $\mathrm{PSL}_2(\mathbb{C})$ to connected components in Γ. But $\mathrm{PSL}_2(\mathbb{C})$ is connected since it is a topological quotient of a connected space, namely $\mathrm{SL}_2(\mathbb{C})$. Therefore, Ψ must send $\mathrm{PSL}_2(\mathbb{C})$ to a connected component of Γ. Which one should it be? It is clear that it ought to be the connected component of an element of Γ which belongs to $\Psi(\mathrm{PSL}_2(\mathbb{C}))$. So take $\mathrm{I} \in \Gamma$, which surely belongs to $\mathrm{Im}\,\Psi$ since Ψ is a homomorphism of groups.

Recall we proved that $\Gamma_+^\uparrow$ is the connected component of the identity. This means we have the inclusion $\Psi(\mathrm{PSL}_2(\mathbb{C})) \subseteq \Gamma_+^\uparrow$. But we are willing to show that equality holds. If we show that $\Psi(\mathrm{PSL}_2(\mathbb{C}))$ is closed in Γ we will have shown the equality.

$\Psi(\mathrm{PSL}_2(\mathbb{C}))$ is an open subgroup of Γ, and it follows that every coset of $\Psi(\mathrm{PSL}_2(\mathbb{C}))$ is also open. However, the complement of $\Psi(\mathrm{PSL}_2(\mathbb{C}))$ is the union of all nontrivial cosets, so it is open. It follows that $\Psi(\mathrm{PSL}_2(\mathbb{C}))$ is also closed.

In other words, this is telling us that the image of the map Ψ is the restricted Lorentz group. Summing up everything until now, we get that Ψ is a smooth bijective open homomorphism between $\mathrm{PSL}_2(\mathbb{C})$ and $\Gamma_+^\uparrow$, i.e. an isomorphism of Lie groups. Hence Ψ is the isomorphism we have been chasing,

$$\Psi : \mathrm{PSL}_2(\mathbb{C}) \xrightarrow{\cong} \Gamma_+^\uparrow.$$

4 Conjugacy Classes of the Restricted Lorentz Group

The idea of conjugacy class is somewhat useful in the context of the rotation group $\mathrm{SO}(3,\mathbb{R})$, because it provides a formal background to the intuitive idea of angle. If we think of a rotation of angle θ in three dimensional space, we can picture many types of them. They are going to differ one from another basically by the choice of rotation axis we are making. But we obviously agree that they all have in common the fact that they rotate by the same *amount*, i.e. by an angle of θ. To formalise this, we just need to comprehend what does it really mean for two rotations to have the same *angle*. We would not be surprised if "having the same angle" was an equivalence relation in $\mathrm{SO}(3,\mathbb{R})$. So two rotations $R, L \in \mathrm{SO}(3,\mathbb{R})$ are going to be said to have the same angle when there exists a third rotation $S \in \mathrm{SO}(3,\mathbb{R})$ such that L can be expressed as rotating by S, then applying R, and then undoing the rotation initially carried out by S. We will write it as $R \sim L$, and this condition can be put in a formal statement as

$$R \sim L \;\Leftrightarrow\; \exists S \in \mathrm{SO}(3,\mathbb{R}) \,|\, S^{-1}RS = L.$$

The twiddle relation can be checked to be an equivalence relation. It gives us criteria for classifying rotations in space (indeed any equivalence relation in $\mathrm{SO}(3,\mathbb{R})$ would). The quotient set $\mathrm{SO}(3,\mathbb{R})/\sim$ gives a description of all angles, and basically gives us an idea of the amount of rotation being done.
We mimic the same strategy, but this time applied to the group $\Gamma_+^\uparrow$, hoping that we will get a good intuition on how rotations of spacetime can be different, modding out by transformations of coordinates between two inertial observers.

4.1 Conjugacy Classes in $\mathbf{PSL}_2(\mathbb{C})$

Recall in the vector space $M_{n\times n}(\mathbb{R})$ we can define an equivalence relation $\sim$ as follows. Let $A, B \in M_{n\times n}(\mathbb{R})$, then

$$A \sim B \;\Leftrightarrow\; \exists \text{ invertible } U \in M_{n\times n}(\mathbb{R}) \text{ such that } U^{-1}AU = B.$$

This equivalence relation is called *conjugacy*, and by saying two matrices are conjugate we will mean they are related by $\sim$. Once we define the conjugacy relation, we are mainly interested in finding its conjugacy classes.

In this section we aim to give a description of the conjugacy classes of the restricted Lorentz group. Just to make it clear, we write down what we mean by a conjugacy class of an element A in $\Gamma_+^\uparrow$,

$$[A] = \{B \in \Gamma_+^{\uparrow} \mid \exists U \in \Gamma_+^{\uparrow}\ A = UBU^{-1}\}.$$

Recall last section ended with this isomorphism of Lie groups $\Psi : \mathrm{PSL}_2\,\mathbb{C} \longrightarrow \Gamma_+^{\uparrow}$. This means that from the classification of conjugacy classes in $\mathrm{PSL}_2(\mathbb{C})$ we can obtain all the information about conjugacy classes in $\Gamma_+^{\uparrow}$. Let us review what we already know from conjugacy classes in $\mathrm{SL}_2(\mathbb{C})$, and that will lead to a better understanding of conjugacy classes in $\mathrm{PSL}_2(\mathbb{C})$. If we let

$$D_\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix},\ \lambda \in \mathbb{C}^\times$$

and

$$E_+ = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},\ E_- = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix},$$

we have that any matrix in $\mathrm{SL}_2(\mathbb{C})$ is conjugate to either E_+, E_- or D_λ for some $\lambda \in \mathbb{C}^\times$. To make ourselves clear, observe that we are not saying that the classes of the elements E_+, E_- and D_λ for $\lambda \in \mathbb{C}^\times$ are all different. Indeed, they are all different except for the classes of the elements D_λ and $D_{\frac{1}{\lambda}}$, $\lambda \in \mathbb{C}^\times$, which are the same.

Notice that in $\mathrm{PSL}_2(\mathbb{C})$ the matrices D_λ and $D_{-\lambda}$ become equal, hence their conjugacy class becomes the same. Observe also that E_+ and E_- lie in the same conjugacy class in $\mathrm{PSL}_2(\mathbb{C})$. This can be easily checked by letting

$$U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and checking that $UE_+U^{-1} = -E_-$. Recall in $\mathrm{PSL}_2(\mathbb{C})$ the matrices E_- and $-E_-$ become the same element, which is conjugate to E_+. So, essentially the conjugacy classes in $\mathrm{PSL}_2(\mathbb{C})$ are the ones with representatives D_λ with $\lambda \in \mathbb{C}^\times$ and E_+. Note that the classes of the form D_λ, $\lambda \in \mathbb{C}^\times$ are all different from the class of E_+, but by no means they are all different one from another.

