Appendix - The Lorentz group, Mobius transformations and the sky of a rapidly moving observer

1. The Lorentz group

Recall that the group of all isometries of a Minkowski spacetime is the so-called Poincaré group. The Lorentz group is the subgroup of the Poincaré group formed by all linear isometries, or, equivalently, all isometries which fix the origin. Consequently the Lorentz group determines the relation between the observations of two inertial observers at a given event in a general curved spacetime.

If $\{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is an orthonormal basis for the Minkowski 4-spacetime and

$$v = v^0\mathbf{e}_0 + v^1\mathbf{e}_1 + v^2\mathbf{e}_2 + v^3\mathbf{e}_3$$

is a vector, then

$$\begin{aligned}
\langle v, v\rangle &= -\left(v^0\right)^2 + \left(v^1\right)^2 + \left(v^2\right)^2 + \left(v^3\right)^2 \\
&= \begin{pmatrix} v^0 & v^1 & v^2 & v^3 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix} \\
&= x^t \eta x
\end{aligned}$$

where x is the column vector of v's components and $\eta = diag\,(-1, 1, 1, 1)$. If L is a Lorentz transformation and Λ its matrix representation with respect to the chosen basis, then one must have

$$\langle Lv, Lv\rangle = \langle v, v\rangle \Leftrightarrow (\Lambda x)^t\,\eta\,(\Lambda x) = x^t\eta x \Leftrightarrow x^t\left(\Lambda^t\eta\Lambda\right)x = x^t\eta x.$$

Since this must hold for all $x \in \mathbb{R}^4$ and both $\Lambda^t\eta\Lambda$ and η are symmetric matrices, we conclude that

Proposition 1.1. *The Lorentz group is (isomorphic to)*

$$O\,(3,1) = \left\{\Lambda \in GL(4) : \Lambda^t\eta\Lambda = \eta\right\}.$$

Example 1.2. *If $R \in O\,(3)$ then*

$$\widetilde{R} = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}$$

satisfies

$$\begin{aligned}
\widetilde{R}^t\eta\widetilde{R} &= \begin{pmatrix} 1 & 0 \\ 0 & R^t \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \\
&= \begin{pmatrix} -1 & 0 \\ 0 & R^tR \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & I \end{pmatrix} = \eta
\end{aligned}$$

and thus $\widetilde{R} \in O\,(3,1)$. It is easy to see that in fact

$$\widetilde{O}(3) = \left\{\widetilde{R} \in O\,(3,1) : R \in O\,(3)\right\}$$

is a subgroup of $O\left(3,1\right)$ *isomorphic to* $O\left(3\right)$*. For instance, since*

$$\begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 0 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \in O(3)$$

for any $\theta \in \mathbb{R}$*, we know that*

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & -\sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & \sin\theta & 0 & \cos\theta \end{pmatrix} \in O(3,1);$$

this Lorentz transformation is said to be a rotation about $\mathbf{e}_3$ by an angle θ*.*

Example 1.3. *Not all Lorentz transformations are rotations. For instance, defining*

$$B = \begin{pmatrix} \cosh u & 0 & 0 & \sinh u \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh u & 0 & 0 & \cosh u \end{pmatrix}$$

one sees that

$$\begin{aligned} B^t\eta B &= \begin{pmatrix} \cosh u & 0 & 0 & \sinh u \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh u & 0 & 0 & \cosh u \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cosh u & 0 & 0 & \sinh u \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh u & 0 & 0 & \cosh u \end{pmatrix} \\ &= \begin{pmatrix} -\cosh u & 0 & 0 & \sinh u \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh u & 0 & 0 & \cosh u \end{pmatrix} \begin{pmatrix} \cosh u & 0 & 0 & \sinh u \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh u & 0 & 0 & \cosh u \end{pmatrix} \\ &= \begin{pmatrix} \sinh^2 u - \cosh^2 u & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \cosh^2 u - \sinh^2 u \end{pmatrix} = \eta \end{aligned}$$

and therefore $B \in O(3,1)$*. This Lorentz transformation is said to be a* boost in the $\mathbf{e}_3$ direction by a hyperbolic angle u*.*

Let us now recall briefly what is meant by *active* and *passive* transformations. Setting

$$E = \begin{pmatrix} \mathbf{e}_0 & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{pmatrix}$$

it is clear that

$$v = v^0\mathbf{e}_0 + v^1\mathbf{e}_1 + v^2\mathbf{e}_2 + v^3\mathbf{e}_3 = Ex$$

and consequently

$$Lv = L\left(Ex\right) = E\left(\Lambda x\right).$$

In particular,

$$LE = \begin{pmatrix} L\mathbf{e}_0 & L\mathbf{e}_1 & L\mathbf{e}_2 & L\mathbf{e}_3 \end{pmatrix} = L\left(EI\right) = E\left(\Lambda I\right) = E\Lambda.$$

Thus in the *new* orthonormal frame $E' = LE$ the *same* vector v has new coordinates x' such that

$$v = Ex = E'x' \Leftrightarrow Ex = E\Lambda x'$$

i.e.,

$$x' = \Lambda^{-1} x.$$

Thus if Λ represents an *active* Lorentz transformation L, Λ^{-1} represents the corresponding *passive* transformation, yielding the coordinates of any vector on the orthonormal frame obtained by applying the active transformation to the vectors of the initial orthonormal frame.

Example 1.4. *Let B represent a boost in the $\mathbf{e}_3$ direction by a hyperbolic angle u; then an event with coordinates*

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

in the initial frame E will have coordinates

$$\begin{aligned}
\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} &= \begin{pmatrix} \cosh u & 0 & 0 & \sinh u \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh u & 0 & 0 & \cosh u \end{pmatrix}^{-1} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \\
&= \begin{pmatrix} \cosh u & 0 & 0 & -\sinh u \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh u & 0 & 0 & \cosh u \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \\
&= \begin{pmatrix} t\cosh u - z\sinh u \\ x \\ y \\ z\cosh u - t\sinh u \end{pmatrix}
\end{aligned}$$

in the transformed frame E'. In particular,

$$z' = 0 \Leftrightarrow z\cosh u - t\sinh u = 0 \Leftrightarrow z = t\tanh u$$

and we see that the transformed frame corresponds to an inertial observer moving with speed $v = \tanh u$ with respect to the inertial observer represented by the initial frame.

If $\Lambda \in O(3,1)$ then

$$\Lambda^t \eta \Lambda = \eta \Rightarrow \det\left(\Lambda^t \eta \Lambda\right) = \det \eta = -1 \Leftrightarrow -(\det \Lambda)^2 = -1 \Leftrightarrow \det \Lambda = \pm 1.$$

Now consider the four matrices

$$I, \Sigma = \begin{pmatrix} I & 0 \\ 0 & -1 \end{pmatrix}, \Theta = \begin{pmatrix} -1 & 0 \\ 0 & I \end{pmatrix}, \Omega = \Sigma\Theta,$$

all of which are trivially in $O(3,1)$. We see that

$$\det I = -\det \Sigma = -\det \Theta = \det \Omega = 1$$

and consequently there are matrices in $O(3,1)$ with either value of the determinant. Since the determinant is a continuous function, it follows that $O(3,1)$ has at least two connected components.

