Notes on the structure of the Lorentz group
(preliminary and unfinished version)

Claudio Bartocci

Contents

1 Generalities 2

2 The polar decomposition theorem for the Lorentz group 3

3 An interlude: the case of \( O(1, 1) \) 8

4 The topological structure of \( O(p, 1) \) 10

5 Hyperbolic geometry 14

6 The Poincaré disc 17

©2014 Claudio Bartocci
1 Generalities

Let \((V, g)\) be a pseudo-Euclidean vector space of signature \((p, q)\); the dimension of \(V\) is \(n = p + q\). As we proved in [2, §8], after suitably choosing an orthonormal basis, the matrix representing the pseudo-Euclidean product \(g\) is a diagonal matrix of the form:

\[
\eta_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}
\]

where \(I_r, r = p, q\), is the \(r \times r\) identity matrix.

The group of \(g\)-isometries of \(V\) is isomorphic to the matrix group

\[
O(p, q) = \{ \Lambda \in GL(p + q; \mathbb{R}) \mid \Lambda^T \eta_{p,q} \Lambda = \eta_{p,q} \} .
\]

As customary, we denote by \(SO(p, q)\) the subgroup of \(O(p, q)\) consisting of matrices whose determinant is equal to 1. When \(q = 0\), we recover the familiar definition of the orthogonal group:

\[
O(n, 0) \equiv O(n) = \{ \Lambda \in GL(n; \mathbb{R}) \mid \Lambda^T I_n \Lambda = I_n \} .
\]

It is immediate to see that, if \(\Lambda \in O(p, q)\), then \(\Lambda^T \in O(p, q)\) as well; moreover, since \(\eta_{p,q} = -\eta_{q,p}\), the inverse of \(\Lambda\) is

\[
\Lambda^{-1} = \eta_{p,q} \Lambda^T \eta_{p,q} .
\]

When \(R \in O(n)\), the previous formula reduces to the usual one, i.e. \(R^{-1} = R^T\).

From the fact that \(\eta_{p,q} = -\eta_{q,p}\), it follows that the groups \(O(p, q)\) and \(O(q, p)\) are isomorphic for any \(p, q\); in particular, \(O(n) = O(0, n)\). We shall make use of the standard immersions \(O(p, q) \hookrightarrow O(p + 1, q)\) and \(O(p, q) \hookrightarrow O(p, q + 1)\), which are defined, respectively, by the assignments

\[
\Lambda \in O(p, q) \mapsto \begin{pmatrix} 1 & 0 \\ 0 & \Lambda \end{pmatrix} \in O(p + 1, q)
\]

and

\[
\Lambda \in O(p, q) \mapsto \begin{pmatrix} \Lambda & 0 \\ 0 & 1 \end{pmatrix} \in O(p, q + 1) .
\]

In particular, when \(R \in O(p)\), we have \(\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} \in O(p, q + 1)\).

**Theorem 1.1.** The matrix group \(O(p, q)\) has a natural structure of closed Lie subgroup of \(GL(p + q; \mathbb{R})\); its dimension is

\[
\dim O(p, q) = \frac{(p + q)(p + q - 1)}{2} .
\]

The standard immersions \(O(p, q) \hookrightarrow O(p + 1, q)\) and \(O(p, q) \hookrightarrow O(p, q + 1)\) are embeddings.
Proof. Let $\Sigma_{p+q}$ be the vector space of real $(p+q) \times (p+q)$ symmetric matrix; its dimension is $\frac{(p+q)(p+q+1)}{2}$. Let us introduce the map

$$H : GL(p+q, \mathbb{R}) \to \Sigma_{p+q}$$

defined by $H(\Lambda) = \Lambda^T \eta_{p,q} \Lambda$. Then $O(p,q) = H^{-1}(\eta_{p,q})$. The map $H$ is smooth and its differential at the point $\Lambda$ is given by

$$dH_\Lambda(\alpha) = \frac{d}{dt} \bigg|_{t=0} H(\Lambda \exp(t\alpha)) = \alpha^T \Lambda^T \eta_{p,q} \Lambda + \Lambda^T \eta_{p,q} \Lambda \alpha$$

for any $\alpha \in \mathfrak{gl}(p+q)$. Now, if $\Lambda$ satisfies the relation $\Lambda^T \eta_{p,q} \Lambda = \eta_{p,q}$, we get

$$dH_\Lambda(\alpha) = \alpha^T \eta_{p,q} + \eta_{p,q} \alpha.$$

This shows that the differential $dH_\Lambda : \mathfrak{gl}(p+q) \to \Sigma_{p+q}$ is surjective: indeed, given any symmetric matrix $S$, it suffices to take $\alpha = (S\eta_{p,q})/2$. By [6, Theorem 1.38], it follows that $O(p,q) = H^{-1}(\eta_{p,q})$ has a unique structure of embedded submanifold of $GL(p+q, \mathbb{R})$ whose dimension is

$$(p+q)^2 - \frac{(p+q)(p+q+1)}{2} = \frac{(p+q)(p+q-1)}{2}.$$ 

It is easy to conclude (see e.g. [6, Theorem 3.34]) that $O(p,q)$ is a (closed) Lie subgroup of $GL(p+q, \mathbb{R})$. The last claim is straightforward. ▲

2 The polar decomposition theorem for the Lorentz group

From now on, we restrict our attention to the (generalized) Lorentz group $O(p,1)$. In order to make notation less cumbersome, the matrix $\eta_{p,1}$ will be indicated simply by $\eta$; vectors in $\mathbb{R}^{p+1}$ will be denoted by $U$, $V$, $X$, $Y$, ..., and they will be thought as column vectors; so, e.g., we shall write $X = (X_1, \ldots, X_{p+1})^T$. We fix the immersion $\mathbb{R}^p \to \mathbb{R}^{p+1}$ defined by $(x_1, \ldots, x_p) \to (x_1, \ldots, x_p, 0)$; column vectors in $\mathbb{R}^p$ will be denoted by $b$, $c$, ..., $u$, $v$, $x$, $y$, ...; hence, a vector $X$ can be written also in the block form $X = \begin{bmatrix} x \\ \alpha \end{bmatrix}$ with $\alpha \in \mathbb{R}$. The Lorentz group $O(p,1)$ is the isometry group of the Minkowski product (defined by $\eta$) on $\mathbb{R}^{p+1}$, namely the pseudo-Euclidean inner product defined by

$$X \star \eta Y = \sum_{i,j=1}^{p+1} \eta_{ij} X_i Y_j = \sum_{i=1}^p X_i Y_i - X_{p+1} Y_{p+1}.$$ 

The usual Euclidean inner product between two vectors $x$, $y$ in $\mathbb{R}^p$ will be denoted by $x \cdot y$. So, if $X = \begin{bmatrix} x \\ \alpha \end{bmatrix}$ and $Y = \begin{bmatrix} y \\ \beta \end{bmatrix}$, one has $X \star \eta Y = x \cdot y - \alpha \beta$.

It is convenient to write a matrix $\Lambda \in O(p,1)$ in the block form

$$\Lambda = \begin{pmatrix} A & b \\ c^T & \alpha \end{pmatrix},$$

where $A$ is a $p \times p$ matrix, $b$ is a $p \times 1$ vector, $c$ is a $1 \times p$ vector, and $\alpha \in \mathbb{R}$.

From now on, we restrict our attention to the (generalized) Lorentz group. In order to make notation less cumbersome, the matrix $\eta_{p,1}$ will be indicated simply by $\eta$; vectors in $\mathbb{R}^{p+1}$ will be denoted by $U$, $V$, $X$, $Y$, ..., and they will be thought as column vectors; so, e.g., we shall write $X = (X_1, \ldots, X_{p+1})^T$. We fix the immersion $\mathbb{R}^p \to \mathbb{R}^{p+1}$ defined by $(x_1, \ldots, x_p) \to (x_1, \ldots, x_p, 0)$; column vectors in $\mathbb{R}^p$ will be denoted by $b$, $c$, ..., $u$, $v$, $x$, $y$, ...; hence, a vector $X$ can be written also in the block form $X = \begin{bmatrix} x \\ \alpha \end{bmatrix}$ with $\alpha \in \mathbb{R}$. The Lorentz group $O(p,1)$ is the isometry group of the Minkowski product (defined by $\eta$) on $\mathbb{R}^{p+1}$, namely the pseudo-Euclidean inner product defined by

$$X \star \eta Y = \sum_{i,j=1}^{p+1} \eta_{ij} X_i Y_j = \sum_{i=1}^p X_i Y_i - X_{p+1} Y_{p+1}.$$ 

The usual Euclidean inner product between two vectors $x$, $y$ in $\mathbb{R}^p$ will be denoted by $x \cdot y$. So, if $X = \begin{bmatrix} x \\ \alpha \end{bmatrix}$ and $Y = \begin{bmatrix} y \\ \beta \end{bmatrix}$, one has $X \star \eta Y = x \cdot y - \alpha \beta$.