Note 4.1. Throughout this section we have been managing conjugacy classes in $\mathrm{PSL}_2(\mathbb{C})$ with the following strategy. First, we described the conjugacy classes in $\mathrm{SL}_2(\mathbb{C})$, and then we projected the classes into the quotient $\mathrm{PSL}_2(\mathbb{C})$, hoping that these will in fact be included in some conjugacy classes in $\mathrm{PSL}_2(\mathbb{C})$. It is worth commenting a few lines on why this business works out well. Call $\mathfrak{p} : \mathrm{SL}_2(\mathbb{C}) \twoheadrightarrow \mathrm{PSL}_2(\mathbb{C})$ the canonical projection, we want to see that

$$\mathfrak{p}[A]_{\mathrm{SL}_2(\mathbb{C})} \subset [\mathfrak{p}A]_{\mathrm{PSL}_2(\mathbb{C})},$$

for all $A \in \mathrm{SL}_2(\mathbb{C})$. Pick $B \in \mathfrak{p}[A]_{\mathrm{SL}_2(\mathbb{C})}$, then there exists $U \in \mathrm{SL}_2(\mathbb{C})$ such that $A = UBU^{-1}$. It is pretty clear that in the quotient $\mathfrak{p}(A)$ will be conjugated to $\mathfrak{p}(B)$ since we will have $\mathfrak{p}(A) = \mathfrak{p}(U)\mathfrak{p}(B)\mathfrak{p}(U)^{-1}$.

At this point we could leave it there, and get all the information we want to know about the conjugacy classes in $\Gamma_+^\uparrow$ via the isomorphism $\Psi : \mathrm{PSL}_2(\mathbb{C}) \longrightarrow \Gamma_+^\uparrow$. Nevertheless, if we want to get a better understanding of the physical interpretation of the conjugacy classes, it turns out we have to dig in how the groups $\mathrm{PSL}_2(\mathbb{C})$ and $\Gamma_+^\uparrow$ act on *special* subsets of $\mathbb{C}^2$ and $\mathbb{R}^4$ respectively. These are going to end up being somehow what we call the Riemann sphere and the celestial sphere. The following consists of introducing these two very particular spaces.

4.2 The Celestial Sphere and the Riemann Sphere

Notice that $\mathbb{R}^\times$ acts on the null cone C. This means, for all $\lambda \in \mathbb{R}^\times$ we get an automorphism $C \longrightarrow C$ by sending $x \longmapsto \lambda x$, where $x \in C$. We call a *ray* an equivalence class $[x] = x\,\mathbb{R}^\times$, in other words a ray is an orbit of the action. By the properties of group structure of $\mathbb{R}^\times$, we know that the set of rays or orbits of points in C form a partition of C. The associated equivalence relation is defined by putting $x \propto y$ if and only if there exists a $\lambda \in \mathbb{R}^\times$ such that $y = \lambda x$.

Definition 4.2. The *celestial sphere* $\mathcal{C}$ is defined to be the quotient of the action of $\mathbb{R}^\times$ in C, so $\mathcal{C} = \{[x] = x\,\mathbb{R}^\times \mid x \in C\}$. We can think of $\mathcal{C}$ as a topological space homeomorphic to S^2 sitting in $\mathbb{R}^4$.

We will refer verbally as celestial sphere to both $\mathcal{C}$ and its image under the well-defined map that $p : \mathbb{R}^4 \xrightarrow{\cong} V$ induces on $\mathcal{C}$. We introduce a bit of notation to discriminate these two isomorphic topological spaces, $\mathcal{C}' = \{p(x) \mid [x] \in \mathcal{C}\}/(\mathbb{R}^\times) = \{A \in V \mid \det A = 0,\ A \neq 0\}/(\mathbb{R}^\times)$. Hence we will usually talk about elements in $\mathcal{C}'$ as null rays, i.e. sets of the form $\mathbb{R}^\times \cdot A$, where $A \in V$ is non-zero and $\det A = 0$.

A common way to visualise the celestial sphere is by thinking of it as the Riemann sphere. We go into a specific characterisation of the Riemann sphere.

Definition 4.3. Call $\mathcal{P}$ the set of 1-dimensional complex subspaces of $\mathbb{C}^2$. We can write it as

$$\mathcal{P} = \left\{\mathbb{C} \cdot (z, w) \mid \text{where } (z, w) \in \mathbb{C}^2, (z, w) \neq (0, 0)\right\}.$$

We claim that $\mathcal{P}$ is in bijective correspondence with the Riemann sphere. In a more detailed fashion, we build such bijection $\mathcal{S} : \mathcal{P} \longrightarrow \mathbb{C} \cup \{\infty\}$. Let $L \in \mathcal{P}$, we can assume $L = \mathbb{C} \cdot (z, w)$ with $(z, w) \neq (0, 0)$, then

$$\mathcal{S}(L) = \begin{cases} z/w & \text{if} \quad w \neq 0 \\ \infty & \text{if} \quad w = 0. \end{cases}$$

So from now on, when we refer to $\mathcal{P}$, we shall think of it as the Riemann sphere.

Let us observe that $\mathrm{PSL}_2(\mathbb{C})$ acts on $\mathcal{P}$. It is pretty clear that we get an action $\mathfrak{H}$: $\mathrm{GL}_2(\mathbb{C}) \longrightarrow \mathrm{Aut}(\mathcal{P})$ defined by $\mathfrak{H}(A)(L) = A(L)$, for all $A \in \mathrm{GL}_2(\mathbb{C})$ and $L \in \mathcal{P}$. This is because invertible matrices send 1-dimensional subspaces to 1-dimensional subspaces. Also, the identity matrix is sent by $\mathfrak{H}$ to the identity permutation on $\mathcal{P}$. Finally, because the product of matrices is associative, we can say that $\mathfrak{H}(AB) = \mathfrak{H}(A) \circ \mathfrak{H}(B)$. It can be easily seen that the kernel of the group homomorphism $\mathfrak{H}$ is the set of non-zero multiples of the identity, and we called that Z. Hence by the first isomorphism theorem, we get a unique group isomorphism $\mathfrak{h} : \mathrm{GL}_2(\mathbb{C})/Z \longrightarrow \mathrm{Aut}(\mathcal{P})$. But recall $\mathrm{GL}_2(\mathbb{C})/Z = \mathrm{PSL}_2(\mathbb{C})$, thus not only can we say that $\mathrm{PSL}_2(\mathbb{C})$ acts on $\mathcal{P}$, but we can ensure the action provided is *transitive*. This property actually falls directly from the fact that we have provided a surjective group homomorphism from $\mathrm{PSL}_2(\mathbb{C})$ to $\mathrm{Aut}(\mathcal{P})$. In a more clear way, if we take L and $\overline{L}$ in $\mathcal{P}$, there is always an automorphism $\ell \in \mathrm{Aut}(\mathcal{P})$ bringing L to $\overline{L}$, and because $\mathfrak{h}$ is surjective, we ensure the existence of a $g \in \mathrm{PSL}_2(\mathbb{C})$ such that $\mathfrak{h}(g) = \ell$. So overall, given any two elements $L, \overline{L} \in \mathcal{P}$ there exists a group element $g \in \mathrm{PSL}_2(\mathbb{C})$ such that $\mathfrak{h}(g)$ satisfies $\overline{L} = \mathfrak{h}(g)(L)$.

