THE POINCARÉ GROUP

LUCY HOROWITZ

Abstract. The Poincaré group is the (Lie) group of transformations that are allowed in the theory of special relativity. In this paper we will show that it is one of five possibilities for such a group, and note that it is the only one that agrees with experiment. To this end this we will use some of the basic theory of Lie groups, Lie algebras, and representation theory.

Contents

1. Introduction

In Newtonian mechanics and Galilean relativity we assume that the speed of light is infinite, or that it arrives at its destination instantaneously. This was determined to be false by Ole Rømer, a Dutch astronomer, while he was observing eclipses of Jupiter's moon Io. Of course, if light has a finite speed, then it must be that a moving light source will emit *faster* light than a stationary one! This notion was in turn determined to be false by Albert Michelson and Edward Moreley in 1887 when they attempted to measure the motion of the Earth with respect to the luminiferous aether, a hypothetical medium through which light waves were thought to propagate.

Einstein derived the Lorentz transformations (though he was not the first to do so) in his 1905 paper Zur Elektrodynamik bewegter Körper [3] by assuming what are referred to today as Einstein's postulates of special relativity and considering two observers moving relative to each other with constant velocity. The postulates are as follows:

Assumptions 1.1. First, the **Principle of Relativity** states that in every inertial frame of reference, the laws of physics are the same. That is, the same equations describe the same physical phenomena no matter which (non-accelerating) observer measures them. Moreover, there is no such thing as a "preferred frame of reference,"

Date: August 28, 2021.

which means that no inertial frame is inherently more "correct" or more "physical" than any other.

Second, Einstein assumed that the speed of light is both finite and invariant in all inertial reference frames. This value is currently defined in SI units as exactly $c = 299{,}792{,}458$ meters per second.

A consequence of the Principle of Relativity is that we should only consider affine transformations. An affine transformation is one for which the image of a straight world-line is a straight world-line. Note that an affine transformation is essentially composed of a linear transformation and a translation.

These assumptions allow for a rather elegant derivation of the Lorentz transformations by considering two inertial observers and applying simple algebraic substitutions to their coordinate labels for the same point in spacetime. However, assuming from the beginning that the speed of light is finite and invariant seems too high-level to be an assumption. It would be much more gratifying if it were possible to *deduce* that the speed of light must be finite and invariant just by considering more basic properties of spacetime, and indeed it is possible. The first to do so was Wladimir Ignatowski in his 1910 paper Das Relativitätsprinzip [7]. Many other scientists and mathematicians have built on his methods since their original publication. New derivations and refinements of earlier proofs are presented regularly, and it is on one such proof published just last year by Jean-Phillippe Anker and François Ziegler [1] that I will focus in this paper.

The proof presented by Anker and Ziegler shows that the Lorentz group $\mathrm{O}(3,1)$ is a candidate for the symmetry group of spacetime when considering only linear transformations. Here I present the suggested generalization of that proof to include affine transformations, and consequently show that the Poincaré group is a candidate for the symmetry group of spacetime.

By convention, capital letters will tend to denote elements of Lie groups or the Lie groups themselves, while lowercase letters will tend to denote variables, indices, and elements of Lie algebras. The reader is assumed to have an understanding of basic linear algebra, and for the rest of the material I hope that this paper is largely self-contained.

2. Preliminary Definitions

Definition 2.1. A **group** is a set G together with a binary operation $*$ such that

- $(a * b) * c = a * (b * c)$ for all $a, b, c \in G$. That is, $*$ is associative;
- There exists an element $e \in G$ such that $e * a = a * e = a$ for all $a \in G$. This is called the identity element of G;
- For every $a \in G$ there exists an inverse element $a^{-1} \in G$ such that $a^{-1} * a = a * a^{-1} = e$.