It is convenient to write a matrix $\Lambda \in O(p,1)$ in the block form

$$\Lambda = \begin{pmatrix} A & b \\ c^T & \alpha \end{pmatrix},$$

where $A$ is a $p \times p$ matrix, $b$ is a $p \times 1$ vector, $c$ is a $1 \times p$ vector, and $\alpha \in \mathbb{R}$.
where \( A \in GL(p) \), \( b \) and \( c \) are (column) vectors in \( \mathbb{R}^p \), and \( \sigma \in \mathbb{R} \). Notice that the matrices \( \eta \) and \( \xi = \left( \begin{array}{c|c} -I_p & 0 \\ \hline 0 & 1 \end{array} \right) \) are elements of \( O(p, 1) \); so, if a matrix \( \Lambda \) as above lies in \( O(p, 1) \), then the matrices
\[
\eta \Lambda = \left( \begin{array}{c|c} A & b \\ \hline -c^T & -\sigma \end{array} \right), \quad \Lambda \eta = \left( \begin{array}{c|c} A & -b \\ \hline c^T & -\sigma \end{array} \right),
\]
\[
\xi \Lambda = \left( \begin{array}{c|c} -A & -b \\ \hline c^T & \sigma \end{array} \right), \quad \Lambda \xi = \left( \begin{array}{c|c} -A & b \\ \hline -c^T & \sigma \end{array} \right)
\]
lie in \( O(p, 1) \) as well. This means that, in our subsequent reasonings, we shall be free to make the substitutions \((b, \sigma) \rightarrow (-b, -\sigma), (c, \sigma) \rightarrow (-c, -\sigma), (A, b) \rightarrow (-A, -b)\), or \((A, c) \rightarrow (-A, -c)\).

The inversion formula (1.1) yields some useful relations between \( A, b, c, \sigma \). One has
\[
\Lambda^T = \left( \begin{array}{c|c} A^T & c \\ \hline b^T & \sigma \end{array} \right)
\]
and
\[
\Lambda^{-1} = \eta \Lambda^T \eta = \left( \begin{array}{c|c} A^T & -c \\ \hline -b^T & \sigma \end{array} \right).
\]
So, since \( \Lambda^{-1} = \left( \begin{array}{c|c} I_p & 0 \\ \hline 0 & 1 \end{array} \right) \), we get
\[
AA^T - bb^T = I_p; \quad (2.1)
\]
\[
-Ac + \sigma b = 0; \quad (2.2)
\]
\[
-c^Tc + \sigma^2 = 1. \quad (2.3)
\]
Notice that, if \( b = (b_1, \ldots, b_p) \) and \( c = (c_1, \ldots, c_p) \), then \( c^Tc \) is the scalar \( c \cdot c = \sum_{i=1}^p c_i^2 \) and \( bb^T \) is the \((p \times p)\)-matrix whose \((i, j)\) entry is \( b_i b_j \). Interchanging the role of \( \Lambda \) and \( \Lambda^{-1} \) (i.e. making the substitutions \( A \rightarrow A^T \) and \( b \rightarrow -c \)), we obtain the relations
\[
A^T A - cc^T = I_p; \quad (2.4)
\]
\[
-A^T b + \sigma c = 0; \quad (2.5)
\]
\[
-b^T b + \sigma^2 = 1. \quad (2.6)
\]

It comes as no surprise that equations (2.1) -- (2.6) are invariant under the substitutions \((b, \sigma) \rightarrow (-b, -\sigma), (c, \sigma) \rightarrow (-c, -\sigma), (A, b) \rightarrow (-A, -b), (A, c) \rightarrow (-A, -c)\).

Equation (2.3) — or, equivalently, equation (2.6) — entails that \( \sigma^2 \geq 1 \). It is an easy matter to prove the following elementary, though very useful result.

**Lemma 2.1.** The following conditions are equivalent:
1) \( \sigma^2 = 1 \);
2) \( b = 0 \);
3) \( c = 0 \);
4) \( A \in O(p) \).
Proof. ▲

We now wish to apply the polar decomposition theorem for the real general linear group \([2, \text{Theorem 8.8}]\) and express any Lorentz matrix as the product of an orthogonal matrix and a symmetric positive definite matrix. To this aim, we slavishly follow the proof of that theorem: to get the positive definite matrix associated with \(\Lambda \in O(p, 1)\) we first compute the product \(\Lambda^T \Lambda\) and then take its square root.

By direct computation — making use of equations (2.4), (2.5), (2.6), and (2.3) — one gets

\[
\Lambda^T \Lambda = \begin{pmatrix} A^T & c \\ e^T & \sigma \end{pmatrix} \begin{pmatrix} A & b \\ c & \sigma \end{pmatrix} = \begin{pmatrix} A^T A + cc^T & A^T b + \sigma c \\ b^T A + \sigma e^T & b^T b + \sigma^2 \end{pmatrix} =
\begin{pmatrix} I_p + 2cc^T & 2\sigma c \\ 2\sigma e^T & e^T e + \sigma^2 \end{pmatrix}.
\]

(2.7)

The case \(\sigma^2 = 1\) is trivial. Let us assume that \(\sigma^2 > 1\); one has the following helpful identities:

\[
(I_p + \frac{cc^T}{1 + \sigma}) (I_p + \frac{cc^T}{1 + \sigma}) = I_p + 2\frac{cc^T}{1 + \sigma} + \frac{c(e^T e)}{(1 + \sigma)^2} = I_p + cc^T; \tag{2.8}
\]

\[
(I_p + \frac{cc^T}{1 + \sigma}) c = c + \frac{c(e^T c)}{1 + \sigma} = c + \frac{c(\sigma^2 - 1)}{1 + \sigma} = \sigma c. \tag{2.9}
\]

Of course, equations (2.8) and (2.9) are invariant under the substitution \((c, \sigma) \rightarrow (-c, -\sigma)\).

The square root of \(\Lambda^T \Lambda\) has to be a positive definite symmetric matrix: in particular, its \((p, p)\) entry must be positive. Thus, we have to distinguish two cases:

if \(\sigma > 1\), \(\sqrt{\Lambda^T \Lambda} = \begin{pmatrix} I_p + \frac{cc^T}{1 + \sigma} & c \\ e^T & \sigma \end{pmatrix}\),

(2.10)

if \(\sigma < -1\), \(\sqrt{\Lambda^T \Lambda} = \begin{pmatrix} I_p + \frac{cc^T}{1 + \sigma} & -c \\ -e^T & -\sigma \end{pmatrix}\).

(2.11)

Lemma 2.2. Let \(\Lambda = \begin{pmatrix} A & b \\ e^T & \sigma \end{pmatrix}\) be any matrix in \(O(p, 1)\). Let \(e = \frac{\sigma}{|\sigma|}\) be the sign of \(\sigma\). Then, the matrix

\[
S_\Lambda = \begin{pmatrix} I_p + \frac{cc^T}{1 + \sigma} & \frac{cc^T}{1 + \sigma} \\ e^T e & e^T \sigma \end{pmatrix} = \begin{pmatrix} \sqrt{I_p + \frac{cc^T}{1 + \sigma}} & \frac{cc^T}{1 + \sigma} \\ e^T e & e^T \sigma \end{pmatrix}
\]

(2.12)

is positive definite and symmetric and \(S_\Lambda^2 = \Lambda^T \Lambda\). If \(\sigma^2 > 1\), an orthonormal basis of eigenvectors for the matrix (2.12) is given by the vectors \(\begin{pmatrix} v_1 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} v_{p-1} \\ 0 \end{pmatrix}\), \(U_1, U_2\), where \(\{v_1, \ldots, v_{p-1}\}\) is an orthonormal basis for the space \(c^1 = \{w \in \mathbb{R}^p \mid c \cdot w = 0\} \subset \mathbb{R}^p\) and

\[
U_1 = \frac{1}{2} \begin{pmatrix} c \\ \sqrt{\sigma^2 - 1} \end{pmatrix}, \quad U_2 = \frac{1}{2} \begin{pmatrix} c \\ -\sqrt{\sigma^2 - 1} \end{pmatrix}.
\]
The corresponding eigenvalues are 1 with multiplicity \( p - 1 \) and \( \epsilon \left( \sigma \pm \sqrt{\sigma^2 - 1} \right) \).