Moreover, $\Gamma_+^\uparrow$ acts on $\mathcal{C}$. In order to proceed, we first need to observe how an element of $\Gamma_+^\uparrow$ gives a permutation of $\mathcal{C}$. We need to show that indeed an element of $\Gamma_+^\uparrow$ sends rays to rays. First, check that any $A \in \Gamma_+^\uparrow$ leaves the null cone C invariant. This means, for all $x \in C$ then $Ax \in C$, since $(Ax, Ax) = (x, x) = 0$. Because after all, A is a linear map, it sends subspaces to subspaces. Thus the result, A sends rays to rays. Therefore, we get this map

$$\mathfrak{g} : \Gamma_+^\uparrow \longrightarrow \mathrm{Aut}(\mathcal{C}),$$

defined by $\mathfrak{g}(A)([x]) = [Ax]$, for all $A \in \Gamma_+^\uparrow$ and $[x] \in \mathcal{C}$. We shall prove that $\mathfrak{g}$ is a group action. Obviously, the identity matrix is mapped to the identity permutation in $\mathcal{C}$. The product of two matrices in $\Gamma_+^\uparrow$ is mapped by $\mathfrak{g}$ to the composition of the two permutations assigned to each factor, this falls from associativity of product of matrices.

Now the situation is the following. On the one hand, we have two isomorphic Lie groups $\mathrm{PSL}_2(\mathbb{C})$ and $\Gamma_+^\uparrow$. Though they are essentially the *same* object, the former group *moves* things that live in $\mathbb{C}^2$ (which is another way of saying that $\mathrm{PSL}_2(\mathbb{C})$ is an order 2 matrix group over $\mathbb{C}$) while the restricted Lorentz group *moves* things living in $\mathbb{R}^4$.

On the other hand, we have actions $\mathfrak{h} : \mathrm{PSL}_2(\mathbb{C}) \longrightarrow \mathrm{Aut}\,\mathcal{P}$ and $\mathfrak{g} : \Gamma^{\uparrow}_{+} \longrightarrow \mathrm{Aut}\,\mathcal{C}$. We believe we can define an isomorphism from $\mathcal{P}$ to $\mathcal{C}$ in a way that preserves the group actions defined above. This is, we want to define a map

$$\Theta : \mathcal{P} \longrightarrow \mathcal{C}$$

such that

$$\Theta(\mathfrak{h}(g)(L)) = \mathfrak{g}(\Psi(g))(\Theta(L)),$$

where $g \in \mathrm{PSL}_2(\mathbb{C})$ and $L \in \mathcal{P}$. Assuming that such Θ can be found, we express the compatibility with the actions in a more intuitive way. Considering the isomorphism $\Theta^* : \mathrm{Aut}\,\mathcal{P} \longrightarrow \mathrm{Aut}\,\mathcal{C}$ induced by the bijection $\Theta : \mathcal{P} \longrightarrow \mathcal{C}$ defined in the natural way

$$\begin{array}{rccc} \Theta^*(\alpha) : & \mathcal{C} & \longrightarrow & \mathcal{C} \\ & x & \longmapsto & (\Theta \circ \alpha \circ \Theta^{-1})(x), \end{array}$$

for any $\alpha \in \mathrm{Aut}\,\mathcal{P}$. Then asking that Θ preserves the actions of $\mathrm{PSL}_2(\mathbb{C})$ on $\mathcal{P}$ and $\Gamma^{\uparrow}_{+}$ on $\mathcal{C}$ can be formulated as making the diagram

$$\begin{tikzcd} \mathrm{PSL}_2(\mathbb{C}) \arrow[r, "\Psi"] \arrow[d, "\mathfrak{h}"'] & \Gamma^{\uparrow}_{+} \arrow[d, "\mathfrak{g}"] \\ \mathrm{Aut}\,\mathcal{P} \arrow[r, "\Theta^*"] & \mathrm{Aut}\,\mathcal{C} \end{tikzcd}$$

commute.

Before finding Θ, we take an intermediate step by defining $\Theta' : \mathcal{P} \longrightarrow \mathcal{C}'$ as follows

$$\Theta'\left(\mathbb{C} \cdot (z, w)\right) = \mathbb{R}^{\times} \cdot \begin{pmatrix} z\overline{z} & z\overline{w} \\ \overline{z}w & w\overline{w} \end{pmatrix},$$

for all $(z, w) \in \mathbb{C}^2 \setminus \{(0,0)\}$. Notice that as we have just defined it, $\Theta'\left(\mathbb{C} \cdot (z, w)\right)$ is the set of all non-zero real multiples of an Hermitian matrix with null determinant. Therefore it is clear that Θ' sends 1-dimensional subspaces of $\mathbb{C}^2$ to elements in $\mathcal{C}'$. We are going to show later that Θ' is the map we are looking for, since it is compatible with the action of $\mathrm{PSL}_2(\mathbb{C})$ on $\mathcal{P}$ and the action of $\Gamma'^{\uparrow}_{+}$ on $\mathcal{C}'$.

The following proposition tells us that Θ' is precisely the map that we are searching.

Proposition 4.4. The map Θ' as defined above is an isomorphism between $\mathcal{P}$ and $\mathcal{C}'$. Furthermore, it preserves the action of $\mathrm{PSL}_2(\mathbb{C})$ on $\mathcal{P}$ and the action of $\Gamma'^{\uparrow}_{+}$ on $\mathcal{C}'$.

Proof. We attempt to define the inverse of the map Θ', and we shall call it $\Theta'_* : \mathcal{C}' \longrightarrow \mathcal{P}$. Let R be a null ray in $\mathcal{C}'$, i.e. $R = \mathbb{R}^\times \cdot A$ for some non-zero Hermitian matrix A with determinant zero. Notice that, if we view A as an element of $\mathrm{End}(\mathbb{C}^2)$, the previous conditions on A imply that $\dim_{\mathbb{C}} \operatorname{Im} A = 1$. So we are tempted to directly set $\Theta'_*(R) = \operatorname{Im} A$. And indeed this is going to be the final answer, but we still have to show it is a well-defined choice. Suppose we have another matrix in the null ray R, say $B \in R$. Then there exists $\lambda \in \mathbb{R}^\times$ such that $B = \lambda A$, and so $\operatorname{Im} B = \lambda \operatorname{Im} A = \operatorname{Im} A$. And by checking this we have seen that our definition of $\Theta'_*(R)$ does not depend upon the choice of representative of the equivalence class R.