Also, if I, S, T, U are the Lorentz transformations represented by I, Σ, Θ, Ω then

$$I\mathbf{e}_0 = S\mathbf{e}_0 = -T\mathbf{e}_0 = -U\mathbf{e}_0 = \mathbf{e}_0.$$

Now if L is a Lorentz transformation then define

$$f(L) = \langle \mathbf{e}_0, L\mathbf{e}_0 \rangle .$$

Since

$$\langle L\mathbf{e}_0, L\mathbf{e}_0 \rangle = \langle \mathbf{e}_0, \mathbf{e}_0 \rangle = -1$$

one gets from the backwards Schwartz inequality,

$$|f(L)| = |\langle \mathbf{e}_0, L\mathbf{e}_0 \rangle| \geq |\mathbf{e}_0| \, |L\mathbf{e}_0| = 1.$$

Since

$$f(I) = f(S) = -f(T) = -f(U) = -1$$

we see that I and S cannot belong to the same connected component of the Lorentz group as T and U. Thus $O(3,1)$ has at least four distinct connected components. We summarize this in the following

Proposition 1.5. *$O(3,1)$ is the disjoint union of the four open sets*

$$\begin{aligned} O_+^{\uparrow}(3,1) &= \{\Lambda \in O(3,1) : \det \Lambda = -f(\Lambda) = 1\}; \\ O_+^{\downarrow}(3,1) &= \{\Lambda \in O(3,1) : \det \Lambda = f(\Lambda) = 1\}; \\ O_-^{\uparrow}(3,1) &= \{\Lambda \in O(3,1) : \det \Lambda = f(\Lambda) = -1\}; \\ O_-^{\downarrow}(3,1) &= \{\Lambda \in O(3,1) : -\det \Lambda = f(\Lambda) = 1\}. \end{aligned}$$

Informally, $O_+^{\uparrow}(3,1)$ is the set of Lorentz transformations which preserve both orientation and time orientation; $O_+^{\downarrow}(3,1)$ is the set of Lorentz transformations which preserve orientation but reverse time orientation (and consequently must reverse space orientation as well); $O_-^{\uparrow}(3,1)$ is the set of Lorentz transformations which reverse orientation but preserve time orientation (hence reversing space orientation); and $O_-^{\downarrow}(3,1)$ is the set of Lorentz transformations which reverse both orientation and reverse time orientation (hence preserving space orientation).

Exercise 1.6. *Show that (i) $O_+^{\downarrow}(3,1) = TO_+^{\uparrow}(3,1)$; (ii) $O_-^{\uparrow}(3,1) = SO_+^{\uparrow}(3,1)$; (iii) $O_-^{\downarrow}(3,1) = UO_+^{\uparrow}(3,1)$.*

Of these disjoint open subsets of the Lorentz group only $O_+^{\uparrow}(3,1)$ contains the identity, and can therefore be a subgroup.

Exercise 1.7. *Show that $O_+^{\uparrow}(3,1)$ is a subgroup of $O(3,1)$ ($O_+^{\uparrow}(3,1)$ is called the group of proper Lorentz transformations).*

It is possible to prove that $O_+^{\uparrow}(3,1)$ is connected (but not simply connected; as we will see, $\pi_1\left(O_+^{\uparrow}(3,1)\right) = \mathbb{Z}_2$).

2. Stereographic projection

Recall that

$$S^2 = \left\{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\right\}.$$

The points $N = (0,0,1)$ and $S = (0,0,-1)$ are said to be the *North* and *South poles* of S^2, and will play special roles in what follows.

We define

$$\alpha = \left\{(x,y,z) \in \mathbb{R}^3 : z = 0\right\}$$

and identify α with $\mathbb{C}$ by identifying $(x,y,0)$ with $\zeta = x+iy$. The *stereographic projection* $\zeta : S^2 \backslash \{N\} \to \mathbb{C}$ is the map that to each $(x,y,z) \in S^2 \backslash \{N\}$ associates the intersection ζ of the line through $(0,0,1)$ and (x,y,z) with α. Thus

$$\zeta\left(x,y,z\right) = \lambda \frac{x+iy}{\left(x^2+y^2\right)^{\frac{1}{2}}}$$

where

$$\frac{\lambda}{1} = \frac{\left(x^2+y^2\right)^{\frac{1}{2}}}{1-z}$$

i.e.,

$$\zeta\left(x,y,z\right) = \frac{x+iy}{1-z}.$$

Introducing spherical coordinates (r,θ,φ) in an appropriate open set of $\mathbb{R}^3$ through the inverse coordinate transformation

$$\begin{aligned} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{aligned}$$

we see that S^2 is the level set $r=1$ and hence (θ,φ) are local coordinates in the corresponding open set in S^2. Thus we can write

$$\zeta\left(\theta,\varphi\right) = \frac{\sin\theta\cos\varphi + i\sin\theta\sin\varphi}{1-\cos\theta} = \frac{\sin\theta}{1-\cos\theta}e^{i\varphi}.$$

One can think of this as a coordinate transformation in S^2. The derivative of this transformation is seen to be given by

$$\begin{aligned} d\zeta &= \frac{\cos\theta\left(1-\cos\theta\right)-\sin^2\theta}{\left(1-\cos\theta\right)^2}e^{i\varphi}d\theta + i\frac{\sin\theta}{1-\cos\theta}e^{i\varphi}d\varphi \\ &= -\frac{1}{1-\cos\theta}e^{i\varphi}d\theta + i\frac{\sin\theta}{1-\cos\theta}e^{i\varphi}d\varphi \end{aligned}$$

and hence

$$\begin{aligned} d\zeta d\overline{\zeta} &= \frac{1}{\left(1-\cos\theta\right)^2}d\theta^2 + \frac{\sin^2\theta}{\left(1-\cos\theta\right)^2}d\varphi^2 \\ &= \frac{1}{\left(1-\cos\theta\right)^2}ds^2 \end{aligned}$$

where ds^2 is the usual line element of S^2. Since

$$1+\zeta\overline{\zeta} = 1+\frac{\sin^2\theta}{\left(1-\cos\theta\right)^2} = \frac{1-2\cos\theta+\cos^2\theta+\sin^2\theta}{\left(1-\cos\theta\right)^2} = \frac{2}{1-\cos\theta}$$

we see that

$$\begin{aligned} ds^2 &= \frac{4}{\left(1+\zeta\overline{\zeta}\right)^2}d\zeta d\overline{\zeta} \\ &= \frac{4}{\left(1+x^2+y^2\right)}\left(dx^2+dy^2\right). \end{aligned}$$

Thus if one sees the stereographic projection as a map $\zeta : S^2\setminus\{N\} \to \mathbb{C} \approx \alpha \approx \mathbb{R}^2$ we see that it is a *conformal map*, i.e., it satisfies

$$\langle u_p, v_p\rangle_{S^2} = \Omega^2(p)\,\langle \zeta_* u_p, \zeta_* v_p\rangle_{\mathbb{R}^2}$$

for all $u_p, v_p \in T_pS^2$ and all $p \in S^2\setminus\{N\}$. Another way of putting this is to say that the stereographic projection maps circles on T_pS^2 to circles on $T_{\zeta(p)}\mathbb{R}^2$ (or that it maps infinitesimal circles to infinitesimal circles).

A *circle* in S^2 is just a geodesic sphere, i.e., the image through the exponential map of a circle on some tangent space. It is easy to see that any circle is the intersection of $S^2 \subset \mathbb{R}^3$ with some plane $\beta \subset \mathbb{R}^3$.

Proposition 2.1. *If $\gamma \subset S^2$ is a circle then $\zeta(\gamma) \subset \mathbb{C}$ is either a straight line or a circle depending on whether or not $N \in \gamma$.*

Exercise 2.2. *Prove proposition 2.1.*

3. Complex structure of S^2

Obviously one can define another stereographic projection $\widetilde{\zeta} : S^2\setminus\{S\} \to \mathbb{C}$ by associating to each $(x,y,z) \in S^2\setminus\{S\}$ the intersection $\widetilde{\zeta}$ of the line through $(0,0,1)$ and (x,y,z) with α. Crucially, however, one now identifies α with $\mathbb{C}$ by identifying $(x,y,0)$ with $\widetilde{\zeta} = x - iy$. Thus

$$\widetilde{\zeta}(x,y,z) = \widetilde{\lambda}\frac{x-iy}{(x^2+y^2)^{\frac{1}{2}}}$$

where

$$\frac{\widetilde{\lambda}}{1} = \frac{\left(x^2+y^2\right)^{\frac{1}{2}}}{1+z}$$

i.e.,

$$\widetilde{\zeta}(x,y,z) = \frac{x-iy}{1+z}.$$

Notice that on $S^2\setminus\{N,S\}$ one has

$$\zeta\widetilde{\zeta} = \frac{x^2+y^2}{1-z^2} = 1$$

and consequently $\widetilde{\zeta}\circ\zeta^{-1} : \mathbb{C}\setminus\{0\} \to \mathbb{C}\setminus\{0\}$ is the map $\zeta \mapsto \frac{1}{\zeta}$. In addition to being smooth, this map is complex analytic.