The full definition of a manifold is rather involved, and it is not particularly useful for the purposes of this paper. In lieu of a formal definition, we will say that an n-dimensional **smooth manifold** is a set that is locally homeomorphic to $\mathbb{R}^n$ and on which we can "do calculus." All we will need from manifolds in this paper is the next definition, a simple lemma or two, and dimensionality.

Definition 2.2. A **Lie group** (pronounced "lee") is a group that also happens to be a smooth manifold on which the group operation and the inverse operation are both smooth.

Example 2.3. The **general linear group**, denoted $\mathrm{GL}(n, \mathbb{R})$, is the group of invertible $n \times n$ matrices with real entries. It forms an n^2-dimensional Lie group.

All of the Lie groups we will consider in this paper are real matrix Lie groups, so they are subgroups/submanifolds of $\mathrm{GL}(n, \mathbb{R})$ for some n.

Example 2.4. The **orthogonal group**, denoted $\mathrm{O}(n)$, is the group of orthogonal $n \times n$ matrices. That is, it consists of $n \times n$ matrices A with real entries such that $A^\intercal A = AA^\intercal = I_n$, the $n \times n$ identity matrix. These matrices, when applied to vectors $\vec{v} \in \mathbb{R}^n$, preserve the usual Euclidean norm $||\vec{v}||$.

Definition 2.5. A **Lie algebra** is a vector space $\mathfrak{g}$ together with an operation called the **Lie bracket** $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ such that

- The bracket is bilinear: for all $x, y, z \in \mathfrak{g}$ and $\alpha, \beta \in \mathbb{R}$,
$$[\alpha x + \beta y, z] = \alpha[x, z] + \beta[y, z] \text{ and}$$
$$[x, \alpha y + \beta z] = \alpha[x, y] + \beta[y, z];$$
- $[x, x] = 0$ for all $x \in \mathfrak{g}$;
- The Jacobi identity is satisfied:
$$[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$$
for all $x, y, z \in \mathfrak{g}$.

Every Lie group has an associated Lie algebra.

Proposition 2.6. *For G a matrix Lie group, the associated Lie algebra (denoted $\mathfrak{g}$) is the set of all matrices X (not necessarily invertible) such that the matrix exponential $e^{tX} \in G$ for all $t \in \mathbb{R}$.*

Under the commutator bracket (i.e. define $[a, b] = ab - ba$), the set
$$\{X \in M_n(\mathbb{R}) \mid e^{tX} \in G \text{ for all } t \in \mathbb{R}\}$$
does in fact form a Lie algebra. The Lie algebra of a Lie group is essentially the tangent space of the manifold at the group's identity element with the extra structure of the bracket.

Example 2.7. The **general linear algbera**, denoted $\mathfrak{gl}(n, \mathbb{R})$, is the Lie algebra associated to the general linear group. It consists of *all* $n \times n$ matrices with real entries.

3. Statement of Theorem

First we must define the semidirect product, which serves to incorporate translations into our collection of allowed transformations.

Definition 3.1. Let H and N be groups and let $\phi : H \to \mathrm{Aut}(N)$ be a group homomorphism. The **semidirect product** $H \ltimes_\phi N$ is a group on $H \times N$ with operation $*$ given by
$$(h_1, n_1) * (h_2, n_2) = (h_1 h_2, n_1 \phi(h_1) n_2).$$
It is easy to see that the identity of this group is the ordered pair consisting of the identities of H and N while the inverse of an element (h, n) is $(h^{-1}, \phi(h^{-1}) n^{-1})$.

In this paper, we will only consider semidirect products of matrix Lie groups with Euclidean space as a group under addition. This allows us to use a nice matrix construction of the semidirect product. Moreover, the group homomorphism ϕ is "natural," and is the same for all semidirect products of Lie groups considered here. Thus we will only write $\ltimes$ without reference to ϕ because it is implicit in the construction.