We now have to check that $\Theta'_* \circ \Theta' = \mathrm{id}_{\mathcal{P}}$ and $\Theta' \circ \Theta'_* = \mathrm{id}_{\mathcal{C}'}$. First, let $(z, w) \in \mathbb{C}^2 \setminus \{(0,0)\}$ and check

$$(\Theta'_* \circ \Theta')(\mathbb{C}(z,w)) = \Theta'_* \left(\mathbb{R}^\times \begin{pmatrix} z\overline{z} & z\overline{w} \\ \overline{z}w & w\overline{w} \end{pmatrix} \right)$$

Suppose now that $z \neq 0$, in which case we have

$$(\Theta'_* \circ \Theta')(\mathbb{C}(z,w)) = \mathbb{C}(z\overline{z}, w\overline{z}) = \mathbb{C}(z,w).$$

If $w \neq 0$ we have

$$\mathbb{C}(z\overline{z}, w\overline{z}) = \mathbb{C}(z\overline{w}, w\overline{w}) = \mathbb{C}(z,w)).$$

Hence in any of the cases $\Theta'_* \circ \Theta' = \mathrm{id}_{\mathcal{P}}$ holds.

Let now $A = (a_{ij})_{1 \le i,j \le 2} \in V$ such that $A \neq 0$ and $\det A = 0$. Notice that this very last condition is equivalent to $a_{11}a_{22} = a_{21}\overline{a_{21}}$. We will also take into account that $a_{11} \in \mathbb{R}$, since A is Hermitian, in other words $a_{11} = \overline{a_{11}}$. We distinguish two cases, according to the value of a_{11}. If $a_{11} \neq 0$, then

$$(\Theta' \circ \Theta'_*)(\mathbb{R}^\times A) = \Theta'(\operatorname{Im} A) = \Theta'(\mathbb{C}(a_{11}, a_{21})) = \mathbb{R}^\times \begin{pmatrix} a_{11}\overline{a_{11}} & a_{11}\overline{a_{21}} \\ \overline{a_{11}}a_{21} & a_{21}\overline{a_{21}} \end{pmatrix} = \mathbb{R}^\times \begin{pmatrix} a_{11} & \overline{a_{21}} \\ a_{21} & a_{22} \end{pmatrix}.$$

Hence we get the desired result in this case. Let us check what happens whenever $a_{11} = 0$. In this case, by the determinant condition we deduce that $a_{12} = a_{21} = 0$. Hence we are forced to have $a_{22} \in \mathbb{R}^\times$, since we cannot have $A = 0$. The image of A in that case will be the 1-dimensional linear subspace $\mathbb{C}(0, a_{22})$. Hence,

$$(\Theta' \circ \Theta'_*)(\mathbb{R}^\times A) = \mathbb{R}^\times \begin{pmatrix} 0 & 0 \\ 0 & a_{22}^2 \end{pmatrix} = \mathbb{R}^\times A.$$

Therefore, we can conclude that $\Theta' \circ \Theta'_* = \mathrm{id}_{\mathcal{C}'}$.

Now we have to show that Θ' is compatible with the group actions of $\mathrm{PSL}_2\,\mathbb{C}$ on $\mathcal{P}$ and Γ' on $\mathcal{C}'$, namely $\mathfrak{h}$ and $\mathfrak{g}'$. In fact we show that Θ' preserves the action provided by $\mathrm{GL}_2(\mathbb{C})$

on $\mathcal{P}$ and Γ' on $\mathcal{C}'$. In other words, if we let $X \in \mathrm{GL}_2(\mathbb{C})$ and $\mathbb{C}(z, w) \in \mathcal{P}$ then we must show

$$\Theta'(\mathfrak{H}(X)(\mathbb{C}(z, w))) = \mathfrak{g}(\mathfrak{A}(X))(\Theta'(\mathbb{C}(z, w))).$$

In simpler words, this is equivalent to showing

$$\Theta' X \mathbb{C}(z, w) = X\Theta'(\mathbb{C}(z, w))X^\dagger.$$

If we let $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{C})$, then the last condition can be written as

$$\mathbb{R}^\times \begin{pmatrix} (az+bw)\overline{(az+bw)} & (az+bw)\overline{(cz+dw)} \\ \overline{(az+bw)}(cz+dw) & (cz+dw)\overline{(cz+dw)} \end{pmatrix} = \mathbb{R}^\times \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z\overline{z} & z\overline{w} \\ \overline{z}w & w\overline{w} \end{pmatrix} \begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix}.$$

These *rays* (i.e. elements in $\mathcal{C}'$) are indeed the same since their representatives are equal. This means, the equality of sets before holds because the identity of matrices

$$\begin{pmatrix} (az+bw)\overline{(az+bw)} & (az+bw)\overline{(cz+dw)} \\ \overline{(az+bw)}(cz+dw) & (cz+dw)\overline{(cz+dw)} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z\overline{z} & z\overline{w} \\ \overline{z}w & w\overline{w} \end{pmatrix} \begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix}$$

is satisfied.

□

Note 4.5. We should not be worried about the fact that the equivariant isomorphism Θ' we found isn't quite what we promised it would be. We wanted an equivariant isomorphism between $\mathcal{P}$ and $\mathcal{C}$, but instead we provided an equivariant isomorphism between $\mathcal{P}$ and $\mathcal{C}'$. But this is easily solved because we shall remember that $p : \mathbb{R}^4 \longrightarrow V$ induced an isomorphism between $\mathcal{C}$ and $\mathcal{C}'$ which preserved the actions of the $\Gamma_+^\uparrow$ on $\mathcal{C}$ and the action of $\Gamma'^\uparrow_+$ on $\mathcal{C}'$. So overall, we get that the following diagram

$$\begin{tikzcd}
\mathrm{PSL}_2(\mathbb{C}) \arrow[r, "\tilde{\mathfrak{A}}"] \arrow[rr, bend left, "\Psi"] \arrow[d, "\mathfrak{h}"'] & \Gamma'^\uparrow_+ \arrow[r, "q_*"] \arrow[d, "\mathfrak{g}'"] & \Gamma_+^\uparrow \arrow[d, "\mathfrak{g}"] \\
\mathrm{Aut}\,\mathcal{P} \arrow[r, "\Theta'_*"'] \arrow[rr, bend right, "\Theta_*"'] & \mathrm{Aut}\,\mathcal{C}' \arrow[r, "q_*"'] & \mathrm{Aut}\,\mathcal{C}
\end{tikzcd}$$

commutes, since the two little squares do. And so we have checked that the map $\Theta_* : \mathrm{Aut}\,P \longrightarrow \mathrm{Aut}\,\mathcal{C}$ induced by $\Theta = q \circ \Theta'$ is equivariant.

So at this point is where we can give a physical meaning to the conjugacy classes in $\Gamma_+^\uparrow$, translating the behaviours of the conjugacy classes in $\mathrm{PSL}_2(\mathbb{C})$ on the Riemann sphere into behaviours of conjugacy classes in $\Gamma_+^\uparrow$ on the celestial sphere.