Definition 3.1. *The set*

$$\mathcal{A} = \left\{\left\{\mathbb{C}, \zeta^{-1}\right\}, \left\{\mathbb{C}, \widetilde{\zeta}^{-1}\right\}\right\}$$

is said to be an analytic atlas *for S^2, which is then said to possess the structure of a (1-dimensional)* complex manifold*.*

Clearly having a complex structure is a stronger requirement than having a differentiable structure. When a manifold possesses this kind of structure the natural functions to consider are no longer smooth functions:

Definition 3.2. *A map $f : S^2 \to S^2$ is said to be* complex analytic iff *both complex functions of complex variable $\zeta\circ f\circ\zeta^{-1}$ and $\widetilde{\zeta}\circ f\circ\widetilde{\zeta}^{-1}$ are complex analytic.*

Let $f : S^2 \to S^2$ be a complex analytic automorphism. If $f(N) = N$ then $g = \zeta \circ f \circ \zeta^{-1}$ must be holomorphic in $\mathbb{C}$ and

$$\lim_{|\zeta| \to +\infty} |g(\zeta)| = +\infty.$$

If $f(N) = p' \neq N$, then $f(p'') = N$ for some $p'' \neq N$. If $\zeta' = \zeta(p')$ and $\zeta'' = \zeta(p'')$ then g will have a singularity at ζ'' as we must have

$$\lim_{\zeta \to \zeta'} |g(\zeta)| = +\infty$$

and will necessarily satisfy

$$\lim_{|\zeta| \to +\infty} |g(\zeta)| = \zeta'.$$

We conclude that any complex analytic automorphism of S^2 can be represented by an analytic function on $\mathbb{C}$ with at most one singularity and with a well defined limit as $|\zeta| \to +\infty$. This is often summarized by extending g to $\mathbb{C} \cup \{\infty\}$ and writing $g(\infty) = \infty$ in the case $f(N) = N$ and $g(\infty) = \zeta'$, $g(\zeta'') = \infty$ in the case $f(N) \neq N$. Notice that one can identify S^2 with $\mathbb{C} \cup \{\infty\}$ and hence f with g. This could have been done by using the South pole chart, and one should be careful to stress which chart is being used.

Example 3.3. *Let $f : S^2 \to S^2$ be represented by $g : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ given by $g(\zeta) = \zeta + b$, with $b \neq 0$ (thus $g(\infty) = \infty$). Clearly f is bijective and $\zeta \circ f \circ \zeta^{-1} = g\,|_{\mathbb{C}}$ is holomorphic. As for $\widetilde{\zeta} \circ f \circ \widetilde{\zeta}^{-1}$, it is given on the overlap of the North and South pole charts by*

$$h\left(\widetilde{\zeta}\right) = \frac{1}{g\left(\frac{1}{\widetilde{\zeta}}\right)} = \frac{1}{\frac{1}{\widetilde{\zeta}} + b} = \frac{\widetilde{\zeta}}{b\widetilde{\zeta} + 1}$$

and since $f(N) = N$ and $\widetilde{\zeta}(N) = 0$ the above expression is valid also for $\widetilde{\zeta} = 0$. Thus

$$\widetilde{\zeta} \circ f \circ \widetilde{\zeta}^{-1} : \mathbb{C} \backslash \left\{-\frac{1}{b}\right\} \to \mathbb{C} \backslash \left\{\frac{1}{b}\right\}$$

is seen to be holomorphic on its domain, and hence f is a complex analytic automorphism. Notice by the way that h can be extended to $\mathbb{C} \cup \{\infty\}$ by setting $h\left(-\frac{1}{b}\right) = \infty$, $h(\infty) = \frac{1}{b}$. These are the South pole chart versions of $g(-b) = 0$, $g(0) = b$.

Exercise 3.4. *Show that the functions represented by $a\zeta$ ($a \neq 0$) and $\frac{1}{\zeta}$ are complex analytic diffeomorphisms.*

Clearly any composition of complex analytic automorphisms is a complex analytic automorphism. Let g represent a complex analytic automorphism. If $g(\infty) \neq \infty$ then $g(a) = \infty$ for some $a \in \mathbb{C}$. Consequently

$$g_1(\zeta) = g\left(a + \frac{1}{\zeta}\right)$$

represents a complex analytic automorphism satisfying $g_1(\infty) = \infty$. If $g_1(0) = b \neq 0$, then

$$g_2(\zeta) = g_1(\zeta) - b$$

satisfies $g_2(\infty) = \infty$ and $g_2(0) = 0$. Thus g_2 must be holomorphic in $\mathbb{C}$. On the other hand, the function

$$h_2\left(\widetilde{\zeta}\right) = \frac{1}{g_2\left(\frac{1}{\widetilde{\zeta}}\right)}$$

must also be holomorphic in $\mathbb{C}$. If $k \geq 1$ is the order of the zero of g_2 at the origin, then $\frac{1}{g_2}$ has a pole of order k at the origin, and consequently its Laurent series is

$$\frac{1}{g_2(\zeta)} = \sum_{i=-k}^{+\infty} a_i \zeta^i.$$

Thus the Laurent series of h_2 is

$$h_2\left(\widetilde{\zeta}\right) = \sum_{i=-\infty}^{k} a_{-i} \widetilde{\zeta}^i$$

and we conclude that $a_i = 0$ for $i \geq 1$. Consequently,

$$g_2(\zeta) = \frac{1}{\frac{a_{-k}}{\zeta^k} + \dots + a_0} = \frac{\zeta^k}{a_{-k} + \dots + a_0 \zeta^k}$$

and in order for this function to be holomorphic one must have $a_{-k+1} = \dots = a_0 = 0$. Thus

$$g_2(\zeta) = c\zeta^k$$

for some $c \in \mathbb{C} \setminus \{0\}$, and since g_2 must be bijective in $\mathbb{C}$ we conclude that $k = 1$. Notice that

$$g_2(\zeta) = c\zeta$$

yields

$$h_2\left(\widetilde{\zeta}\right) = \frac{\widetilde{\zeta}}{c}$$

and hence h_2 is indeed holomorphic.

It is now easy to prove

Proposition 3.5. *Any complex analytic automorphism of S^2 is a composition of automorphisms represented by $\frac{1}{\zeta}$ and $a\zeta + b$ ($a \neq 0$).*

Exercise 3.6. *Use this and proposition 2.1 to prove that any complex analytic automorphism of S^2 sends circles to circles.*

4. Mobius transformations

Definition 4.1. *The group of* Mobius transformations *is the group $\mathcal{M}$ of all complex analytic automorphisms of S^2 .*

To understand the importance of this group notice that

$$\begin{aligned}
ds^2(g(\zeta)) &= \frac{4}{(1+g\overline{g})^2} dg d\overline{g} \\
&= \frac{4}{(1+g(\zeta)\overline{g}(\zeta))^2} g'(\zeta)\overline{g}'(\zeta) d\zeta d\overline{\zeta} \\
&= \frac{g'\overline{g}'\left(1+\zeta\overline{\zeta}\right)^2}{(1+g\overline{g})^2} \frac{4}{\left(1+\zeta\overline{\zeta}\right)^2} d\zeta d\overline{\zeta} \\
&= \frac{g'\overline{g}'\left(1+\zeta\overline{\zeta}\right)^2}{(1+g\overline{g})^2} ds^2(\zeta).
\end{aligned}$$

In other words, complex analytic automorphisms are conformal. Indeed, it can be shown that the group of all complex analytic automorphisms of S^2 is the same as the group of conformal orientation preserving differentiable automorphisms of S^2.