Proposition 3.2. *Let H be a group of $n \times n$ matrices. Then $H \ltimes \mathbb{R}^n$ is isomorphic to the group*

$$\left\{ \begin{pmatrix} A & \vec{v} \\ \vec{0}^\intercal & 1 \end{pmatrix} \mid A \in H, \vec{v} \in \mathbb{R}^n \right\}$$

where the group homomorphism $\phi : H \to Aut(\mathbb{R}^n) = GL(n, \mathbb{R})$ in the semidirect product is the inclusion map.

Proof. Recall that the underlying set of this semidirect product is just

$$\{(A, \vec{v}) \mid A \in H, \vec{v} \in \mathbb{R}^n\}.$$

The bijective map is therefore a natural one:

$$(A, \vec{v}) \mapsto \begin{pmatrix} A & \vec{v} \\ \vec{0}^\intercal & 1 \end{pmatrix}.$$

Now we must show that this map respects the group operation. Let $A, B \in H$ and $\vec{v}, \vec{w} \in \mathbb{R}^n$. By definition, $(A, \vec{v}) * (B, \vec{w}) = (AB, \vec{v} + A\vec{w})$. Applying the bijection,

$$(AB, \vec{v} + A\vec{w}) \mapsto \begin{pmatrix} AB & \vec{v} + A\vec{w} \\ \vec{0}^\intercal & 1 \end{pmatrix}.$$

If we apply the bijection to $(A, \vec{v})$ and $(B, \vec{w})$ before investigating their product, we find that

$$\begin{pmatrix} A & \vec{v} \\ \vec{0}^\intercal & 1 \end{pmatrix} \begin{pmatrix} B & \vec{w} \\ \vec{0}^\intercal & 1 \end{pmatrix} = \begin{pmatrix} AB & \vec{v} + A\vec{w} \\ \vec{0}^\intercal & 1 \end{pmatrix}$$

under the usual (block) matrix multiplication. Thus the two groups are in fact isomorphic and whenever we write $H \ltimes \mathbb{R}^n$, we think of it as this collection of $(n+1) \times (n+1)$ matrices. $\square$

The element $(A, \vec{v}) \in H \ltimes \mathbb{R}^n$ corresponds to applying the linear transformation A then translating by $\vec{v}$. Despite the fact that this group consists of $(n+1) \times (n+1)$ real matrices, we do not need to increase the dimension of our model of spacetime. We can think of these matrices as acting on $\mathbb{R}^n \times \{1\}$, so that some vector $\vec{x} \in \mathbb{R}^n$ transforms as follows:

$$\begin{pmatrix} A & \vec{v} \\ \vec{0}^\intercal & 1 \end{pmatrix} \begin{pmatrix} \vec{x} \\ 1 \end{pmatrix} = \begin{pmatrix} A\vec{x} + \vec{v} \\ 1 \end{pmatrix}$$

Note that because the operation is matrix multiplication, H must consist only of invertible matrices. So too does $H \ltimes \mathbb{R}^n$: simply computing the determinant shows that

$$\det \begin{pmatrix} A & \vec{v} \\ \vec{0}^\intercal & 1 \end{pmatrix} = \det A.$$

Since $A \in H \subseteq \mathrm{GL}(n, \mathbb{R})$, it must have nonzero determinant by definition.

Now for the statement of the theorem, which comes from Anker and Zeigler [1]. We make frequent reference to $\mathbb{R}^{n+1}$ rather than $\mathbb{R}^n$ because we wish to draw attention to the fact that spacetime treats the single time coordinate differently from the n spatial coordinates.