4.3 Giving Meaning to Conjugacy Classes in $\Gamma_+^\uparrow$

First of all, note that eigenvectors of $U \in \mathrm{PSL}_2(\mathbb{C})$ give rise to fixed points for the action of U on $\mathcal{P}$. But remember we just built this isomorphism $\Theta : \mathcal{P} \longrightarrow \mathcal{C}$ which is equivariant, so this means that fixed points for the action of U in $\mathcal{P}$ correspond bijectively to fixed points for the action of $\Psi(U)$ on $\mathcal{C}$. As the reader can already suspect, conjugacy classes come to play provided that eigenvector analysis of elements in $\mathrm{PSL}_2(\mathbb{C})$ is essentially the same as understanding of conjugacy classes in $\mathrm{PSL}_2(\mathbb{C})$.
We are going to cover all the possible cases, meaning that we are going to study how does $\Psi(U)$ look like in terms of geometry of spacetime for a representative U of any conjugacy class in $\mathrm{PSL}_2(\mathbb{C})$. For that we want to break the problem down into three distinctive cases:

4.3.1 The Diagonalisable Over $\mathbb{C}$ Case

Let A be a matrix in $\mathrm{SL}_2(\mathbb{C})$ conjugate to D_λ for $\lambda \in \mathbb{C} \setminus \{1, -1, 0\}$. This means that A has two linearly independent eigenvectors. The one dimensional subspaces generated by such eigenvectors are invariant under the action of A. So this is telling us that the two linearly independent eigenvectors of A yield to two distinct fixed points for the action of A on $\mathcal{P}$. Then we know that there are only two fixed points for the action of $B = \Psi(\mathfrak{p}A)$ on $\mathcal{C}$. We can take u_1 and u_2 to be linearly independent vectors in $\mathbb{R}^4$ generating the two fixed null rays by $\Psi(\mathfrak{p}A)$. Because u_1 and u_2 generate 1-dimensional real subspaces lying in the null cone C, we can assume that u_1 and u_2 have strictly positive time component (i.e. $u_1 \cdot e_1$ and $u_2 \cdot e_1$ are strictly positive numbers).

Proposition 4.6. Let $\mathbf{x}$ and $\mathbf{y}$ be two non-zero linearly independent vectors in the null cone. Then $(\mathbf{x}, \mathbf{y}) \neq 0$.

Proof. Let $\mathbf{x} = (t, x)$ and $\mathbf{y} = (s, y)$ be two non-zero linearly independent vectors in the null cone. We can assume $t, s > 0$. We shall first see that x and y are forced to be linearly independent vectors in $\mathbb{R}^3$. Suppose they are not, so there must exist a $\lambda \in \mathbb{R}^\times$ for which $x = \lambda y$. Hence $\|x\|^2 = \lambda^2 \|y\|^2$. Recall both $\mathbf{x}$ and $\mathbf{y}$ are in the null cone, so we have $t^2 = \|x\|^2$ and $s^2 = \|y\|^2$. Observe

$$t^2 = \|x\|^2 = \lambda^2 \|y\|^2 = \lambda^2 s^2.$$

Recall we had $t, s > 0$ so the equation above implies $t = \lambda s$. Using this, we see

$$\mathbf{x} = (t, x) = (\lambda s, \lambda y) = \lambda(s, y) = \lambda \mathbf{y}.$$

And this is a contradiction, since our assumption was that $\mathbf{x}$ and $\mathbf{y}$ were linearly independent vectors in $\mathbb{R}^4$.

We shall remark that it is fine to assume t, s are strictly positive numbers. Suppose we are given $\mathbf{x} = (t, x)$ and $\mathbf{y} = (s, y)$ where, for example, $t < 0$ and $s > 0$. Then we see that $-\mathbf{x}$ is still a vector in the null cone whose first coordinate is now strictly positive. In which case we know $(-\mathbf{x}, \mathbf{y}) \neq 0$ and so we can say $(\mathbf{x}, \mathbf{y}) = -(-\mathbf{x}, \mathbf{y}) \neq 0$. We proceed similarly on the other cases.

Once we know that x and y are going to be linearly independent vectors in $\mathbb{R}^3$ we proceed,

$$(\mathbf{x}, \mathbf{y}) = ts - \langle x, y \rangle > ts - \|x\|\|y\| = 0.$$

We have used Cauchy-Schwarz inequality. Recall that equality holds if and only if vectors involved are linearly dependent, but in our case we are assuming x, y linearly independent so strict inequality will hold.

□

From the last proposition, we have that our choice of u_2 can be such that $(u_1, u_2) = 1$. Now B fixes the null rays generated by u_1 and u_2, and at the same time it preserves the relation $(u_1, u_2) = 1$. So there is a unique $a \in \mathbb{R}^+$ such that $Bu_1 = au_1$. We can assume a is positive because $B \in \Gamma_+^\uparrow$, so it cannot send the future null cone to the past null cone. Also, B must preserve the null ray generated by u_2, i.e. Bu_2 is going to be a multiple of u_2, but from the condition $(Bu_1, Bu_2) = 1$ we are forced to have $Bu_2 = a^{-1}u_2$.
Now B clearly preserves the subspace $\mathcal{U} = \langle u_1, u_2 \rangle$. Let us check that it also preserves the subspace $\mathcal{V} = \{v \in \mathbb{R}^4 \ : \ (v, u_1) = (v, u_2) = 0\}$, which can be said to be the orthogonal complement of $\mathcal{U}$ under the Minkowski product. This is why we can sometimes refer to $\mathcal{V}$ as $\mathcal{U}^0$. Let $v \in V$, then

$$(Bv, u_1) = (Bv, a^{-1}Bu_1) = a^{-1}(Bv, Bu_1) = a^{-1}(v, u_1) = 0.$$

Similarly it can be seen that $(Bv, u_2) = 0$. Note that $\mathcal{V}$ is a 2-dimensional real subspace of $\mathbb{R}^4$. In order to carry on with our procedure we would need to prove that the quadratic form q is negative definite on the subspace $\mathcal{V}$, in order to pick a basis for $\mathcal{V}$ formed by two space-like vectors. We summarise this result in the following proposition.

Proposition 4.7. The quadratic form q is negative definite on $\mathcal{V}$. This is, for any $\mathbf{x}$ in $\mathcal{V}$ we have $\mathsf{q}(\mathbf{x}) = (\mathbf{x}, \mathbf{x}) < 0$.

Proof. Let us write $\mathbf{x} \in \mathcal{V}$ as $\mathbf{x} = (t, x)$ like we usually do, and we set $u_1 = (u_1^0, \mathrm{u}_1)$ and $u_2 = (u_2^0, \mathrm{u}_2)$ we can write the two linear equations defining $\mathbf{x} \in \mathcal{V}$ as

$$\begin{cases} tu_1^0 - \langle x, \mathrm{u}_1 \rangle = 0 \\ tu_2^0 - \langle x, \mathrm{u}_2 \rangle = 0 \end{cases}$$

We can multiply both equations to get a new equation

$$t^2 u_1^0 u_2^0 = \langle x, \mathrm{u}_1\rangle\langle x, \mathrm{u}_2\rangle.$$

We can take the absolute value both sides, and apply the Cauchy-Schwarz on the right hand side

$$t^2|u_1^0||u_2^0| = |\langle x, \mathrm{u}_1\rangle||\langle x, \mathrm{u}_2\rangle| \leq \|x\|^2\|\mathrm{u}_1\|\|\mathrm{u}_2\|.$$

Now observe that since u_1 is a null vector, it satisfies $(u_1^0)^2 = \|\mathrm{u}_1\|^2$. Therefore $|u_1^0| = \|\mathrm{u}_1\|$. Same argument applies for u_2. From the inequation above we deduce that $t^2 \leq \|x\|^2$, hence the result follows. $\square$

So we can choose a basis of $\mathcal{V}$ given by two vectors $v_3, v_4 \in \mathcal{V}$ such that $(v_3, v_3) = (v_4, v_4) = -1$ and $(v_3, v_4) = 0$. Also, by defining $v_1 = \frac{u_1+u_2}{\sqrt{2}}$ and $v_2 = \frac{u_1-u_2}{\sqrt{2}}$ we get that v_1, v_2 form a basis of $\mathcal{U}$ satisfying $(v_1, v_1) = 1$ and $(v_2, v_2) = -1$.