As we've seen, the Mobius group is generated by compositions of automorphisms represented by $\frac{1}{\zeta}$ and $a\zeta + b$ ($a \neq 0$). All of these are of the form

$$\frac{a\zeta + b}{c\zeta + d}$$

with $ad - bc \neq 0$ (notice that if $ad - bc = 0$ the above expression yields a constant function). Conversely, all automorphisms represented by functions of the kind above can be obtained as compositions of the automorphisms which generate the Mobius group. This is obvious if $c = 0$; if $c \neq 0$, on the other hand, one has

$$\begin{aligned} \frac{a\zeta + b}{c\zeta + d} &= \frac{ac\zeta + bc + ad - ad}{c\zeta + d} \\ &= a + \frac{bc - ad}{c\zeta + d}. \end{aligned}$$

Consequently all of the above functions represent Mobius transformations.

Consider the map $H : GL(2, \mathbb{C}) \to \mathcal{M}$ defined by

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

Exercise 4.2. *Show that H is a group homomorphism.*

In particular this proves that the set of all complex analytic automorphisms represented by the functions of the kind we've considered is in fact $\mathcal{M}$.

To compute the kernel of H we solve the equation

$$H \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \zeta \Leftrightarrow \frac{a\zeta + b}{c\zeta + d} = \zeta \Leftrightarrow b = c = 0 \text{ and } a = d.$$

Thus $\ker H = \{aI : a \in \mathbb{C} \setminus \{0\}\}$.

We know that $\mathcal{M}$ is isomorphic to

$$\frac{GL(2, \mathbb{C})}{\ker H}.$$

Let A be a representative of an equivalence class in this quotient group. Since $\det(aA) = a^2 \det A$ and $\det A \neq 0$ (as $A \in GL(2, \mathbb{C})$) we see that each equivalence class has at least one representative E with determinant 1. In fact, since $\det(aE) = a^2$, we see that each equivalence class has exactly two such representatives, $\pm E$. Since

$$SL(2, \mathbb{C}) = \{A \in GL(2, \mathbb{C}) : \det A = 1\}$$

is trivially a subgroup of $GL(2, \mathbb{C})$, we therefore conclude that $\mathcal{M}$ is isomorphic to

$$\frac{SL(2, \mathbb{C})}{\{\pm I\}}.$$

From now on we can represent any Mobius transformation by a function

$$g(\zeta) = \frac{a\zeta + b}{c\zeta + d}$$

satisfying $ad - bc = 1$.

Exercise 4.3. *Show that given such a representation every Mobius transformation with $a+d \neq \pm 2$ has exactly 2 fixed points, and every Mobius transformation with $a+d=\pm 2$ has exactly 1 fixed point. (Consider the cases $c \neq 0$ and $c=0$ separately).*

Suppose that ζ_0, ζ_1 and ζ_2 are three distinct complex numbers. Then the Mobius transformation represented by

$$g(\zeta)=\frac{\zeta_1-\zeta_0}{\zeta_1-\zeta_2}\cdot\frac{\zeta-\zeta_0}{\zeta-\zeta_2}$$

satisfies $g(\zeta_0)=0$, $g(\zeta_1)=1$ and $g(\zeta_2)=\infty$. Furthermore, if h is the representation of any other Mobius transformation satisfying the same conditions then $i=h\circ g^{-1}$ represents a Mobius transformation satisfying $i(0)=0$, $i(1)=1$ and $i(\infty)=\infty$. Setting

$$i(\zeta)=\frac{a\zeta+b}{c\zeta+d}$$

we see that

$$\begin{aligned} &i(0)=0 \Rightarrow b=0;\\ &i(\infty)=\infty \Rightarrow c=0;\\ &i(1)=1 \Rightarrow \frac{a}{d}=1 \end{aligned}$$

i.e., i is the identity, and hence $h=g$.

Exercise 4.4. *Use this to prove that any Mobius transformation is completely determined by the three (distinct) images of three distinct points in S^2.*

5. Mobius transformations and the Lorentz group

If $\{\mathbf{e}_0,\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$ is an orthonormal basis for Minkowski spacetime and

$$v=v^0\mathbf{e}_0+v^1\mathbf{e}_1+v^2\mathbf{e}_2+v^3\mathbf{e}_3$$

is a vector, then we associate to v (and to this basis) the matrix

$$V=\frac{1}{\sqrt{2}}\begin{pmatrix} v^0+v^3 & v^1+iv^2 \\ v^1-iv^2 & v^0-v^3\end{pmatrix} \tag{1}$$

(we shall explain the $\frac{1}{\sqrt{2}}$ factor in identification (1) later). Notice that $V\in\mathbb{H}_2$ (here $\mathbb{H}_2$ is the set of all Hermitian 2×2 complex matrices, i.e., all 2×2 complex matrices V satisfying $V^*=V$); in fact, the map defined above is a bijection between Minkowski spacetime and $\mathbb{H}_2$. This map is useful because

$$\det V=\frac{1}{2}\left(\left(v^0\right)^2-\left(v^3\right)^2-\left(v^1\right)^2-\left(v^2\right)^2\right)=-\frac{1}{2}\langle v,v\rangle\,.$$

As is well known $GL(2,\mathbb{C})$ acts on $\mathbb{H}_2$ through the so-called adjoint action,

$$g\cdot V=gVg^*$$

for all $g\in GL(2,\mathbb{C})$, $V\in\mathbb{H}_2$, as

$$\left(gVg^*\right)^*=\left(g^*\right)^*V^*g^*=gVg^*.$$

On the other hand,

$$\det\left(gVg^*\right)=\det g\det V\det g^*=\left|\det g\right|^2\det V$$

and thus this action preserves the determinant *iff* $|\det g| = 1$. Now any matrix $g \in GL(2,\mathbb{C})$ satisfying $|\det g| = 1$ is of the form

$$g = e^{i\frac{\theta}{2}} h$$

where

$$\det g = e^{i\theta}$$

and $h \in SL(2,\mathbb{C})$, and

$$g \cdot V = gVg^* = \left(e^{i\frac{\theta}{2}} h\right)^* V \left(e^{i\frac{\theta}{2}} h\right) = e^{-i\frac{\theta}{2}} h^* V e^{i\frac{\theta}{2}} h = h^* V h = h \cdot V.$$

Thus one gets all determinant-preserving adjoint actions of $GL(2,\mathbb{C})$ on $\mathbb{H}_2$ from the elements of $SL(2,\mathbb{C})$.

Notice that $\mathbb{H}_2$ is a vector space, and the identification (1) is clearly a linear isomorphism. On the other hand, the adjoint action of $SL(2,\mathbb{C})$ on $\mathbb{H}_2$ is easily seen to be by linear determinants-preserving maps (or, using the identification (1), by linear isometries). We therefore have a map $H : SL(2,\mathbb{C}) \to O(3,1)$. This map is a group homomorphism, as

$$H(gh)\, v = ghV(gh)^* = ghVh^*g^* = g\left(hVh^*\right) g^* = H(g)\, H(h)\, v$$

for all $V \in \mathbb{H}_2$, $g, h \in SL(2,\mathbb{C})$ (we use our identification to equate vectors on Minkowski space to Hermitian 2×2 matrices).

Exercise 5.1. *Prove that* $\ker H = \{\pm I\}$.