Theorem 3.3. *Let $n \geq 2$ and define*

$$K = \left\{ \begin{pmatrix} B & \vec{0} \\ \vec{0}^\intercal & \pm 1 \end{pmatrix} \mid B \in O(n) \right\}.$$

Let G be a subgroup of $GL(n+1,\mathbb{R}) \ltimes \mathbb{R}^{n+1}$ such that

$$G \cap \left(\left\{ \begin{pmatrix} A & \vec{0} \\ \vec{0}^\intercal & x \end{pmatrix} \mid A \in GL(n,\mathbb{R}), x \in \mathbb{R}^\times \right\} \ltimes \mathbb{R}^{n+1} \right) = K \ltimes \mathbb{R}^{n+1}$$

and let

$$\mathfrak{p}_\sigma = \left\{ \begin{pmatrix} 0_n & \vec{v} \\ \sigma \vec{v}^\intercal & 0 \end{pmatrix} \mid \vec{v} \in \mathbb{R}^n \right\} \text{ and } \mathfrak{p}_\infty = \left\{ \begin{pmatrix} 0_n & \vec{0} \\ \vec{v}^\intercal & 0 \end{pmatrix} \mid v \in \mathbb{R}^n \right\}.$$

Either $G = K \ltimes \mathbb{R}^{n+1}$ or there exists $\sigma \in \mathbb{R} \cup \{\infty\}$ such that $G = K \exp(\mathfrak{p}_\sigma) \ltimes \mathbb{R}^{n+1}$.

The hypotheses of Theorem 3.3 encapsulate the *assumptions* about the nature of spacetime itself that are key to formulating the theory of special relativity.

First of all, the condition that G be a sub*group* of some other group is a fairly natural one. This is because what we are really doing with these transformations is jumping between inertial observers. The world would be a very strange place if Alice could not describe her own point of view, or if she could describe what Bob sees but Bob could not describe what Alice sees. This notion encapsulates the necessity of requiring the inclusion of the identity ("doing nothing") and of inverses.

The specific group that must contain G, namely $\mathrm{GL}(n+1,\mathbb{R}) \ltimes \mathbb{R}^{n+1}$, relates to the assumption that we consider *affine* transformations of spacetime. Loosely, these are linear transformations (i.e. ones that preserve the origin) together with translations. What is important to note is that such transformations take straight lines to straight lines. Specifically, if the *worldline* of a particle (its trajectory through spacetime) is straight in one reference frame, a transformation allowed in G must take it to another straight line. This essentially guarantees that Newton's first law, the law of inertia, is satisfied. Because we are doing special relativity, we consider neither accelerating bodies nor forces, so every particle's and every observer's worldline should always be straight, and no inertial (non-accelerating) reference frame is privileged over any other.

Finally, the intersection condition boils down to requiring that G includes the usual Euclidean symmetries (e.g. rotations and reflections) along with time flips and time skips. Moreover, every other transformation in G will in some sense mix space and time. A good example of such a transformation is a Lorentz boost, which (in the case $\sigma > 0$) describes how observers moving at constant velocity with respect to each other will measure the coordinates of the same event. It leads to phenomena called "time dilation" and "length contraction" in which a stationary (note that this is a relative term) observer will measure a uniformly moving observer's clock to tick slow and their meterstick (when pointed in the direction of motion) to shrink along the direction of motion.

A physical interpretation of the result is in Section 6 so that we can discuss the machinery we used to prove Theorem 3.3 simultaneously with its interpretation.

4. Outline of Proof

We begin with some lemmas on properties of semidirect products and define precisely *which* semidirect product we mean when we write $H \ltimes \mathbb{R}^{n+1}$ for some

Lie group or Lie algebra H. We then show that G does admit a Lie algebra and moreover find an explicit expression for it in terms of the $\mathfrak{p}_\sigma$ from above. Then we compute the normalizer of the Lie algebra of G and use the fact that it must contain G to find the possible forms of elements of G for different values of σ.

5. Proof of Theorem

Note that both $\mathrm{GL}(n+1,\mathbb{R})$ and $\mathbb{R}^{n+1}$ are Lie groups. First, we prove the following lemmas concerning the semidirect product.