**Proof.** By Lemma 2.1, if \( \sigma^2 = 1 \), then \( \epsilon = 0 \); so, in this case, \( S_\Lambda = I_{p+1} \), which is symmetric and positive definite. Assume now that \( \sigma^2 > 1 \); therefore, \( \epsilon \neq 0 \). It is obvious that, for any vector \( \mathbf{v} \) such that \( \epsilon \mathbf{v} = \mathbf{v}^T \mathbf{v} = 0 \), one has

\[
\begin{pmatrix}
I_p + \frac{\epsilon \mathbf{e}^T}{1 + \epsilon \sigma} & \epsilon \mathbf{e} \\
\epsilon \mathbf{e} & \epsilon \sigma
\end{pmatrix}
\begin{pmatrix}
\mathbf{v} \\
0
\end{pmatrix}
=
\begin{pmatrix}
\mathbf{v} \\
0
\end{pmatrix},
\]

so that each \( \mathbf{v} \) is an eigenvector of eigenvalue 1. On the other hand, taking in mind the identities (2.10) and (2.11), we get

\[
\begin{pmatrix}
I_p + \frac{\epsilon \mathbf{e}^T}{1 + \epsilon \sigma} & \epsilon \mathbf{e} \\
\epsilon \mathbf{e} & \epsilon \sigma
\end{pmatrix}
\begin{pmatrix}
\frac{\mathbf{e}}{\sqrt{\sigma^2 - 1}} \\
\frac{1 + \epsilon \sigma}{\sqrt{\sigma^2 - 1}}
\end{pmatrix}
=
\begin{pmatrix}
\frac{\mathbf{e}}{\sqrt{\sigma^2 - 1}} + \frac{\epsilon \mathbf{e}^T \mathbf{e}}{\epsilon \sigma} \\
\frac{1 + \epsilon \sigma}{\sqrt{\sigma^2 - 1}} + \epsilon \sigma
\end{pmatrix}
=
\begin{pmatrix}
\frac{\mathbf{e} + \epsilon \sqrt{\sigma^2 - 1}}{\epsilon \sqrt{\sigma^2 - 1}} \\
\frac{1 + \epsilon \sigma}{\epsilon \sqrt{\sigma^2 - 1}}
\end{pmatrix}
=
\begin{pmatrix}
\frac{\mathbf{e} + \epsilon \sqrt{\sigma^2 - 1}}{\epsilon \sqrt{\sigma^2 - 1}} \\
\frac{1 + \epsilon \sigma}{\epsilon \sqrt{\sigma^2 - 1}}
\end{pmatrix},
\]

It follows that \( \mathbf{U}_1 \) is an eigenvector for \( S_\Lambda \), having eigenvector is \( \epsilon \left( \sigma + \sqrt{\sigma^2 - 1} \right) \). An analogous computation can be carried out for \( \mathbf{U}_2 \), showing that the corresponding eigenvalue is \( \epsilon \left( \sigma - \sqrt{\sigma^2 - 1} \right) \). Notice that, since \( \sqrt{\sigma^2 - 1} < |\sigma| \), all eigenvalues are positive, so that \( S_\Lambda \) is positive definite. Equations (2.8) and (2.9) show (1) that \( S_\Lambda^2 = \Lambda^T \Lambda \).

The orthogonal matrix \( R_\Lambda \) associated with \( \Lambda \) can now be readily obtained (cf. [2, Theorem 8.8]). If we let \( R_\Lambda = \Lambda S_\Lambda^{-1} \), then

\[
R_\Lambda^T R_\Lambda = S_\Lambda^{-1} \Lambda^T \Lambda S_\Lambda^{-1} = S_\Lambda^{-1} S_\Lambda S_\Lambda^{-1} = I_{p+1}.
\]

To get an explicit expression for \( R_\Lambda \), it more convenient to compute \( R^T = R^{-1} = \)
\[ S_{\Lambda} \Lambda^{-1} : \]

\[
R^T = \begin{pmatrix}
I_p + \frac{cc^T}{1+\varepsilon\sigma} & cc \\
cc^T & \varepsilon\sigma
\end{pmatrix}
\begin{pmatrix}
A^T & -c \\
b^T & \sigma
\end{pmatrix} = \\
\begin{pmatrix}
A^T + \frac{cc^T A^T}{1+\varepsilon\sigma} - cc b^T \\
cc^T A^T - \varepsilon\sigma b^T
\end{pmatrix} = \\
\begin{pmatrix}
A^T + \frac{\sigma cb^T}{1+\varepsilon\sigma} - cc b^T \\
\varepsilon\sigma b^T - \varepsilon\sigma b^T
\end{pmatrix} = \\
\begin{pmatrix}
A^T - \frac{cc b^T}{1+\varepsilon\sigma} & 0 \\
0 & \varepsilon
\end{pmatrix}.
\] (2.13)

Putting all together, we have reached the following fundamental result.

**Theorem 2.3 (Polar decomposition theorem).** Any \( \Lambda = \begin{pmatrix} A & b \\ c^T & \sigma \end{pmatrix} \) in the Lorentz group \( O(p,1) \) can be uniquely factored as \( \Lambda = R \Lambda S \Lambda \), where \( \varepsilon = \frac{\sigma}{|\sigma|} \), \( R \Lambda = \begin{pmatrix} A - \frac{cb c^T}{1+\varepsilon\sigma} & 0 \\ 0 & \varepsilon \end{pmatrix} \) is a matrix in \( O(p+1) \), \( \varepsilon = \frac{\sigma}{|\sigma|} \), and \( S \Lambda = \begin{pmatrix} \sqrt{I_p + cc^T} & \varepsilon c \\ cc^T & \varepsilon \sigma \end{pmatrix} \) is a positive definite symmetric matrix.

\[ \text{Proof.} \] The matrix \( Q \Lambda = A - \frac{cb c^T}{1+\varepsilon\sigma} \) lies in \( O(p) \), as it can be easily checked by direct computation using equations (2.1) – (2.6). We have the following identity:

\[
Q c = Ac - \frac{cb c^T}{1+\varepsilon\sigma} = \sigma b - \frac{cb (\sigma^2 - 1)}{1+\varepsilon\sigma} = \varepsilon b. 
\] (2.14)

For any vector \( u \) in \( \mathbb{R}^p \), let us define the scalar \( \gamma(u) = \sqrt{u^T u + 1} = \sqrt{u \cdot u + 1} \geq 1 \) and the matrix

\[
\Lambda_u = \begin{pmatrix}
\sqrt{I_p + uu^T} & u \\
\frac{uu^T}{u^T} & \gamma(u)
\end{pmatrix} = \begin{pmatrix}
I_p + \frac{uu^T}{u^T} & u \\
\frac{uu^T}{u^T} & \gamma(u)
\end{pmatrix}.
\] (2.15)

Clearly, any matrix \( S \Lambda \) appearing in Theorem 2.3 corresponds to one and only one \( \Lambda_u \), with \( u = cc \) (indeed, \( cc^T (cc) = c^T c \) and \( \varepsilon \sigma = \gamma(u) \)).
Corollary 2.4. The mapping $u \mapsto \Lambda_u$ yields an embedding $\mathbb{R}^p \to O(p+1)$. It follows that the group $SO^+(p,1)$ is not compact. ▲

Lemma 2.2 tells us that, whenever $u \neq 0$, a basis of orthonormal vectors for $\Lambda_u$ is provided by the vectors $\{v_1, \ldots, v_{p-1}\}$, $V_1$, $V_2$, where $\{v_1, \ldots, v_{p-1}\}$ is an orthonormal basis for the space $\mathbf{u} \perp \subset \mathbb{R}^p$ and

$$V_1 = \frac{1}{2} \left( \begin{array}{c} u \\ \sqrt{\gamma(u)^2 - 1} \end{array} \right), \quad V_2 = \frac{1}{2} \left( \begin{array}{c} u \\ \sqrt{\gamma(u)^2 - 1 - 1} \end{array} \right).$$

The corresponding eigenvalues are 1 with multiplicity $p-1$ and $\gamma(u) \pm \sqrt{\gamma(u)^2 - 1}$. As $\gamma(u) \geq 1$, there is a unique real number $a(u) \geq 0$ such that

$$\cosh a(u) = \gamma(u).$$

So, we have that

$$\sqrt{\gamma(u)^2 - 1} = \sinh a(u),$$
$$\gamma(u) + \sqrt{\gamma(u)^2 - 1} = \cosh a(u) + \sinh a(u) = \exp a(u),$$
$$\gamma(u) - \sqrt{\gamma(u)^2 - 1} = \cosh a(u) - \sinh a(u) = \exp(-a(u)).$$

In conclusion, with respect to the given basis of eigenvectors, the matrix $\Lambda_u$ takes the diagonal form

$$\begin{pmatrix} I_{p-1} & 0 & 0 \\ 0 & \exp a(u) & 0 \\ 0 & 0 & \exp(-a(u)) \end{pmatrix}.$$  

3 An interlude: the case of $O(1,1)$

The group $O(1,1)$, being one-dimensional, is susceptible of an extremely simple description. Nevertheless, the study of this toy example can be instructive to understand the topological structure of $O(p,1)$ with $p > 1$.