We now prove that $SL(2,\mathbb{C})$ is simply connected. In order to do so we'll need the following quite useful

Lemma 5.2. *Any matrix* $g \in GL(n,\mathbb{C})$ *with* $\det g > 0$ *may be written as* $g = RDS$*, where* $R, S \in SU(n)$ *and* D *is a diagonal matrix with diagonal elements in* $\mathbb{R}^+$*.*

Recall that

$$SU(n) = SL(n,\mathbb{C}) \cap U(n) = \{R \in GL(n,\mathbb{C}) : RR^* = I \text{ and } \det R = 1\}.$$

To prove this lemma we notice that if $g \in GL(n,\mathbb{C})$ then g^*g is a nonsingular positive Hermitian matrix, as

$$(g^*g)^* = g^*(g^*)^* = g^*g$$

and

$$v^*g^*gv = (gv)^* gv > 0$$

for all $v \in \mathbb{C}^n \setminus \{0\}$. Thus there exist $S \in SU(n)$ and a diagonal matrix Λ with diagonal elements in $\mathbb{R}^+$ such that

$$g^*g = S^*\Lambda S.$$

Moreover, we can write $\Lambda = D^2$ with D is a diagonal matrix with diagonal elements in $\mathbb{R}^+$. Therefore

$$\begin{aligned}
&g^*g = S^*DDS \\
&\Leftrightarrow g^*gS^*D^{-1} = S^*DDSS^*D^{-1} \\
&\Leftrightarrow gS^*D^{-1} = (g^*)^{-1} S^*D \\
&\Leftrightarrow gS^{-1}D^{-1} = \left(g^{-1}\right)^* S^*D \\
&\Leftrightarrow \left(DSg^{-1}\right)^{-1} = \left(DSg^{-1}\right)^*
\end{aligned}$$

i.e.,

$$DSg^{-1} \in U(n).$$

If $\det g > 0$ then clearly $\det\left(DSg^{-1}\right) > 0$ and consequently

$$DSg^{-1} \in SU(n) \Leftrightarrow R = \left(DSg^{-1}\right)^{-1} \in SU(n)$$

with

$$gS^{-1}D^{-1} = R \Leftrightarrow g = RDS.$$

In particular if $g \in SL(2,\mathbb{C})$ then we must have $\det D = 1$ and hence

$$D = \begin{pmatrix} x & 0 \\ 0 & \frac{1}{x} \end{pmatrix}$$

for some $x \in \mathbb{R}^+$. Notice that since x and $\frac{1}{x}$ are the eigenvalues of g^*g, they are uniquely determined up to ordering.

Let $g : [0,1] \to SL(2,\mathbb{C})$ be a continuous path satisfying $g(0) = g(1) = I$. For each value of t one can use the decomposition above to get

$$g(t) = R(t) \begin{pmatrix} x(t) & 0 \\ 0 & \frac{1}{x(t)} \end{pmatrix} S(t)$$

and it is clear that $x(t)$ is continuous and $x(0) = x(1) = 1$. Since $\mathbb{R}^+$ is simply connected we can continuously deform this closed path into the constant path $x(t) = 1$, thus continuously deforming $g(t)$ into $R(t)S(t) \in SU(2)$ (which consequently is a continuous closed path, even if $R(t)$ and $S(t)$ by themselves are not). We conclude that if $SU(2)$ is simply connected then $SL(2,\mathbb{C})$ is simply connected as well.

Exercise 5.3. *Show that*

$$SU(2) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} : (a,b) \in \mathbb{C}^2 \text{ and } |a|^2 + |b|^2 = 1 \right\}$$

and that therefore $SU(2)$ is a smooth manifold diffeomorphic to S^3. Conclude that $SU(2)$ (and hence $SL(2,\mathbb{C})$) is simply connected.

A similar technique can be employed to show that $O_+^{\uparrow}(3,1)$ is pathwise connected: if L is a proper Lorentz transformation then clearly

$$L\mathbf{e}_0 = \cosh u \mathbf{e}_0 + \sinh u \mathbf{e}$$

for some $u \geq 0$ and $\mathbf{e} \in (\mathbf{e}_0)^{\perp}$. If R is any rotation (i.e., any proper Lorentz transformation preserving $\mathbf{e}_0$) sending $\mathbf{e}_3$ to $\mathbf{e}$, we have

$$R^{-1}L\mathbf{e}_0 = \cosh u \mathbf{e}_0 + \sinh u \mathbf{e}_3.$$

Thus if B is a boost in the $\mathbf{e}_3$ direction by a hyperbolic angle u, we have

$$B^{-1}R^{-1}L\mathbf{e}_0 = \mathbf{e}_0$$

and consequently $S = B^{-1}R^{-1}L$ is a rotation, and $L = RBS$.

Exercise 5.4. *Use the decomposition above to show that $O_+^{\uparrow}(3,1)$ is pathwise connected. However, one cannot use this decomposition to conclude that $O_+^{\uparrow}(3,1)$ is simply connected (in a similar fashion to what was done for $SL(2,\mathbb{C})$). Why not?*

Exercise 5.5. *Show that $H(SL(2,\mathbb{C})) \subseteq O_+^{\uparrow}(3,1)$ (hint: start by showing that H is continuous).*

We now compute the dimension of $SL\left(2,\mathbb{C}\right)$ by computing its tangent space at the identity. Let $g:\left(-\varepsilon,\varepsilon\right)\to SL\left(2,\mathbb{C}\right)$ be a path satisfying $g\left(0\right)=I$. If we set

$$g\left(t\right)=\left(\begin{array}{cc} a\left(t\right) & b\left(t\right) \\ c\left(t\right) & d\left(t\right) \end{array}\right)$$

we have

$$a\left(0\right)=d\left(0\right)=1,b\left(0\right)=c\left(0\right)=0$$

and

$$\begin{aligned}
&a\left(t\right)d\left(t\right)-c\left(t\right)d\left(t\right)=1\\
&\Rightarrow \dot{a}\left(t\right)d\left(t\right)+a\left(t\right)\dot{d}\left(t\right)-\dot{c}\left(t\right)d\left(t\right)-c\left(t\right)\dot{d}\left(t\right)=0\\
&\Rightarrow \dot{a}\left(0\right)+\dot{d}\left(0\right)=0
\end{aligned}$$

(where the dot represents differentiation with respect to t), indicating that $T_I SL\left(2,\mathbb{C}\right)$ can be identified with the vector space of traceless 2×2 complex matrices. This vector space has real dimension 6, and therefore we conclude that $SL\left(2,\mathbb{C}\right)$ is a 6-dimensional real manifold.

Analogously we determine the dimension of $O_+^{\uparrow}\left(3,1\right)$ by computing its tangent space at the identity. If $\Lambda:\left(-\varepsilon,\varepsilon\right)\to O_+^{\uparrow}\left(3,1\right)$ is a path satisfying $\Lambda\left(0\right)=I$ then

$$\begin{aligned}
\Lambda^t\left(t\right)\eta\Lambda\left(t\right) &= \eta\\
&\Rightarrow \dot{\Lambda}^t\left(t\right)\eta\Lambda\left(t\right)+\Lambda^t\left(t\right)\eta\,\dot{\Lambda}\left(t\right)=0\\
&\Rightarrow \dot{\Lambda}^t\left(0\right)\eta+\eta\,\dot{\Lambda}\left(0\right)=0
\end{aligned}$$

and we then see that $T_I O_+^{\uparrow}\left(3,1\right)$ can be identified with the vector space of 4×4 real matrices A satisfying

$$A^t\eta+\eta A=0\Leftrightarrow\left(\eta A\right)^t+\eta A=0$$

i.e., such that ηA is skew-symmetric. Since η is nonsingular, we conclude that the dimension of $T_I O_+^{\uparrow}\left(3,1\right)$ is equal to the dimension of the vector space of 4×4 real skew-symmetric matrices, i.e., 6.

Both $SL\left(2,\mathbb{C}\right)$ and $O_+^{\uparrow}\left(3,1\right)$ are connected Lie groups, and $H:SL(2,\mathbb{C})\to O_+^{\uparrow}\left(3,1\right)$ is a Lie group homomorphism (i.e., is a smooth map which is a group homomorphism). Because they have the same dimension and $\ker H$ is finite it follows that H is surjective in a neighbourhood of the identity, i.e., is a *local isomorphism*.

It is a theorem by Lie that up to topology all locally isomorphic connected Lie groups are the same. More accurately, two locally isomorphic connected Lie groups have the same *universal covering*, where the universal covering of a connected Lie group G is the unique Lie group U which is locally isomorphic to G and simply connected. In that case there exists a surjective projection homomorphism $h:U\to G$ extending uniquely the local isomorphism.

In our case one then has that $SL(2,\mathbb{C})$ is the universal covering of $O_+^{\uparrow}\left(3,1\right)$, H is surjective and

$$O_+^{\uparrow}\left(3,1\right)=\frac{SL(2,\mathbb{C})}{\ker H}=\frac{SL\left(2,\mathbb{C}\right)}{\{\pm I\}}=\mathcal{M}.$$

We summarize this in the following

Theorem 5.6. *The group of proper Lorentz transformations $O_+^{\uparrow}\left(3,1\right)$ is isomorphic to the group of Mobius transformations $\mathcal{M}$.*

It may sound a bit strange that transformations between proper inertial observers are the same thing as conformal motions of the 2-sphere. Actually this relation is surprisingly natural, as we shall see.