Well now, if we think about it, we have got an alternative basis for $\mathbb{R}^4$, composed by the vectors which we have called v_i, for $1 \leq i \leq 4$. We convince ourselves that it is indeed a basis of $\mathbb{R}^4$. It is clear that $\mathcal{U} = \langle v_1, v_2\rangle$ and $\mathcal{V} = \langle v_3, v_4\rangle$, so it would be enough to show that $\mathbb{R}^4 = \mathcal{U} \oplus \mathcal{V}$. We already know that the dimensions of $\mathcal{U}$ and $\mathcal{V}$ add up to $4 = \dim \mathbb{R}^4$, so we only need to see $\mathcal{U} \cap \mathcal{V} = \{0\}$. Let $x \in \mathcal{U} \cap \mathcal{V}$, we can assume $x = \lambda_1 u_1 + \lambda_2 u_2$, since $x \in \mathcal{U}$. The condition $x \in \mathcal{V}$ forces the following situation to happen. First, $(x, u_1) = \lambda_2 = 0$ and $(x, u_2) = \lambda_1 = 0$. And so $x = 0$.

So now consider the linear map $C : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ which sends the standard basis to the basis made up with v_i's we have just worked out, and define $\mathbf{C}$ to be the matrix form of C in the standard basis. In other words, C maps $e_i \mapsto v_i$ for all $1 \leq i \leq 4$. Notice that $(e_i, e_i) = (v_i, v_i) = (Ce_i, Ce_i)$ for all $1 \leq i \leq 4$, therefore C preserves the quadratic form on a basis of $\mathbb{R}^4$, therefore it preserves the quadratic form everywhere by polarisation. In other words, $C \in \Gamma$. Further, $\mathbf{C}_{11} = C(e_1) \cdot e_1 = v_1 \cdot e_1 = \frac{u_1+u_1}{\sqrt{2}} \cdot e_1 > 0$, since we chose u_1, u_2 to have strictly positive time component. This is telling us that C preserves the direction of time. We aim to see that it also preserves the orientation of space, i.e. $\det C$ is positive. Assume that the change of basis C has negative determinant. Then replace v_4 in the basis $(v_i)_{1\leq i\leq 4}$ by $-v_4$. Now the sign of the determinant of C will change, so we will get $\det C > 0$. Note C will still preserve the quadratic form since $(-v_4, -v_4) = (v_4, v_4) = (e_4, e_4) = -1$. All together combined results in $\det C = 1$. And so we have checked everything we need to say $C \in \Gamma_+^\uparrow$.

If we work out a matrix $\mathbf{B}$ for the linear transformation B in the basis $(v_i)_{1\leq i\leq 4}$, then $\mathbf{C}^{-1}\mathbf{B}\mathbf{C}$ is an element of the restricted Lorentz group, and it is indeed the matrix of B

in the standard basis. Let us work out what the coefficients of the matrix **B** should be. Notice that from the relations

$$v_1 = \frac{u_1 + u_2}{\sqrt{2}},\ v_2 = \frac{u_1 - u_2}{\sqrt{2}},$$

we are able to get

$$u_1 = \frac{v_1 + v_2}{\sqrt{2}},\ u_2 = \frac{v_1 - v_2}{\sqrt{2}}.$$

Hence it easy to determine the value of $B(v_1)$ as a linear combination of v_1 and v_2, by just substituting

$$B(v_1) = B\left(\frac{u_1 + u_2}{\sqrt{2}}\right) = \frac{au_1 + a^{-1}u_2}{\sqrt{2}} = \frac{(a + a^{-1})v_1 + (a - a^{-1})v_2}{2}.$$

Similarly, we find $B(u_2)$

$$B(v_2) = B\left(\frac{u_1 - u_2}{\sqrt{2}}\right) = \frac{au_1 - a^{-1}u_2}{\sqrt{2}} = \frac{(a - a^{-1})v_1 + (a + a^{-1})v_2}{2}.$$

We now proceed to find $B(v_3)$ and $B(v_4)$. Because we know that B leaves $\mathcal{V}$ invariant, we can assume $B(v_3)$ and $B(v_4)$ are going to be of the form

$$\begin{aligned} B(v_3) &= \alpha v_3 + \beta v_4 \\ B(v_4) &= \gamma v_3 + \delta v_4. \end{aligned}$$

Taking into account the restrictions $(v_3, v_3) = (v_4, v_4) = -1$, $(v_3, v_4) = 0$ and the fact that B must preserve the quadratic form, we obtain the following three restrictions for α, β, γ and δ:

$$\begin{cases} \alpha^2 + \beta^2 &= 1 \\ \gamma^2 + \delta^2 &= 1 \\ \alpha\gamma + \beta\delta &= 0. \end{cases}$$

Moreover, recall $B = \Psi(\mathfrak{p}A)$ is an element of $\Gamma_+^\uparrow$, hence has determinant 1. This condition happens to require

$$\alpha\delta - \gamma\beta = 1.$$

There are actually no more restrictions for the action of B on $\mathcal{V}$, which means that any $\alpha, \beta, \gamma, \delta$ satisfying the last four constraints will be valid. The set of solutions can be parametrised by $\theta \in \mathbb{R}$ the following way

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

The notation suggests strongly that the action of B on $\mathcal{V}$ is a rotation of space of angle θ along v_2. Note that B is a specific element in $\Gamma_+^\uparrow$, which means that it determines a unique $\theta \in [0, 2\pi)$ by $\theta = \widehat{v_3 B(v_3)} = \widehat{v_4 B(v_4)}$.