Theorem 2.3 tells us that any matrix $\Lambda = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ can be factored as the product $\Lambda = R_\Lambda S_\Lambda$, where

$$R_\Lambda = \begin{pmatrix} a & \frac{bc}{1 + \sigma} & 0 \\ \frac{1 + \sigma}{1 + \sigma} & 0 \\ 0 & 0 & c \end{pmatrix},$$
$$S_\Lambda = \begin{pmatrix} \sqrt{1 + \sigma^2} & cc \\ \sigma & cc \end{pmatrix}.$$
On the other hand, since $\cosh a \epsilon \sigma$ is isomorphic to $\cosh a$, according to the sign of $\sigma$. By letting $a(u) = \arcsinh u$, one has $\gamma(u) = \cosh a(u)$, so that any $\Lambda_u$ can be written in the form
\[ \Lambda_u = \begin{pmatrix} \cosh a(u) & \sinh a(u) \\ \sinh a(u) & \cosh a(u) \end{pmatrix}. \]

On the other hand, since $a = \frac{ebc}{1 + e^2}$ must be an element of $O(1) = \{1, -1\}$, we see that $R_A$ can only be one of the following 4 matrices:
\[ I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad -\eta = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \eta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad -I = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}. \]

More precisely, by using the identities (2.1) – (2.6), it is easy to enumerate all possible cases:
\[ \begin{cases} 
\text{if } \det \Lambda = 1 \text{ and } \sigma \geq 1, \text{ then } R_A = I; \\
\text{if } \det \Lambda = 1 \text{ and } \sigma \leq -1, \text{ then } R_A = -I; \\
\text{if } \det \Lambda = -1 \text{ and } \sigma \geq 1, \text{ then } R_A = -\eta; \\
\text{if } \det \Lambda = -1 \text{ and } \sigma \leq -1, \text{ then } R_A = \eta. 
\end{cases} \tag{3.1} \]

The condition $\det \Lambda = 1$ is equivalent to either of the following conditions: (i) $a = \sigma$; (ii) $b = c (\checkmark)$.

**Lemma 3.1.** The space $\Sigma(1, 1) = \{\Lambda_u \mid u \in \mathbb{R}\}$ is a (closed) Lie subgroup of $O(1, 1)$, which is isomorphic to $\mathbb{R}$. Thus, $\Sigma(1, 1)$ is the connected component of the identity in $O(1, 1)$.

**Proof.** Let $a_1 = a(u_1) = \arcsinh u_1$ and $a_2 = a(u_2) = \arcsinh u_2$; then, one has
\[
\Lambda_{a_1} \Lambda_{a_2} = \begin{vmatrix} \cosh a_1 & \sinh a_1 \\ \sinh a_1 & \cosh a_1 \end{vmatrix} \begin{vmatrix} \cosh a_2 & \sinh a_2 \\ \sinh a_2 & \cosh a_2 \end{vmatrix} = \\
\begin{vmatrix} \cosh a_1 \cosh a_2 + \sinh a_1 \sinh a_2 & \cosh a_1 \sinh a_2 + \sinh a_1 \cosh a_2 \\ \cosh a_1 \sinh a_2 + \sinh a_1 \cosh a_2 & \cosh a_1 \cosh a_2 + \sinh a_1 \sinh a_2 \end{vmatrix} = \\
\begin{vmatrix} \cosh(a_1 + a_2) & \sinh(a_1 + a_2) \\ \sinh(a_1 + a_2) & \cosh(a_1 + a_2) \end{vmatrix} = \Lambda_{\sinh(a_1 + a_2)}. 
\]

Therefore, $\Sigma(1, 1)$ is a Lie subgroup of $O(1, 1)$ and the map $\mathbb{R} \to \Sigma(1, 1)$ defined by $x \mapsto \Lambda_{\sinh x}$ is a Lie group isomorphism. Since $\mathbb{R}$ is connected and $I \in \Sigma(1, 1)$, we conclude that $\Sigma(1, 1)$ is the connected component of the identity in $O(1, 1)$. □

*Caveat lector:* when $p > 1$, the product of two matrices $\Lambda_{a_1}, \Lambda_{a_2}$ in $O(p, 1)$ may even fail to be symmetric!

We introduce the following notation: $\Sigma(1, 1) = SO^+(1, 1)$.

As a consequence of Lemma 3.1 and of the enumeration (3.1), we have that $SO^+(1, 1)$ is the subgroup of $O(1, 1)$ consisting of matrices having $\sigma \geq 1$ and determinant equal
to 1. Let us now take the disjoint union
\[ O^+(1, 1) = SO^+(1, 1) \cup -\eta \cdot SO^+(1, 1); \]
it is clear that \( O(1, 1) \) is the subgroup of \( O(1, 1) \) consisting of matrices having \( \sigma \geq 1 \).
Also the subgroup \( SO(1, 1) \), whose elements are the matrices \( \Lambda \) such that \( \det \Lambda = 1 \), has two connected components, because one has
\[ SO(1, 1) = SO^+(1, 1) \cup -I \cdot SO^+(1, 1) \] (disjoint union).
Obviously, \( SO^+(1, 1) = O^+(1, 1) \cap SO(1, 1) \).

In conclusion, we have proved the following result.

**Corollary 3.2.** The group \( O(1, 1) \) has 4 connected components, namely
\[ SO^+(1, 1) = \Sigma(1, 1), \ -I \cdot SO^+(1, 1), \ -\eta \cdot SO^+(1, 1), \ \eta \cdot SO^+(1, 1). \]

To get a clearer idea of the structure of the group \( O(1, 1) \), we can analyze its action on \( \mathbb{R}^2 \). Since \( O(1, 1) \) is the isometry group of the Minkowski inner product \( X \cdot M Y = X_1 Y_1 - X_2 Y_2 \), its orbits in the \((X_1, X_2)\) plane are contained in the level sets of the quadratic form \( X_1^2 - X_2^2 \). Consider the branch \( \Gamma^+ \) of the hyperbola \( X_1^2 - X_2^2 = -1 \) lying in the upper half-plane: the matrix
\[
\begin{pmatrix}
\cosh a & \sinh a \\
\sinh a & \cosh a
\end{pmatrix}
\]
takes the point \((0, 1)\) to the point \((\sinh a, \cosh a)\), which lies on \( \Gamma^+ (1) \) as well. The action of \( SO^+(1, 1) \) on \( \Gamma^+ \) is transitive (i.e. every point is reached) and without fixed point, so that it establishes a homeomorphism from \( SO^+(1, 1) \) onto \( \Gamma^+ (1) \) (of course, both spaces are homeomorphic to \( \mathbb{R} \)). Analogously, one shows that \( \Gamma^+ \) is homeomorphic to the subspace \( -\eta \cdot SO^+(1, 1) \), while the branch \( \Gamma^- \) of the hyperbola \( X_1^2 - X_2^2 = -1 \) lying in the lower half-plane turns out to be homeomorphic to both \( -I \cdot SO^+(1, 1) \) and \( \eta \cdot SO^+(1, 1) \). Summing up, the group \( O(1, 1) \) is a trivial double cover of the disjoint union \( \Gamma^+ \cup \Gamma^- \); in other words, it is homeomorphic to the space \((\Gamma^+ \cup \Gamma^-) \times O(1)\).

### 4 The topological structure of \( O(p, 1) \)

\( A \) vector \( X = \begin{pmatrix} x \\ \alpha \end{pmatrix} \) in the Minkowski space \( \mathbb{R}^{p+1} \) is said to be
- **space-like** if \( X \cdot M X > 0 \);
- **time-like** if \( X \cdot M X < 0 \);
- **light-like** (or null) if \( X \cdot M X = 0 \);
- **positive** (resp. **negative**) if \( \alpha > 0 \) (resp. \( \alpha < 0 \)).
The set of light-like vectors is a two-sheeted cone in \( \mathbb{R}^{p+1} \), which is called the light-cone. The sheet consisting of positive (resp. negative) light-like vectors is denoted by \( C^\ast (\text{resp. } C^-) \).