6. Lie algebra of the Lorentz group

If G is a Lie group, its tangent space at the identity $\mathfrak{g} = T_I G$ can be given the structure of an algebra (called the *Lie algebra* of G) by introducing the so-called *Lie bracket*. In all the cases we've seen G was a group of matrices, and hence $\mathfrak{g}$ was a vector space of matrices. In this case the Lie bracket is just the ordinary commutator of two matrices: if $A, B \in \mathfrak{g}$ then

$$[A, B] = AB - BA.$$

It is a theorem by Lie that two Lie groups have the same Lie algebra *iff* they are locally isomorphic. Thus to study the Lie algebra $\mathfrak{o}\,(3,1)$ of the Lorentz group $O\,(3,1)$ we can simply study the Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ of $SL(2,\mathbb{C})$. We saw that $\mathfrak{sl}(2,\mathbb{C})$ is the space of all traceless 2×2 complex matrices, and thus it is not only a real vector space of dimension 6 but also a complex vector space of complex dimension 3. A convenient complex basis for $\mathfrak{sl}(2,\mathbb{C})$ is given by the so-called *Pauli matrices*,

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

These are Hermitian traceless square roots of the identity: one has

$$(\sigma_k)^2 = I$$

for $k = 1, 2, 3$. In fact, $\Sigma = \{I, \sigma_1, \sigma_2, \sigma_3, iI, i\sigma_1, i\sigma_2, i\sigma_3\}$ form a group under matrix multiplication.

Exercise 6.1. *Check that the multiplication table*

$$\begin{array}{cccc}
\cdot & \sigma_1 & \sigma_2 & \sigma_3 \\
\sigma_1 & I & i\sigma_3 & -i\sigma_2 \\
\sigma_2 & -i\sigma_3 & I & i\sigma_1 \\
\sigma_3 & i\sigma_2 & -i\sigma_1 & I
\end{array}$$

is correct. Use it to show that Σ is indeed a group and complete its multiplication table, and to check that the commutation relations

$$\begin{aligned}
[\sigma_1, \sigma_2] &= 2i\sigma_3; \\
[\sigma_2, \sigma_3] &= 2i\sigma_1; \\
[\sigma_3, \sigma_1] &= 2i\sigma_2
\end{aligned}$$

hold.

To get a real basis for $\mathfrak{sl}(2,\mathbb{C})$ we can take the matrices

$$\begin{aligned} B_k &= \frac{1}{2}\sigma_k; \\ R_k &= -\frac{i}{2}\sigma_k \end{aligned}$$

($k = 1, 2, 3$), where the $\frac{1}{2}$ factors were introduced to simplify the commutation relations. The elements of a basis of a Lie algebra are often called *generators* of the algebra.

Exercise 6.2. *Show that the commutation relations*

$$\begin{aligned} &[B_1, B_2] = -R_3;\ [B_2, B_3] = -R_1;\ [B_3, B_1] = -R_2; \\ &[R_1, R_2] = R_3;\ [R_2, R_3] = R_1;\ [R_3, R_1] = R_2; \\ &[B_1, R_2] = B_3;\ [B_2, R_3] = B_1;\ [B_3, R_1] = B_2; \\ &[R_1, B_2] = B_3;\ [R_2, B_3] = B_1;\ [R_3, B_1] = B_2 \end{aligned}$$

hold.

Notice in particular that the real space spanned by $\{R_1, R_2, R_3\}$ is closed with respect to the Lie bracket, and thus forms a *Lie subalgebra* of $\mathfrak{sl}(2,\mathbb{C})$. This corresponds to the Lie subgroup $SU(2,\mathbb{C})$ of $SL(2,\mathbb{C})$ (or alternatively to the Lie subgroup $SO(3)$ of $O_+^\uparrow(3,1)$), as we shall see.

If G is a Lie group of matrices and $\mathfrak{g}$ is its Lie algebra then $e^{At} \in G$ for all $A \in \mathfrak{g}$ and $t \in \mathbb{R}$, and in fact all elements of G are of this form. Then the entire Lie group can be obtained from its Lie algebra by exponentiation (this is the basic fact underlying Lie's theorems).

The Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ can thus be made to act on Minkowski space through the so-called *infinitesimal action*

$$\begin{aligned} A \cdot v &= \frac{d}{dt}\left(e^{At} \cdot v\right)|_{t=0} \\ &= \frac{d}{dt}\left(e^{At}V\left(e^{At}\right)^*\right)|_{t=0} \\ &= \frac{d}{dt}\left(e^{At}Ve^{A^*t}\right)|_{t=0} \\ &= \left(Ae^{At}Ve^{A^*t} + e^{At}VA^*e^{A^*t}\right)|_{t=0} \\ &= AV + AV^* \end{aligned}$$

(where once again we use identification (1)). In particular if A is one of the above generators, and noticing that since the Pauli matrices are Hermitian one has

$$\begin{aligned} (B_k)^* &= B_k; \\ (R_k)^* &= -R_k \end{aligned}$$

($k = 1, 2, 3$), we see that

$$\begin{aligned} B_k \cdot v &= B_kV + VB_k = \{B_k, V\}; \\ R_k \cdot v &= R_kV + VR_k = [R_k, V]. \end{aligned}$$

Example 6.3. *One then has*

$$
\begin{aligned}
B_3 \cdot v &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix} v^0+v^3 & v^1+iv^2 \\ v^1-iv^2 & v^0-v^3 \end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix} v^0+v^3 & v^1+iv^2 \\ v^1-iv^2 & v^0-v^3 \end{pmatrix} \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
&= \frac{1}{2\sqrt{2}}\begin{pmatrix} v^0+v^3 & v^1+iv^2 \\ -v^1+iv^2 & -v^0+v^3 \end{pmatrix} + \frac{1}{2\sqrt{2}}\begin{pmatrix} v^0+v^3 & -v^1-iv^2 \\ v^1-iv^2 & -v^0+v^3 \end{pmatrix} \\
&= \frac{1}{\sqrt{2}}\begin{pmatrix} v^0+v^3 & 0 \\ 0 & -v^0+v^3 \end{pmatrix} = \begin{pmatrix} v^3 \\ 0 \\ 0 \\ v^0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix}
\end{aligned}
$$

and it is therefore clear that

$$
\begin{aligned}
e^{uB_3} &= \sum_{n=0}^{+\infty} \frac{1}{n!}\left(\frac{u}{2}\right)^n \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^n \\
&= \sum_{n=0}^{+\infty} \frac{1}{n!}\begin{pmatrix} \left(\frac{u}{2}\right)^n & 0 \\ 0 & \left(-\frac{u}{2}\right)^n \end{pmatrix} \\
&= \begin{pmatrix} e^{\frac{u}{2}} & 0 \\ 0 & e^{-\frac{u}{2}} \end{pmatrix} \in SL(2,\mathbb{C})
\end{aligned}
$$

corresponds to the Lorentz transformation represented by

$$
\begin{aligned}
\exp\left(u\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\right) &= \sum_{n=0}^{+\infty} \frac{u^n}{n!}\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}^n \\
= \sum_{n=0}^{+\infty} \frac{u^{2n}}{(2n)!}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} &+ \sum_{n=0}^{+\infty} \frac{u^{2n+1}}{(2n+1)!}\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\
= \cosh u\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} &+ \sinh u\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\
= \begin{pmatrix} \cosh u & 0 & 0 & \sinh u \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \sinh u & 0 & 0 & \cosh u \end{pmatrix}
\end{aligned}
$$

i.e., a boost in the $\mathbf{e}_3$ *direction by a hyperbolic angle* u *(which can then be identified with the Mobius transformation* $e^u\zeta$*). For this reason one says that* B_3 *generates boosts in the* $\mathbf{e}_3$ *direction.*