So, along the way two real values a and θ appeared as parameters completely determined by $A \in \mathrm{SL}_2(\mathbb{C})$. Knowing these values allow us to write down the matrix of B in the basis $(v_i)_{1 \le i \le 4}$, which is

$$P(a,\theta) = \begin{pmatrix} \frac{a+a^{-1}}{2} & \frac{a-a^{-1}}{2} & 0 & 0 \\ \frac{a-a^{-1}}{2} & \frac{a+a^{-1}}{2} & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix}$$

We shall remark that by our construction, values of (a, θ) could be *any* in $\mathbb{R}^+ \times [0, 2\pi)$. More precisely, if we choose any $(a, \theta) \in \mathbb{R}^+ \times [0, 2\pi)$ then $P(a, \theta) \in \Gamma_+^\uparrow$. Hence, undoing the procedure developed in this section we could be able to find $A \in \mathrm{SL}_2(\mathbb{C})$ such that a is an eigenvalue of $B = \Psi(\mathfrak{p}A)$ and θ is the angle of the rotation $B = \Psi(\mathfrak{p}A)$ applies to a space-like basis. So we start suspecting that there might be some kind of relation between possible values of a and θ and conjugacy classes in $\Gamma_+^\uparrow$. Indeed, we will see in a bit that there is a relation

$$\{(a,\theta) : a \in (0,1], \theta \in [0,2\pi)\} \rightarrowtail \{\text{Conjugacy Classes in } \Gamma_+^\uparrow\}.$$

We find out when two matrices of the form $P(a, \theta)$ and $P(a', \theta')$ are conjugate in $\Gamma_+^\uparrow$. It is a necessary condition for that to happen that $P(a, \theta)$ and $P(a', \theta')$ are conjugate in $M_{4\times 4}(\mathbb{R})$. For such thing to be true we need the eigenvalues of both matrices be the same. Let us find which are the eigenvalues of a matrix of the form $P(a, \theta)$, by considering its characteristic polynomial. This is

$$p_{a,\theta}(x) = \left(x^2 - (a + a^{-1})x + 1\right)\left(x^2 - 2\cos\theta x + 1\right),$$

which has 2 real roots and 2 non-real complex roots. These are

$$(a, a^{-1}, e^{i\theta}, e^{-i\theta})$$

There are only 4 cases in which the eigenvalues of $P(a, \theta)$ are the same as the ones of $P(a', \theta')$. These happen when $a' = a^{\pm 1}$ and $\theta' = \pm\theta$. This is telling us interesting things about the conjugacy classes of $P(a, \theta)$ in $M_{4\times 4}(\mathbb{R})$. Observe that by restricting the domain of a to $(0, 1] \subseteq \mathbb{R}^+$ and the one of θ to $[0, 2\pi)$, the result above tells us that two matrices of the form $P(a, \theta)$ and $P(a', \theta')$ cannot be conjugate in $M_{4\times 4}(\mathbb{R})$, since their eigenvalues are not the same. Hence they won't be conjugate in $\Gamma_+^\uparrow$ either. We can formulate this in other words, by saying that the map

$$\begin{array}{ccc} (0,1] \times [0,2\pi) & \longrightarrow & \Gamma_+^\uparrow/\sim \\ (a,\theta) & \longmapsto & [P(a,\theta)] \end{array}$$

is injective. Meaning that every $(a,\theta) \in (0,1] \times [0,2\pi)$ will give rise to a different conjugacy class in $\Gamma_+^\uparrow$.

We shall now give a physical interpretation to the conjugacy classes obtained in this case. Observe that we found an *observer* for which B looks like a boost in the space-like direction v_1 together with a rotation of angle θ along the v_2 space-like direction. This is formulated equivalently in a more algebraic way saying that we found an element $C \in \Gamma_+^\uparrow$ for which the linear map CBC^{-1} had the matrix form of $P(a,\theta)$. Let us stress the fact that C is representing the change of coordinates between two observers. More specifically, it is sending coordinates of an event recorded by the first observer (i.e. coordinates in the standard basis $(e_i)_{1\leq i\leq 4}$) to coordinates of the same event viewed from the second observer's perspective (i.e. coordinates in the basis $(v_i)_{1\leq i\leq 4}$). And so, it is easy to work out how does the second observer *sees* B act on spacetime. We distinguish two cases, according to possible values of a exhibiting different behaviours.

- Suppose $a = 1$. Then

$$P(1,\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix},$$

  hence second observer sees B as a pure spatial **rotation** along the direction of v_2.

- For $a \in (0,1)$, we can break down the linear transformation B expressed in the basis $((v_i)_{1\leq i\leq 4})$ into two smaller bits. Firstly, it acts as a boost in the space-like direction v_2. Moreover, it rotates the subspace spanned by the the space-like directions v_2, v_3, v_4 along the v_2 direction by an angle of θ. Considering the case $\theta \neq 0$ we obtain what is known as a **screw**. If we take the rotation angle θ to be 0, then we get a transformation known as **boost**.

4.3.2 The Parabolic Case

Let A be conjugated to E_+ in $\mathrm{PSL}_2(\mathbb{C})$, then there is only one 1-dimensional complex subspace which is invariant for A. As a result, we get one single fixed point for the action of A on the Riemann sphere $\mathcal{P}$. Observe that this single fixed point in $\mathcal{P}$ can be *whichever*, since we can always find a conjugated element in $\mathrm{PSL}_2(\mathbb{C})$ to E_+ such that fixes *any* 1-dimensional subspace of $\mathbb{C}^2$. In terms of the geometry of spacetime, we see that $B = \Psi(\mathfrak{p}A)$ fixes a light-like ray, and moreover the elements in this conjugacy class are called **null rotations**. Similarly to what we did before, we work out the canonical form of a null rotation.

We claim that B has an invariant subspace of dimension 2 of the form $\langle k, x\rangle$, where k is a null vector and x is a space-like vector. From a more general standpoint, we know that any endomorphism of $\mathbb{R}^4$ must possess an invariant 2-space. This follows from the fact that the minimal polynomial equation satisfied by the endomorphism factorises, over the real field, into factors of degree at most 2.

We have to bear in mind that, when dealing with Lorentz transformations, there is always at least one invariant subspace generated by some light-like vector. This can be seen as consequence of the fact that any $U \in \mathrm{PSL}_2(\mathbb{C})$ always has an invariant one-dimensional complex subspace of $\mathbb{C}^2$.

So our Lorentz transformation $B = \Psi(\mathfrak{p}A)$ must leave a 2-dimensional subspace of $\mathbb{R}^4$ invariant, which must contain exactly a 1-dimensional subspace in the null cone. It is not hard to prove that this 2-dimensional invariant subspace $\mathtt{V}$ we are after must be of the form $\mathtt{V} = \langle k, x\rangle$, where k is a null vector and x is a space-like vector. It is a demand for $\mathtt{V}$ to contain a 1-dimensional subspace generated by a light-like vector, so we can pick k in the null cone such that $\langle k\rangle \subset \mathtt{V}$. Now let x be a linearly independent vector to k such that $\mathtt{V} = \langle k, x\rangle$. Clearly x cannot be light-like, or else B would have two different fixed null rays. So assume $(x, x) \neq 0$ and observe that we are looking for circumstances under which

$$\mathtt{q}(\lambda k + \mu x) \neq 0,$$

for all $\lambda, \mu \in \mathbb{R}$ excepting the case $\lambda = \mu = 0$. Or equivalently, we want to know when

$$\mathtt{q}(k + \lambda x) \neq 0,$$

for all $\lambda \in \mathbb{R}^\times$. Let us study the possible solution on λ of the following equation:

$$(k + \lambda x, k + \lambda x) = 2\lambda(k, x) + \lambda^2(x, x) = 0.$$

Check that if $(k, x) \neq 0$ then when can always find $\lambda \neq 0$ such that the equation above is satisfied. Instead, when we require $(k, x) = 0$, then the equation above has no solution for $\lambda \in \mathbb{R}^\times$. We now claim that requiring $(k, x) = 0$ and $(x, x) \neq 0$ implies that x must be space-like. If we let $k = (k_0, k_s)$ and similarly $x = (x_0, x_s)$ then $(k, x) = 0$ implies

$$|k_0 x_0| = |\langle k_s, x_s\rangle| \leq \|k_s\|\|x_s\|.$$

Because k is light-like, we have $|k_0| = \|k_s\|$. So we obtain $|x_0| \leq \|x_s\|$, or in other words, x space-like. Notice that we can perfectly choose x to satisfy $(x, x) = -1$.