**Lemma 4.1.** If \( X \) and \( Y \) are both positive, or both negative time-like vectors, then \( X \ast Y < 0 \). Moreover, for any \( t \in [0, 1] \), the linear combination \( (1-t)X + tY \) is a time-like vector with the same parity as \( X \) and \( Y \).

**Proof.** Let us assume \( X = \begin{pmatrix} x \\ a \end{pmatrix} \) and \( Y = \begin{pmatrix} y \\ \beta \end{pmatrix} \) are two positive time-like vectors. Then, the inequality \( X \ast Y < 0 \) is equivalent to the inequality \( \|x\| < \alpha \); likewise, we have \( \|y\| < \beta \). Now, \( X \ast Y = x \ast y - \alpha \beta \). But \( x \ast y \leq \|x\|\|y\| < \alpha \beta \), so that \( X \ast Y < 0 \). For any \( t \in (0, 1) \), the vectors \( (1-t)X \) and \( tY \) lie in \( C^\ast \cup \{0\} \). By contrast, notice that, if \( p > 1 \), the space of space-like vectors is connected.

The case of two negative time-like vectors is readily handled in the same way. ▲

Lemma 4.1 shows that the sets

\[ T^+ = \{ X \in \mathbb{R}^{p+1} \mid X \text{ time-like and positive} \} \]

and \( T^- \) (analogously defined) are convex — in particular connected — topological subspaces of \( \mathbb{R}^{p+1} \); of course, \( T^+ \) and \( T^- \) are disjoint. The boundary of \( T^+ \) is the union of the positive light-cone \( C^\ast \) and the origin: \( \partial T^+ = C^\ast \cup \{0\} \). Likewise, we have \( \partial T^- = C^- \cup \{0\} \). By contrast, notice that, if \( p > 1 \), the space of space-like vectors is connected.

Every transformation \( \Lambda \in O(p, 1) \) preserves the likeness of any vector, but not its parity; in fact, \( \eta \) maps positive vectors to negative ones and conversely. Let us define the subgroup

\[ O^+(p, 1) = \{ \Lambda \in O(p, 1) \mid \Lambda T^+ \subseteq T^+ \}. \]

It is quite immediate to see that \( O^+(p, 1) \) has index 2 in \( O(p, 1) \); indeed, the action of \( O(p, 1) \) on the space of time-like vectors is continuous and this space has precisely two connected components, namely \( T^+ \) and \( T^- \).

**Lemma 4.2.** Let \( \Lambda = \begin{pmatrix} A & b \\ c & \sigma \end{pmatrix} \) be a transformation in \( O(p, 1) \); \( \Lambda \) lies in \( O^+(p, 1) \) if and only if \( \sigma > 0 \).

**Proof.** First notice that, by eqq. (2.3) and (2.6), the vectors \( \begin{pmatrix} c \\ \sigma \end{pmatrix} \) and \( \begin{pmatrix} b \\ \sigma \end{pmatrix} \) are time-like. Assume that \( \sigma > 0 \). Then, for any positive time-like vector \( X = \begin{pmatrix} x \\ a \end{pmatrix} \), we have \( \Lambda X = Y \), where \( Y = \begin{pmatrix} y \\ \beta \end{pmatrix} \) is time-like and \( \beta = c \ast x + \sigma a \). The vectors \( \begin{pmatrix} c \\ \sigma \end{pmatrix} \) and \( X \) are both positive and time-like, so that, by Lemma 4.1, their Minkowski product is positive and time-like as well: \( c \ast x - \sigma a < 0 \). Since \( \sigma a > 0 \), we have also \( |c \ast x| < \|c\|\|x\| < a\sigma \); hence, \( \beta > 0 \), and \( \Lambda \in O^+(p, 1) \). Conversely, assume that \( \Lambda \in O^+(p, 1) \). Then, for any positive
time-like vector \( \mathbf{X} \), the vector \( \Lambda \mathbf{X} = \mathbf{Y} \) is time-like and positive. If we take \( \mathbf{X} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \), we obtain \( \mathbf{Y} = \begin{pmatrix} b \\ \sigma \end{pmatrix} \); therefore, it must be \( \sigma > 0 \).

Let us now consider the subgroup \( SO(p, 1) \) of \( O(p, 1) \) consisting of Lorentz transformations having determinant equal to 1. This subgroup has index 2 in \( O(p, 1) \). Note that \( SO(p, 1) \) is not contained in \( O^+(p, 1) \): in fact, for any \( Q \in O(p) \) with \( \det Q = -1 \), the matrix \( \begin{pmatrix} Q & 0 \\ 0 & -1 \end{pmatrix} \) lies in \( SO(p, 1) \) but not in \( O^+(p, 1) \). It follows that the subgroup

\[
SO^+(p, 1) = O^+(p, 1) \cap SO(p, 1)
\]

has index 4 in \( O(p, 1) \). More precisely, if we introduce the matrix \( \chi = \begin{pmatrix} -1 & 0 \\ 0 & I \end{pmatrix} \), we see that \( O(p, 1) \) admits the following disjoint union decomposition:

\[
O(p, 1) = SO^+(p, 1) \cup \eta SO^+(p, 1) \cup \chi SO^+(p, 1) \cup (\eta \chi) SO^+(p, 1).
\]

By Theorem 2.3, any \( \Lambda = \begin{pmatrix} A & b \\ \epsilon & \sigma \end{pmatrix} \) in \( O(p, 1) \) can be expressed as the product

\[
\Lambda = \begin{pmatrix} Q_\Lambda & 0 \\ 0 & \epsilon \end{pmatrix} \Lambda_u,
\]

where \( \epsilon = \frac{\sigma}{|\sigma|} \), \( Q_\Lambda \in O(p) \) and \( \Lambda_u \) is given by eq. (2.15). It is easy to enumerate all possible cases:

\[
\begin{align*}
\Lambda & \in SO^+(p, 1) \text{ if and only if } Q_\Lambda \in SO(p) \text{ and } \epsilon = 1; \\
\Lambda & \in \eta SO^+(p, 1) \text{ if and only if } \det Q_\Lambda = -1 \text{ and } \epsilon = 1; \\
\Lambda & \in \eta SO^+(p, 1) \text{ if and only if } Q_\Lambda \in SO(p) \text{ and } \epsilon = -1; \\
\Lambda & \in (\eta \chi) SO^+(p, 1) \text{ if and only if } \det Q_\Lambda = -1 \text{ and } \epsilon = -1.
\end{align*}
\]

Using Lemma 4.2 it is immediate to obtain the following disjoint union decomposition

\[
O^+(p, 1) = SO^+(p, 1) \cup \chi SO^+(p, 1).
\]

In order to extend Corollary 3.2 to the general case of \( O(p, 1) \), we have only to show that \( SO^+(p, 1) \) is connected. We introduce the space

\[
H^+_p = \{ \mathbf{X} \in T^* | \mathbf{X} \cdot \mathbf{X} = -1 \},
\]

which is obviously diffeomorphic to \( \mathbb{R}^p \). A diffeomorphism (which we will regard as an identification) is provided by the projection \( \mathbf{X} = \begin{pmatrix} x \\ a \end{pmatrix} \in H^+_p \rightarrow x \in \mathbb{R}^p \), whose inverse is clearly given by the mapping \( x \mapsto \begin{pmatrix} x \\ \sqrt{1 + x \cdot x} \end{pmatrix} \).
Lemma 4.3. The group $SO^+(p, 1)$ acts transitively on $H^+_p$. The isotropy group $\Gamma_N$ of the vector $N = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is the subgroup of transformations $\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}$ with $Q \in SO(p)$.

Proof. Given any vector $X = \begin{pmatrix} x \\ \alpha \end{pmatrix} \in H^+_p$, one has $0 < \alpha = \sqrt{1 + x \cdot x} = \gamma(x)$. The matrix $\Lambda_x$ defined according to eq. (2.15), namely
\[
\Lambda_x = \begin{pmatrix} \sqrt{I_p + xx^T} & x \\ x^T & \gamma(x) \end{pmatrix},
\]
is an element of $SO^+(p, 1)$ and maps $N = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to $X$. The second claim is now a direct consequence of the polar decomposition theorem 2.3, taking into account the enumeration (4.1).

The isotropy group $\Gamma_X$ of any other vector $X$ is a subgroup of $SO^+(p, 1)$ conjugate to $\Gamma_N$. Precisely, if $X = \Lambda N$, then $\Gamma_X = \Lambda^{-1} \Gamma_N \Lambda$.