Exercise 6.4. *Use the same method to show that* B_1 *and* B_2 *generate boosts in the* $\mathbf{e}_1$ *and* $-\mathbf{e}_2$ *directions, respectively, and that* R_1*,* R_2 *and* R_3 *generate rotations about* $-\mathbf{e}_1$*,* $\mathbf{e}_2$ *and* $-\mathbf{e}_3$*. Show that the elements of* $SL(2,\mathbb{C})$ *corresponding to these Lorentz transformations* *by a hyperbolic angle u or an angle θ are*

$$e^{uB_1} = \begin{pmatrix} \cosh\left(\frac{u}{2}\right) & \sinh\left(\frac{u}{2}\right) \\ \sinh\left(\frac{u}{2}\right) & \cosh\left(\frac{u}{2}\right) \end{pmatrix};$$
$$e^{uB_2} = \begin{pmatrix} \cosh\left(\frac{u}{2}\right) & -i\sinh\left(\frac{u}{2}\right) \\ i\sinh\left(\frac{u}{2}\right) & \cosh\left(\frac{u}{2}\right) \end{pmatrix};$$
$$e^{\theta R_1} = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -i\sin\left(\frac{\theta}{2}\right) \\ -i\sin\left(\frac{u}{2}\right) & \cos\left(\frac{u}{2}\right) \end{pmatrix};$$
$$e^{\theta R_2} = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{u}{2}\right) & \cos\left(\frac{u}{2}\right) \end{pmatrix};$$
$$e^{\theta R_3} = \begin{pmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{pmatrix}.$$

Notice in particular that the R_k generators do generate the subgroup of rotations of $O_+^\uparrow(3,1)$. Notice also that a rotation about $\mathbf{e}_3$ by an angle θ is the same thing as a rotation about $-\mathbf{e}_3$ by an angle $-\theta$, and hence can be identified with the Mobius transformation $e^{i\theta}\zeta$.

7. Spinors

If we take a column vector
$$k = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \in \mathbb{C}^2$$
the matrix
$$kk^* = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \begin{pmatrix} \overline{\xi} & \overline{\eta} \end{pmatrix} = \begin{pmatrix} \xi\overline{\xi} & \xi\overline{\eta} \\ \eta\overline{\xi} & \eta\overline{\eta} \end{pmatrix}$$
is Hermitian, as $(kk^*)^* = kk^*$. Thus it represents a vector in Minkowski space. Since
$$\det(kk^*) = \xi\overline{\xi}\eta\overline{\eta} - \xi\overline{\eta}\eta\overline{\xi}$$
we see that it represents a null vector. More explicitly, such vector v satisfies
$$\frac{1}{\sqrt{2}}\begin{pmatrix} v^0+v^3 & v^1+iv^2 \\ v^1-iv^2 & v^0-v^3 \end{pmatrix} = \begin{pmatrix} \xi\overline{\xi} & \xi\overline{\eta} \\ \eta\overline{\xi} & \eta\overline{\eta} \end{pmatrix}$$
and thus
$$\begin{aligned} v^0 &= \frac{1}{\sqrt{2}}\left(\xi\overline{\xi}+\eta\overline{\eta}\right); \\ v^1 &= \frac{1}{\sqrt{2}}\left(\xi\overline{\eta}+\eta\overline{\xi}\right); \\ v^2 &= \frac{1}{i\sqrt{2}}\left(\xi\overline{\eta}-\eta\overline{\xi}\right); \\ v^3 &= \frac{1}{\sqrt{2}}\left(\xi\overline{\xi}-\eta\overline{\eta}\right). \end{aligned}$$

We see that $v^0 > 0$ for any choice of $k \in \mathbb{C}^2\setminus\{0\}$, and thus $\mathbb{C}^2\setminus\{0\}$ parametrizes (non-injectively) a subset of the future light cone of the origin. In fact, it parametrizes the whole future light cone: if we take $v^0 > 0$, the vectors in the light cone with this $\mathbf{e}_0$ component satisfy

$$\left(v^1\right)^2 + \left(v^2\right)^2 + \left(v^3\right)^2 = \left(v^0\right)^2$$

and thus the point

$$\left(\frac{v^1}{v^0}, \frac{v^2}{v^0}, \frac{v^3}{v^0}\right) \in \mathbb{R}^3$$

is in S^2. If $v^3 \neq v^0$ its stereographic projection is

$$\frac{\frac{v^1}{v^0} + i\frac{v^2}{v^0}}{1 - \frac{v^3}{v^0}} = \frac{v^1 + iv^2}{v^0 - v^3}$$

and consequently any vector v in the future light cone is represented by

$$k = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \in \mathbb{C}^2$$

where ξ and η are two complex numbers satisfying

$$|\xi|^2 + |\eta|^2 = \sqrt{2}v^0 \text{ and } \frac{\xi}{\eta} = \frac{v^1 + iv^2}{v^0 - v^3},$$

which can always be arranged.

Exercise 7.1. *Show that if* $v^3 = v^0$ *then* $v^1 = v^2 = 0$ *and*

$$k = \begin{pmatrix} \sqrt{\sqrt{2}v^0} \\ 0 \end{pmatrix}$$

parametrizes v*.*

Exercise 7.2. *Show that if* $k, l \in \mathbb{C}^2 \setminus \{0\}$ *then they parametrize the same null vector iff* $k = e^{i\theta} l$ *for some* $\theta \in \mathbb{R}$*. Conclude that the future light cone of the origin is bijectively parametrized by* $\frac{\mathbb{C}^2 \setminus \{0\}}{S^1}$*.*

Now if $g \in SL(2, \mathbb{C})$ its action on the null vector parametrized by $k \in \mathbb{C}^2 \setminus \{0\}$ is given by

$$g\left(kk^*\right) g^* = (gk)\,(gk)^*$$

i.e., is the null vector parametrized by gk.

Definition 7.3. *A vector* $k \in \mathbb{C}^2$ *plus the map*

$$k \mapsto kk^* = V = v$$

(where v *is a null vector in Minkowski space) is called a* spinor*.*

As we've seen, nonvanishing spinors can be thought of as "square roots" of future-pointing null vectors plus a phase factor. Notice that the way in which spinors parametrize future-pointing null vectors depends on identification (1), which itself depended on the choice of a basis for Minkowski space. Hence spinors are always associated with a basis for Minkowski space. It is also possible to define spinors in General Relativity if the spacetime we are considering is non-compact and has a (smoothly varying) orthonormal frame at each tangent space (which one can identify with a basis of Minkowski space). In that way one gets a vector bundle with fibre $\mathbb{C}^2$ called the *spin bundle*, in which it is possible to define a *spin connection*. Using this connection one can write Einstein's equations in spinor form. Partly because of the simple way in which a Lorentz transformation acting on a null vector parametrized by a spinor is represented by multiplication of the corresponding matrix in $SL(2,\mathbb{C})$ by the spinor, these equations are particularly simple. Many times very complicated solutions of Einstein's equations can be found by using spinor methods (particularly spacetimes possessing certain kinds of congruencies of null geodesics). For more details see [PR].

Finally, notice that the spinors

$$o = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \iota = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

correspond to the null vectors

$$l = \frac{1}{\sqrt{2}}\left(\mathbf{e}_0 + \mathbf{e}_3\right) \text{ and } n = \frac{1}{\sqrt{2}}\left(\mathbf{e}_0 - \mathbf{e}_3\right)$$

and these satisfy the normalization condition

$$\langle l, n \rangle = -1.$$

This is the reason for the $\frac{1}{\sqrt{2}}$ factor in (1).

8. The sky of a rapidly moving observer

Let us now think of a light ray through the origin. All nonvanishing (null) vectors v in this light ray are multiples of each other, and thus

$$\mathbf{e} = \left(\frac{v^1}{v^0}, \frac{v^2}{v^0}, \frac{v^3}{v^0}\right) \in S^2$$

is the same for all of them. Thus the set S^+ of all light rays through the origin is a sphere S^2, which we can identify with $\mathbb{C} \cup \{\infty\}$.

If

$$k = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

is a spinor parametrizing a future-pointing null vector in the light ray, we saw that the stereographic projection of $\mathbf{e}$ is

$$\zeta = \frac{\xi}{\eta}$$

(provided that $v^3 \neq v^0$). Thus to get the point in $\mathbb{C} \cup \{\infty\}$ corresponding to a light ray containing the null vector parametrized by a spinor k using this stereographic projection we have but to divide its components.