Lemma 5.1. *Let L and H be Lie groups. Then the semidirect product $L \ltimes H$ is itself a Lie group.*

Proof. The Cartesian product of two manifolds is a manifold and the semidirect product of two groups is still a group. By definition of the semidirect product, the group operation and the inverse are both component-wise compositions of smooth functions, so the group $L \ltimes H$ is a Lie group. $\square$

This means that the group $\mathrm{GL}(n+1,\mathbb{R}) \ltimes \mathbb{R}^{n+1}$ is in fact a Lie group, and all of our techniques remain valid even across the semidirect product.

Definition 5.2. The **direct sum** of vector spaces V and W over the same field F is the set of all pairs (v, w) for $v \in V$ and $w \in W$ such that for all $\alpha \in F$, all $v_1, v_2 \in V$, and all $w_1, w_2 \in W$,

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_1, w_1 + w_2) \text{ and } \alpha(v_1, w_1) = (\alpha v_1, \alpha w_1).$$

Definition 5.3. A **derivation** of a Lie algebra $\mathfrak{h}$ is a linear transformation $\delta : \mathfrak{h} \to \mathfrak{h}$ such that $\delta([a, b]) = [\delta(a), b] + [a, \delta(b)]$. The set of all such transformations (denoted $\mathrm{Der}(\mathfrak{h})$) is a Lie subalgebra of $\mathfrak{gl}(\mathfrak{h})$.

Lemma 5.4. *Let $\mathfrak{h}$ be a Lie algebra and let $Aut(\mathfrak{h}) = GL(\mathfrak{h})$ be the group of its automorphisms. The Lie algebra of $Aut(\mathfrak{h})$ is exactly $Der(\mathfrak{h})$.*

Proof. First, we can see that $\mathrm{Der}(\mathfrak{g}) \subseteq \mathrm{Lie}(\mathrm{Aut}(\mathfrak{g}))$ because the exponential of any matrix is always invertible. Thus for any derivation d of $\mathfrak{g}$, e^{td} is an invertible transformation of $\mathfrak{g}$ for all $t \in \mathbb{R}$ and is therefore contained in $\mathrm{Aut}(\mathfrak{g})$.

Next we need to show that $\mathrm{Lie}(\mathrm{Aut}(\mathfrak{g})) \subseteq \mathrm{Der}(\mathfrak{g})$. Let $A \in \mathrm{Lie}(\mathrm{Aut}(\mathfrak{g}))$. Then $e^{tA} \in \mathrm{Aut}(\mathfrak{g})$ for all $t \in \mathbb{R}$, so these matrices are all Lie algebra homomorphisms $\mathfrak{g} \to \mathfrak{g}$. Thus for all $x, y \in \mathfrak{g}$,

$$e^{tA}[X, Y] = [e^{tA}X, e^{tA}Y].$$

We can differentiate with respect to t and evaluate at $t = 0$ to see that by the product rule,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{tA}[x, y]\right)\Big|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\left[e^{tA}x, e^{tA}y\right]\right)\Big|_{t=0} \\
A[x, y] &= \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{tA}\left(xe^{tA}y - ye^{tA}x\right)\right)\Big|_{t=0} \\
&= Axy - Ayx + xAy - yAx \\
&= [Ax, y] + [x, Ay]
\end{aligned}$$

so that $A \in \mathrm{Der}(\mathfrak{g})$, as desired. $\square$

Definition 5.5. Let $\mathfrak{l}$ and $\mathfrak{h}$ be Lie algebras and let $\phi : \mathfrak{l} \to \mathrm{Der}(\mathfrak{h})$ be a Lie group homomorphism. The **semidirect product** $\mathfrak{l} \ltimes \mathfrak{h}$ is a Lie algebra on the *vector space* $\mathfrak{l} \oplus \mathfrak{h}$ with Lie bracket

$$[(\ell_1, h_1), (\ell_2, h_2)] = ([\ell_1, \ell_2], [h_1, h_2] + \phi(\ell_1)(h_2) - \phi(\ell_2)(h_1))$$

for all $\ell_1, \ell_2 \in \mathfrak{l}$ and $h_1, h_2 \in \mathfrak{h}$.[4]