Theorem 4.4. There is a diffeomorphism
\[H^+_p = SO^+(p, 1)/SO(p).\]

As a consequence, the group $SO^+(p, 1)$ is connected (therefore, it is the connected component of the identity in $O(p, 1)$).

Proof. The first statement follows from [3, Chap. 1, Prop. 4.6] and from Lemma 4.3: since the action of $SO^+(p, 1)$ on $H^+_p$ is transitive and the isotropy group of any point can be identified with $SO(p)$, there is a diffeomorphism $H^+ = SO^+(p, 1)/SO(p)$. As for the second claim, let us consider the projection $\pi: SO^+(p, 1) \rightarrow H^+_p$. Now, $H^+$ is connected and, for any $X \in H^+_p$, the fibre $\pi^{-1}(X)$ is diffeomorphic to $SO(p)$ and therefore connected. Given any two points $\Lambda_1, \Lambda_2$ in $SO^+(p, 1)$, we can draw paths connecting $\Lambda_1, \Lambda_2$ to $\pi(\Lambda_1)$, $\pi(\Lambda_2)$ and a path connecting $\pi(\Lambda_1)$ to $\pi(\Lambda_2)$. Thus, $SO^+(p, 1)$ is connected.

Actually, it not hard to show that the $SO(p)$-principal bundle $\pi: SO^+(p, 1) \rightarrow H^+_p$ is trivial: $SO^+(p, 1) = H^+_p \times SO(p)$.

Corollary 4.5. The group $O(p, 1)$ has 4 connected components, which are explicitly described in enumeration (4.1).

Using Theorem 4.3 and the decomposition (4.2) we easily get the following result.

Corollary 4.6. The group $O^+(p, 1)$ acts transitively on $H^+_p$. The isotropy group of the vector $N$ is the subgroup of transformations $\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}$ with $Q \in O(p)$.
5 Hyperbolic geometry

In this Section we prove a result of fundamental importance, namely that the group $O^+(p, 1)$ is the isometry group of the $p$-dimensional hyperbolic space.

If $X$, $Y$ are two vectors in $H^+_p$, their Minkowski product, by Lemma 4.1, is a negative real number. So, there exists a unique non-negative real number $\mu(X, Y)$ such that

$$\cosh \mu(X, Y) = -X \cdot_0 Y.$$

Clearly, we have

$$\mu(X, X) = 0,$$
$$\mu(X, Y) > 0 \quad \text{for all } X \neq Y,$$
$$\mu(X, Y) = \mu(Y, X).$$

In order to show that the function $\mu : H^+_p \times H^+_p \to \mathbb{R}^\geq 0$ is a (topological) metric it remains only to prove that it satisfies the triangle inequality.

**Lemma 5.1.** For all vectors $X$, $Y$, $Z$ in $H^+_p$, the triangle inequality

$$\mu(X, Z) \leq \mu(X, Y) + \mu(Y, Z)$$

is satisfied.

**Proof.** Notice that, by definition, $\mu(X, Y) = \mu(\Lambda X, \Lambda Y)$ for any $\Lambda \in O^+(p, 1)$. Since $O^+(p, 1)$ acts transitively on $H^+_p$, there is no restriction in assuming that $Y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Let $X = \begin{pmatrix} x \\ \alpha \end{pmatrix}$ and $Z = \begin{pmatrix} z \\ \beta \end{pmatrix}$. We have $-X \cdot_0 Y = \alpha$ and $-Y \cdot_0 Z = \beta$; then, $\cosh \mu(X, Y) = \alpha$ and $\cosh \mu(Y, Z) = \beta$. Now, $\alpha^2 = \|x\|^2 + 1$, so that $\sinh \mu(X, Y) = \|x\|$; analogously, we get $\sinh \mu(Y, Z) = \|z\|$. One has:

$$-X \cdot_0 Z = -x \cdot z + \alpha \beta =$$
$$= -\|x\| \|z\| \cos \theta + \cosh \mu(X, Y) \cosh \mu(Y, Z) \leq$$
$$\leq \sinh \mu(X, Y) \sinh \mu(Y, Z) + \cosh \mu(X, Y) \cosh \mu(Y, Z) =$$
$$= \cosh (\mu(X, Y) + \mu(Y, Z)).$$

Now, $\cosh \mu(X, Z) = -X \cdot_0 Z$; since the hyperbolic cosine is an increasing function, we conclude that

$$\mu(X, Z) \leq \mu(X, Y) + \mu(Y, Z).$$

▲

The $p$-dimensional hyperbolic space is the metric topological space $\mathbb{H}_p = (H^+_p, \mu(\cdot, \cdot))$.

A map $\psi : \mathbb{H}_p \to \mathbb{H}_p$ is an isometry if it preserves distances, namely if

$$\mu(\psi(X), \psi(Y)) = \mu(X, Y) \quad \text{for all } X, Y \text{ in } \mathbb{H}_p.$$
It is easy to get convinced that any isometry of a metric space is continuous and injective; as it will be shown in Theorem 5.2, any isometry $\psi : \mathbb{H}_p \to \mathbb{H}_p$ happens also to be surjective (a property that, of course, is not satisfied in the general case).

**Theorem 5.2.** A map $\psi : \mathbb{H}_p \to \mathbb{H}_p$ is an isometry if and only if it is the restriction of a transformation $\Lambda \in O^+(p,1)$.

**Proof.** It is clear that map $\psi : \mathbb{H}_p \to \mathbb{H}_p$ is an isometry if and only if $\mu(\mathbf{X}) \cdot \psi(\mathbf{Y}) = \mathbf{X} \cdot \mathbf{Y}$ for all $\mathbf{X}, \mathbf{Y}$ in $\mathbb{H}_p$. Therefore, the restriction of any $\Lambda \in O^+(p,1)$ to $\mathbb{H}_p$ is an isometry. Conversely, let $\psi$ be an isometry, and let us suppose first that $\psi(\mathbf{N}) = \mathbf{N}$. Hence, for any $\mathbf{X} = \begin{pmatrix} \mathbf{x} \\ \alpha \end{pmatrix}$ we have:

$$\psi(\mathbf{X}) \cdot \alpha \mathbf{N} = \psi(\mathbf{X}) \cdot \alpha \psi(\mathbf{N}) = \mathbf{X} \cdot \alpha \mathbf{N} = -\alpha.$$

It follows that $\psi$ has the form $\psi(\mathbf{X}) = \begin{pmatrix} \psi(\mathbf{x}) \\ \alpha \end{pmatrix}$, where the map $\tilde{\psi} : \mathbb{R}^p \to \mathbb{R}^p$ preserves the Euclidean product, namely $\tilde{\psi}(\mathbf{x}) \cdot \tilde{\psi}(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^p$. Therefore, by a well-known basic result in Euclidean geometry, there is $Q \in O(p)$ such that $\tilde{\psi}(\mathbf{x}) = Q \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^p$. Thus, we get $\psi(\mathbf{X}) = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \mathbf{X}$ for all $\mathbf{X} \in \mathbb{H}_p$. Let us now suppose that $\psi(\mathbf{N}) \neq \mathbf{N}$. Since $SO^+(p,1)$ acts transitively on $\mathbb{H}_p$, there exists a transformation $\Lambda_1$ lying in that group such that $\Lambda_1(\psi(\mathbf{N})) = \mathbf{N}$. But then the map $\Lambda_1 \circ \psi : \mathbb{H}_p \to \mathbb{H}_p$ is an isometry fixing the point $\mathbf{N}$; thus, for all $\mathbf{X} \in \mathbb{H}_p$, one has $\Lambda_1 \circ \psi(\mathbf{X}) = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \mathbf{X}$, for some $Q \in O(p)$. Having in mind the decomposition (4.2), we conclude that $\psi(\mathbf{X}) = \Lambda_2 \mathbf{X}$, for some $\Lambda_2 \in O^+(p,1)$. ▲

Let us introduce the isometry group of $\mathbb{H}_p$:

$$\mathcal{I}(\mathbb{H}_p) = \{ \psi : \mathbb{H}_p \to \mathbb{H}_p \mid \psi \text{ isometry} \}.$$

It is possible to prove that $\mathcal{I}(\mathbb{H}_p)$ carries a natural structure of Lie group [5, Chap. 5]. So, Theorem 5.2 can be rephrased in the following way.