If

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

the corresponding Lorentz transformation takes the null vector parametrized by k to the null vector parametrized by

$$gk = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} a\xi + b\eta \\ c\xi + d\eta \end{pmatrix}$$

i.e., takes the light ray represented by ζ to the light ray represented by

$$\frac{a\xi + b\eta}{c\xi + d\eta} = \frac{a\zeta + b}{c\zeta + d}$$

(a Mobius transformation!). Thus we have proved the following

Proposition 8.1. *Any proper Lorentz transformation is completely determined by its action on the set S^+ of light rays through the origin. More specifically, the group of proper Lorentz transformations can be thought of as the group of orientation-preserving conformal motions of the 2-sphere S^+.*

To understand how the skies of two observers are related, one must have two things in mind: the first is that if $g \in SL(2, \mathbb{C})$ represents the *active* Lorentz transformation relating the two observers then this change is accomplished by the corresponding *passive* Lorentz transformation (represented by g^{-1}). The second is that the sky of an observer is not actually S^+, but the image S^- of S^+ under the antipodal map, for the simple reason that an observer places an object whose light is moving in direction $\mathbf{e}$ in position $-\mathbf{e}$ of his celestial sphere.

As we've seen, using spherical coordinates (θ, φ) in (an appropriate open subset of) S^2, the stereographic projection is given by

$$\zeta(\theta, \varphi) = \frac{\sin\theta}{1 - \cos\theta} e^{i\varphi}.$$

Consequently, the antipodal map $(\theta, \varphi) \mapsto (\pi - \theta, \varphi + \pi)$ is given by

$$\zeta \mapsto A(\zeta) = -\frac{\sin\theta}{1 + \cos\theta} e^{i\varphi}$$

and hence

$$\overline{\zeta} A(\zeta) = -\frac{\sin^2\theta}{1 - \cos^2\theta} = -1 \Leftrightarrow A(\zeta) = -\frac{1}{\overline{\zeta}}.$$

Consequently if the active Lorentz transformation relating the two observers corresponds to the Mobius transformation

$$g(\zeta) = \frac{a\zeta + b}{c\zeta + d}$$

of S^+, then the corresponding change of the observer's celestial sphere corresponds to the Mobius transformation

$$A \circ g^{-1} \circ A(\zeta)$$

of S^-. Since

$$g^{-1}(\zeta) = \frac{d\zeta - b}{-c\zeta + a}$$

we have

$$\begin{aligned} A \circ g^{-1} \circ A(\zeta) &= A\left(\frac{-d\overline{\zeta}^{-1} - b}{c\overline{\zeta}^{-1} + a}\right) \\ &= \frac{\overline{c}\zeta^{-1} + \overline{a}}{\overline{d}\zeta^{-1} + \overline{b}} \\ &= \frac{\overline{a}\zeta + \overline{c}}{\overline{b}\zeta + \overline{d}}. \end{aligned}$$

Thus we have proved the following

Theorem 8.2. *If the active Lorentz transformation relating two observers is represented by $g \in SL(2, \mathbb{C})$ then the observers' celestial spheres are related by the Mobius transformation corresponding to g^*.*

Example 8.3. *Recall that*

$$e^{\theta R_3} = \begin{pmatrix} e^{i\frac{\theta}{2}} & 0 \\ 0 & e^{-i\frac{\theta}{2}} \end{pmatrix} \in SL(2,\mathbb{C})$$

corresponds to a rotation about $\mathbf{e}_3$ *by an angle* θ. *Consequently the sky of the rotated observer is given by applying to the sky of the initial observer the Mobius transformation corresponding to*

$$\left(e^{\theta R_3}\right)^* = \begin{pmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{pmatrix}$$

i.e., $e^{-i\theta}\zeta$. *This clearly corresponds to rotating the celestial sphere by an angle* $-\theta$ *about* $\mathbf{e}_3$, *as should be expected (if an observer is rotated one way, he sees his celestial sphere rotating the opposite way).*

Exercise 8.4. *Show that the sky of an observer moving in the* $\mathbf{e}_3$ *direction with speed* $\tanh u$ *is obtained from the sky of an observer at rest by the Mobius transformation* $e^u\zeta$. *Thus objects in the sky of a rapidly moving observer accumulate towards the direction of motion, an effect known as* aberration.

Exercise 8.5. *As seen from Earth, the Sun has an angular diameter of half a degree. What is the angular diameter an observer speeding past the Earth at* 96% *of the speed of light would measure for the Sun?*

Because any proper Lorentz transformation can be decomposed in two rotations and one boost, we see that the general transformation of the sky of an observer is the composition of this aberration effect with two rigid rotations. In addition to the aberration, there is also a Doppler shift due to the fact that the energy of the photon correspondent to the null vector $v = kk^*$ as measured by the two observers is different. The ratio of these energies (which equals the ratio of their frequencies) is

$$\delta = \frac{Lv^0}{v^0}$$

where L is the active Lorentz transformation relating the two observers. If L is represented by

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

then on S^+ one gets

$$\begin{aligned} \delta &= \frac{\|k\|^2}{\|gk\|^2} = \frac{|\xi|^2 + |\eta|^2}{|a\xi + b\eta|^2 + |c\xi + d\eta|^2} \\ &= \frac{|\zeta|^2 + 1}{|a\zeta + b|^2 + |c\zeta + d|^2}. \end{aligned}$$

Exercise 8.6. *Show that if* L *is a rotation then* $\delta = 1$ *(hint: recall that* L *is a rotation* iff $g \in SU(2)$*).*

Exercise 8.7. *Show that if* L *is a boost in the* $\mathbf{e}_3$ *direction by a hyperbolic angle* $u > 0$ *then*

$$\delta = \frac{|\zeta|^2 + 1}{e^u |\zeta|^2 + e^{-u}}.$$

Show that in S^- this becomes

$$\delta = \frac{|\zeta|^2 + 1}{e^{-u}|\zeta|^2 + e^u}.$$

so that the Doppler shift ratio is maximum for $\zeta = \infty$ and minimum for $\zeta = 0$.

In addition to this Doppler shift, one also gets intensity shifts, as both observers get different numbers of photons per unit time from a given direction. This is due not only to the difference in their proper times but also to the fact that their motions differ; however, we shall not pursue this matter any further.

Theorems about Mobius transformations can be readily transformed into theorems about skies. For instance, the theorem stating that any Mobius transformation is completely determined by the image of three distinct points translates as

Theorem 8.8. *If an observer sees three stars in his sky and specifies a new position for each star, there exists a unique observer who sees the three stars in these positions.*

The fact that Mobius transformations are conformal transformations translates into

Theorem 8.9. *Small objects are seen by different observers as having the same exact shape. If $h \in SL(2,\mathbb{C})$ represents the active Lorentz transformation relating the two observers and $g(\zeta)$ is the Mobius transformation corresponding to h^* then the magnification factor at each point of the observer's sky is given by the formula*

$$ds^2(g(\zeta)) = \frac{g'\overline{g}'\left(1+\zeta\overline{\zeta}\right)^2}{(1+g\overline{g})^2} ds^2(\zeta).$$

Exercise 8.10. *Show that for a boost in the $\mathbf{e}_3$ direction by a hyperbolic angle $u > 0$ the magnification factor is given by the formula*

$$ds^2(g(\zeta)) = \frac{e^{2u}\left(1+|\zeta|^2\right)^2}{\left(1+e^{2u}|\zeta|^2\right)^2} ds^2(\zeta)$$

so that it is e^{-u} for $\zeta = \infty$ and e^u for $\zeta = 0$.

Perhaps the most surprising statement of this kind is the sky version of the theorem stating that Mobius transformations take circles into circles:

Theorem 8.11. *If an observer sees a circular outline for any object on his sky then any observer sees a circular outline for this object.*

References

[PR] Penrose, R. & Rindler, W., *Spinors and Space-time*, Vols. 1 & 2, Cambridge University Press (1986).