We now choose a basis $\{k, l, x, y\}$ of $\mathbb{R}^4$, where k and x are the vectors described above; y is a space-like vector linearly independent to x satisfying $(y, y) = -1$ and $(x, y) = 0$, and l is a light-like vector not belonging to the subspace spanned by k. Consider the following facts:

1. The null space $\langle k \rangle$ is invariant for B.
2. The subspaces $\langle k, x \rangle$ and $\langle k, x \rangle^0 = \{v \in \mathbb{R}^4 : (k, v) = (x, v) = 0\} = \langle k, y \rangle$ are invariant for B.

Using the isometric property of $B \in \Gamma_+^\uparrow$ we find that

$$\begin{cases} B(k) = \mu k, \ B(x) = \pm x + \alpha k, \ B(y) = \pm y + \beta k \\ B(l) = \mu^{-1} \left(l \pm \alpha x \pm \beta y + \dfrac{1}{2}(\alpha^2 + \beta^2) k \right) \end{cases} \tag{4}$$

where μ, α, β are scalars. Recall that B is an element of $\Gamma_+^\uparrow$, so it maps the future null cone to itself, therefore $\mu > 0$. We are willing to show that $\mu = 1$, and we need two general results about the restricted Lorentz group to do so.

Proposition 4.8. Assume μ is an eigenvalue of $A \in \Gamma_+^\uparrow$. Then μ^{-1} is also an eigenvalue of A.

Proof. Let A be an element of $\Gamma_+^\uparrow$, for which μ is an eigenvalue. As a remark, we shall say that μ cannot be 0, since any element of the Lorentz group is invertible. Because of the isometric property of the Lorentz transformations, it follows that $A^{-1} = J^{-1} A^T J$. Now observe that $A^{-1} - \mu \, \mathrm{I} = J^{-1}(A^T - \mu \, \mathrm{I}) J$. And so

$$\det(A^{-1} - \mu \, \mathrm{I}) = (\det J)^{-1} \det(A^T - \mu \, \mathrm{I}) \det J = \det(A - \mu \, \mathrm{I}) = 0,$$

taking into account that μ is an eigenvalue of A. Finally, check that

$$A^{-1} - \mu \, \mathrm{I} = A^{-1}(\mathrm{I} - \mu A) = (-\mu A^{-1})(A - \mu^{-1} \, \mathrm{I}).$$

This means that $\det(A - \mu^{-1} \, \mathrm{I}) = 0$, which makes μ^{-1} an eigenvalue of A. $\square$

Assume the eigenvalue μ for which $B(k) = \mu k$ is different than 1, then according to the last proposition, μ^{-1} is an eigenvalue of B as well. If we call $\tilde{k}$ a corresponding eigenvector, we see that because $\mu \neq \mu^{-1}$, k and $\tilde{k}$ must be linearly independent.

Proposition 4.9. Let $A \in \Gamma_+^\uparrow$ such that there exists $\mu \neq \pm 1$ scalar and v vector such that $A(v) = \mu v$. Then the vector v is light-like.

Proof. From the isometric property of A, we have that $(v, v) = (A(v), A(v)) = \mu^2(v, v)$. Because we are assuming $|\mu| \neq 1$, it follows that (v, v) must be 0. □

Since the eigenvalue $|\mu| \neq 1$, we deduce that $\tilde{k}$ is a light-like vector. Notice that we have obtained another invariant 1-subspace of the null cone for B, namely $\langle \tilde{k} \rangle$. But it was a requisite that our Lorentz transformation B only fixed one ray in the null cone. It follows that $\mu = 1$.

It must happen that $\det B = 1$, so it must either take all the upper signs or the lower ones throughout (4). If it takes the lower signs, then we can change basis to $x' = x - \frac{1}{2}\alpha k$ and $y' = y - \frac{1}{2}\beta k$ and get $B(x') = -x'$ and $B(y') = -y'$. Therefore the space-like subspace spanned by x' and y' is invariant for B. The invariance of a space-like 2-subspace can be shown to imply that B is not a null rotation (see reference [2]).

Finally, the only case left to consider is

$$\begin{cases} B(k) = k, \ B(x) = x + \alpha k, \ B(y) = y + \beta k \\ B(l) = l + \alpha x + \beta y + \dfrac{1}{2}(\alpha^2 + \beta^2)k \end{cases}$$

By change of orthonormal basis

$$x' = \frac{\alpha x + \beta y}{\sqrt{\alpha^2 + \beta^2}}, \quad y' = \frac{\alpha x - \beta y}{\sqrt{\alpha^2 + \beta^2}},$$

and using the re-scalings of k and l

$$k' = \sqrt{\alpha^2 + \beta^2} k, \quad l' = \frac{l}{\sqrt{\alpha^2 + \beta^2}},$$

we get that the Lorentz transformation B has a simple matrix form in the basis $\{k', l', x', y'\}$, given by

$$Q = \begin{pmatrix} 1 & \frac{1}{2} & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

As a result, we can say that all null rotations will be conjugated to Q.

4.3.3 The Trivial Case

$A = \mathrm{I}$. Notice that in $\mathrm{SL}_2(\mathbb{C})$, the conjugacy class of the identity matrix is $\{\mathrm{I}, -\mathrm{I}\}$. Hence in $\mathrm{PSL}_2(\mathbb{C})$ the class of the identity element is a singleton. Since Ψ is a homomorphism, we get directly that $\Psi(\mathfrak{p}A) = \mathrm{I} \in \mathrm{GL}_4(\mathbb{R})$, hence this case is trivial.

References

[1] Bernard F.Schutz, *A First Course in General Relativity*, Cambridge University Press 2nd edition, 2009.

[2] Ronald Shaw, *The Conjugacy Classes of the Homogeneous Lorentz Group*, University of Hull, April 1969.

[3] Wikipedia, *Lorentz Group*, 2015, http://en.wikipedia.org/wiki/Lorentz_group.

[4] Wikipedia, *Möbius Transformation*, 2015, http://en.wikipedia.org/wiki/Mbius_transformation.