**Corollary 5.3.** There is a Lie group isomorphism $\mathcal{I}(\mathbb{H}_p) \cong O^+(p,1)$. The subgroup $\mathcal{I}^+(\mathbb{H}_p)$ of orientation-preserving isometries is isomorphic to $SO^+(p,1)$. ▲

The space $\mathbb{H}_p$, as a manifold, can naturally be endowed — by restricting the Lorentzian metric on $\mathbb{R}^{p+1}$ — with a Riemannian metric $h$ (therefore, for any point $\mathbf{X} \in \mathbb{H}_p$, $h_\mathbf{X}$ is a positive-definite inner product on the tangent space $T_\mathbf{X} \mathbb{H}_p$). According to a general procedure valid for all Riemannian metrics, one can associate to $h$ a topological distance function $\mu_h$ such that the distance $\mu_h(\mathbf{X}, \mathbf{Y})$ between any two (sufficiently close) points is equal to the arc length (measured by using $h$) of the geodesic segment joining $\mathbf{X}$ and $\mathbf{Y}$ (see e.g. [4, §1.4]). It turns out that the distance function $\mu_h$ coincides with the distance function $\mu$ we have previously defined. As a consequence, the (topological) isometry group $\mathcal{I}(\mathbb{H}_p)$ coincides with the (differentiable) isometry group of the Riemannian manifold $(\mathbb{H}_p, h)$.

To get a more perspicuous picture of the hyperbolic space $\mathbb{H}_p$, one can construct
its Poincaré's conformal model. Let $B^p$ be the open $p$-dimensional ball in the Euclidean space $\mathbb{R}^{p+1}$. The map:

$$\zeta : \mathbb{H}_p \to B^p$$

$$X = \begin{pmatrix} x \\ a \end{pmatrix} \mapsto \frac{x}{1 + a}$$

is a diffeomorphism. Actually, an easy computation shows that

$$\left\| \frac{x}{1 + a} \right\|^2 = \frac{a - 1}{a + 1} < 1$$

and that the inverse of $\zeta$ is the map

$$\zeta^{-1} : B^p \to \mathbb{H}_p$$

$$v \mapsto \frac{1}{1 - \|v\|^2} \left( \frac{2v}{1 + \|v\|^2} \right). \quad (5.1)$$

Let us embed $B^p$ into $\mathbb{R}^{p+1}$ by mapping the point $v$ to $(v, 0)$. Then, the map $\zeta^{-1} : B^p \to \mathbb{H}_p$ is the stereographic projection of $B^p$ onto $\mathbb{H}_p$ from the point $(0, -1)$, as shown in Figure 1.

We can obviously induce a topological metric on $B^p$ by setting

$$\delta(v, w) = \mu(\zeta^{-1}(v), \zeta^{-1}(w)) \quad \text{for any pair } (v, w) \in B^p.$$ 

It is quite remarkable that the distance $\delta(v, w)$ can be expressed in terms of Euclidean invariants.

**Lemma 5.4.** For any pair $(v, w)$ of points in $B^p$, one has

$$\cosh(\delta(v, w)) = 1 + 2 \frac{\|v - w\|^2}{(1 - \|v\|^2)(1 - \|w\|^2)}. \quad (5.2)$$
Proof. Let $v = \zeta(X)$ and $w = \zeta(Y)$, with $X = \begin{pmatrix} x \\ \alpha \end{pmatrix}$ and $Y = \begin{pmatrix} y \\ \beta \end{pmatrix}$. By definition we have
\[
\cosh(\delta(v, w)) = \cosh(\mu(X, Y)) = -X \cdot M \cdot Y.
\]
The following identity is easily obtained from the expression of $v$ in terms of $X$:
\[
1 - \|v\|^2 = \frac{1}{1+\alpha};
\]
similarly, one has $1 - \|w\|^2 = \frac{2}{1+\beta}$. Then, by direct computation we get:
\[
1 + \frac{2}{1+\alpha} (1 + \beta) \left( \frac{a^2 - 1}{(1+\alpha)^2} + \frac{\beta^2 - 1}{(1+\beta)^2} - \frac{2x \cdot y}{(1+\alpha)(1+\beta)} \right) =
\]
\[
= 1 + \frac{1}{2} (-2 + 2\alpha\beta - 2x \cdot y) = -X \cdot M \cdot Y.
\]

The isometry group of the space $(B^p, \delta(\cdot, \cdot))$ is called the $p$-dimensional Möbius group: this group is obviously isomorphic to the Lorentz group $O(p,1)$ and provides an extremely valuable tool to investigate its structure. We refer the reader to Ratcliffe’s book [5] for a thorough treatment of the general case. Here we limit ourselves, in the next Section, to analyse the 2-dimensional case, which exhibits some special features.

6 The Poincaré disc

The open ball $B^2$ is diffeomorphic to open disc $D$ in the complex plane,
\[
D = \{ z \in \mathbb{C} \mid \|z\| < 1 \}.
\]

As usual, a vector $x = (x_1, x_2)$ is identified with the complex number $z = x_1 + ix_2$; under this identification the Euclidean norm on $B^2$ is equal to the Hermitian norm on $D$, namely $\|x\| = \sqrt{x_1^2 + x_2^2} = \|z\| = \sqrt{zz^*}$. Therefore, we can define the hyperbolic metric on $D$ given by eq. (5.2) in terms of complex quantities only:
\[
\delta(z_1, z_2) = \text{arcosh} \left( 1 + \frac{\|z_1 - z_2\|^2}{(1 - \|z_1\|^2)(1 - \|z_2\|^2)} \right)
\]
for any pair of points $z_1$, $z_2$ in $D$. The disc $D$ endowed with such a metric is called the Poincaré disc.

The group of homeomorphisms of the extended complex plane $\hat{\mathbb{C}} = \mathbb{C} \cup \infty$ taking circles to circles is the group of Möbius transformations, which is generated by the linear fractional transformations
\[
z \mapsto \frac{az + b}{cz + d} \quad a, b, c, d \in \mathbb{C}, \quad ad - bc = 1
\]
The structure of the Lorentz group

The group of Möbius transformations of the form (6.2) will be denoted by \( \mathcal{M}^+ (\mathbb{D}) \). These transformations preserve the natural orientation of the disk.

Lemma 6.1. The Möbius group \( \mathcal{M}^+ (\mathbb{D}) \) is isomorphic to the projective special linear group \( \text{PSL}(2; \mathbb{R}) \).

Proof. Any transformation \( m \) in \( \mathcal{M}^+ (\mathbb{D}) \) of the form \( z \mapsto \frac{az + b}{bz + \bar{a}} \) can be mapped to the matrix \( A_m = \begin{pmatrix} a & b \\ b & \bar{a} \end{pmatrix} \), which lies in \( \text{SL}(2; \mathbb{C}) \) because \( \|a\|^2 - \|b\|^2 = 1 \). Clearly, this correspondence is well-defined only up to the sign of \( A_m \); thus \( A_m \) has to be thought of as an element of group \( \text{PSL}(2; \mathbb{C}) \). In this way, since the composition of two transformations \( m, m' \) is associated with the product matrix \( A_m A_{m'} \), we get an isomorphism \( \mathcal{M}^+ (\mathbb{D}) \to \Gamma \), where \( \Gamma \) is a subgroup of \( \text{PSL}(2; \mathbb{C}) \). Let us now consider the matrix

\[
P = \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}.
\]

The product

\[
P^{-1} A_m P = \frac{1}{2} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} a & b \\ b & \bar{a} \end{pmatrix} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} = \hat{A}_m = \begin{pmatrix} Ra + 3b & \Re b + 3a \\ \Re b - 3a & Ra - 3b \end{pmatrix}
\]

is a matrix with real entries whose determinant is

\[
(\Re a)^2 - (3b)^2 - (\Re b)^2 + (3a)^2 = \|a\|^2 - \|b\|^2 = 1;
\]

This shows that the group \( \Gamma \) is conjugate (in \( \text{PGL}(2; \mathbb{C}) \)) to the group \( \text{PSL}(2; \mathbb{R}) \). As a consequence, we get an isomorphism \( \mathcal{M}^+ (\mathbb{D}) \to \text{SL}(2; \mathbb{R}) \) explicitly given by the assignment \( m \mapsto \hat{A}_m \). ▲

The matrix \( P = \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \) is associated with the Möbius transformation

\[
z \mapsto \frac{z - i}{-iz + 1},
\]

which provides a holomorphic diffeomorphism of the upper-half complex plane \( \mathbb{U} \) onto the disc \( \mathbb{D} \). Thus, Lemma 6.1 implies that \( \mathcal{M}^+ (\mathbb{U}) = \text{PSL}(2; \mathbb{R}) \). We refer the reader to [1] and [5] for further details.
A Möbius transformation \( z \mapsto \frac{az + b}{bz + \bar{a}} \) in \( \mathcal{M}^+(\mathbb{D}) \) can always be factorized into the composition
\[
z \mapsto e^{i\theta} \frac{z + c}{\bar{c}z + 1},
\]
where \( e^{i\theta} = \frac{a}{\bar{a}} \) and \( c = \frac{b}{\bar{a}} \). As \( \|a\|^2 - \|b\|^2 = 1 \), one has \( \|c\| < 1 \). A transformation of type
\[
z \mapsto e^{i\theta} z
\]
acts as a rotation in the complex plane and hence it preserves both the Euclidean distance and the hyperbolic metric (6.1) on \( \mathbb{D} \). On the other hand, the transformations of type
\[
z \mapsto \frac{z + c}{\bar{c}z + 1}
\]
do not preserve, in general, the Euclidean distance but they do preserve the hyperbolic metric (6.1). Though this fact can be shown through a (somewhat tedious) direct computation, we will follow a different route, which has the advantage of bringing to light the existence of an isomorphism of the groups \( \mathcal{M}^+(\mathbb{D}) \) and \( SO^+(2, 1) \) (see Theorem 6.4).

We have introduced earlier an isometry \( \zeta : \mathbb{H}_2 \rightarrow B^2 \), which maps the vector \( X = \begin{pmatrix} x \\ \alpha \end{pmatrix} \) to the point \( \frac{x}{1 + \alpha} \); recall that \( \alpha = \sqrt{\|x\|^2 + 1} > 0 \). By identifying \( B^2 \) and \( \mathbb{D} \) as above, the map \( \zeta \) takes the form
\[
\zeta : \mathbb{H}_2 \rightarrow \mathbb{D}
\]
\[
X = \begin{pmatrix} x_1 \\ x_2 \\ \alpha \end{pmatrix} \mapsto \zeta(X) = \frac{1}{1 + \alpha} (x_1 + i x_2).
\]
It is immediate to deduce the following identity:
\[
\alpha = \frac{1 + \|\zeta(X)\|^2}{1 - \|\zeta(X)\|^2}.
\]
Therefore, by eq. (5.1), the inverse map \( \zeta^{-1} : \mathbb{D} \rightarrow \mathbb{H} \) is given by the assignment
\[
z \mapsto \zeta^{-1}(z) = \begin{pmatrix} \psi(z) \\ \tau(z) \end{pmatrix},
\]
where
\[
\psi(z) = \frac{1}{1 - \|z\|^2} \begin{pmatrix} 2\Re z \\ 2\Im z \end{pmatrix}, \quad \tau(z) = \frac{1 + \|z\|^2}{1 - \|z\|^2}.
\]
Note that the map \( \psi : \mathbb{D} \rightarrow \mathbb{R}^2 \) is one-to-one.

A question that naturally arises at this point is which kind of transformations are induced on \( \mathbb{H}_2 \) by the Möbius transformations on \( \mathbb{D} \) through the map \( \zeta^{-1} \). The answer provided by the next two results could hardly be simpler: they are just Lorentz transformations.
Lemma 6.2. Let \( m_\theta \) be the Möbius transformation given by
\[
    z \mapsto m_\theta(z) = e^{i\theta} z.
\]
Then, for any \( z \in \mathbb{D} \), one has
\[
    \zeta^{-1}(m_\theta(z)) = R_\theta \zeta^{-1}(z),
\]
where
\[
    R_\theta = \begin{pmatrix}
        \cos \theta & -\sin \theta & 0 \\
        \sin \theta & \cos \theta & 0 \\
        0 & 0 & 1
    \end{pmatrix}.
\]

Proof. It is a straightforward computation (\( \Box \)).

Theorem 6.3. Let \( m_c \) be the Möbius transformation given by
\[
    z \mapsto m_c(z) = \frac{z + c}{\bar{c}z + 1},
\]
with \( c \in \mathbb{D} \). Then, for any \( z \in \mathbb{D} \), one has
\[
    \zeta^{-1}(m_c(z)) = \Lambda_{\psi(c)}\zeta^{-1}(z).
\]

Proof. Let us first write down the Lorentz transformation \( \Lambda_{\psi(c)} \). For the sake of brevity, we write \( \psi(c) \) in the form \( \psi(c) = \frac{2}{1 - ||c||^2} c \), where \( c \) is the 2-dimensional vector corresponding to the complex number \( c \). An easy computation yields
\[
    \gamma(\psi(c)) = \sqrt{\psi(c) \cdot \psi(c) + 1} = \frac{1 + ||c||^2}{1 - ||c||^2}.
\]
Hence, by eq. (2.15), we have
\[
    \Lambda_{\psi(c)} = \begin{pmatrix}
        
        I + \frac{\psi(c)\psi(c)^T}{1 + \gamma(\psi(c))} & \psi(c) \\
        \psi(c)^T & \gamma(\psi(c))
    \end{pmatrix} = \begin{pmatrix}
        I + \frac{2}{1 - ||c||^2} c c^T & \frac{2}{1 - ||c||^2} c \\
        \frac{2}{1 - ||c||^2} c^T & 1 + ||c||^2
    \end{pmatrix}.
\]
Using the same convention as for \( \psi(c) \), we write \( \zeta^{-1}(z) \) in the form
\[
    \zeta^{-1}(z) = \frac{1}{1 - ||z||^2} \begin{pmatrix} 2z \\ 1 + ||z||^2 \end{pmatrix}.
\]
We set \( \Lambda_{\psi(c)}\zeta^{-1}(z) = \begin{pmatrix} y \\ \beta \end{pmatrix} \). It will be enough to compute the “spatial” component \( y \) of \( \Lambda_{\psi(c)}\zeta^{-1}(z) \) and then to check that is equal to to the “spatial” component \( \psi(m_c(z)) \) of \( \zeta^{-1}(m_c(z)) \); indeed, eqq. (6.6) and (6.7) ensure that the “temporal” components \( \beta \) and \( \tau(m_c(z)) \) will be equal as well. On one hand, we get
\[
    y = \frac{2z}{1 - ||z||^2} + \frac{4c}{(1 - ||c||^2)(1 - ||z||^2)} c \cdot z + \frac{2(1 + ||z||^2)c}{(1 - ||c||^2)(1 - ||z||^2)}, \tag{6.8}
\]
On the other hand, we find out that the complex number \( w = \psi(m_c(z)) \) is
\[
2 \left( \frac{z + c}{1 - \|z + c\|^2} \right) \frac{z + c}{1 - \|z + c\|^2}
\]
It is immediate to verify that \( \|cz + 1\|^2 - \|z + c\|^2 = (1 - \|c\|^2)(1 - \|z\|^2) \); thus, we have
\[
w = \frac{2(\|z\|^2 + z + c^2 \bar{z} + c)}{(1 - \|c\|^2)(1 - \|z\|^2)}
\]
The identity \( \|c + z\|^2 = \|c\|^2 + \|z\|^2 + \bar{c}z + \bar{z}c = 2c \cdot z \). Therefore we get
\[
\psi(m_c(z)) = \frac{2(\|z\|^2)c}{(1 - \|c\|^2)(1 - \|z\|^2)} + \frac{2z}{1 - \|z\|^2} + \frac{4c}{(1 - \|c\|^2)(1 - \|z\|^2)} c \cdot z = y.
\]
This concludes the proof.

We are now in a position to establish the main result of this Section, which provides a particularly significant characterization of the Lorentz group \( O(2,1) \).

**Theorem 6.4.** The group \( SO^+(2,1) \) is isomorphic to the Möbius group \( \mathcal{M}^+(D) \); therefore, there is an isomorphism
\[
SO^+(2,1) \cong PSL(2; \mathbb{R}). \tag{6.9}
\]
As consequence, \( \mathcal{M}^+(D) \) is the orientation-preserving isometry group of the Poincaré disc.

**Proof.** Theorem 2.3, along with the characterization of \( SO^+(2,1) \) in (4.1), tells us that any transformation in \( SO^+(2,1) \) is the product of a matrix of type \( \Lambda_u \) and an orthogonal matrix of type \( \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \), where \( u \in \mathbb{R}^2 \) and \( Q \in SO(2) \). Thus, because of the invertibility of both \( \zeta \) and \( \psi \), the first claim follows directly from Lemma 6.2 and Theorem 6.3. The isomorphism (6.9) is then an immediate consequence of Lemma 6.1. Since the map \( \zeta: \mathbb{H}_2 \to D \) is an isometry, the last claim follows directly from Corollary 5.3.

**References**