Note that the bracket "agrees" on both $\mathfrak{l}$ and $\mathfrak{h}$:

$$[(\ell_1, 0), (\ell_2, 0)] = [\ell_1, \ell_2]_{\mathfrak{l}} \text{ and } [(0, h_1), (0, h_2)] = [h_1, h_2]_{\mathfrak{h}}.$$

Moreover,

$$[(0, h), (\ell, 0)] = \phi(\ell)(h).$$

Proposition 5.6. *The Lie algebra of the semidirect product of Lie groups $L \ltimes_\phi H$ is the semidirect product of the Lie algebras $\mathfrak{l}$ and $\mathfrak{h}$ of L and H respectively. That is,*

$$Lie(L \ltimes_\phi H) = \mathfrak{l} \ltimes_\psi \mathfrak{h}$$

where $\phi : L \to Aut(H)$ is the group homomorphism associated to the semidirect product of groups and $\psi : \mathfrak{h} \to Der(\mathfrak{l})$ is the Lie algebra homomorphism associated to the semidirect product of Lie algebras.

The substance of this proposition is that the semidirect product of Lie groups "carries over" in a natural way to the associated Lie algebra. In our case, we use a matrix construction that is very similar to the one for the semidirect product of Lie groups, and the next proposition provides a sort of proof of a special case of the previous one.

Proposition 5.7. *Let $\mathfrak{h}$ be a Lie algebra of $n \times n$ matrices. Then $\mathfrak{h} \ltimes \mathbb{R}^n$ is isomorphic to the group*

$$\left\{ \begin{pmatrix} h & \vec{v} \\ \vec{0}^\intercal & 0 \end{pmatrix} \mid h \in \mathfrak{h}, \vec{v} \in \mathbb{R}^n \right\}$$

where the algebra homomorphism $\phi : \mathfrak{h} \to Der(\mathbb{R}^n)$ in the semidirect product is the inclusion map.

Proof. As before, the underlying set of the semidirect product is

$$\{(h, \vec{v}) \mid h \in \mathfrak{h}, \vec{v} \in \mathbb{R}^n\}$$

and our bijection will be simply

$$(h, \vec{v}) \mapsto \begin{pmatrix} h & \vec{v} \\ \vec{0}^\intercal & 0 \end{pmatrix}.$$

Note that the Lie algebra of $\mathbb{R}^n$ considered as a Lie group under vector addition is itself so that by Lemma 5.4, the Lie algebra of $\mathrm{Aut}(\mathbb{R}^n) = \mathrm{GL}(n, \mathbb{R})$ is $\mathrm{Der}(\mathbb{R}^n)$. Now we check if the transformation respects the bracket so that it is an isomorphism. We have

$$\begin{aligned} [(h, \vec{v}), (k, \vec{w})] &= ([h, k], [\vec{v}, \vec{w}] + h\vec{w} - k\vec{v}) \\ &= (hk - kh, h\vec{w} - k\vec{v}) \end{aligned}$$

while

$$\left[\begin{pmatrix} h & \vec{v} \\ \vec{0}^\intercal & 0 \end{pmatrix}, \begin{pmatrix} k & \vec{w} \\ \vec{0}^\intercal & 0 \end{pmatrix}\right] = \begin{pmatrix} h & \vec{v} \\ \vec{0}^\intercal & 0 \end{pmatrix}\begin{pmatrix} k & \vec{w} \\ \vec{0}^\intercal & 0 \end{pmatrix} - \begin{pmatrix} k & \vec{w} \\ \vec{0}^\intercal & 0 \end{pmatrix}\begin{pmatrix} h & \vec{v} \\ \vec{0}^\intercal & 0 \end{pmatrix}$$
$$= \begin{pmatrix} hk - kh & h\vec{w} - k\vec{v} \\ \vec{0}^\intercal & 0 \end{pmatrix}.$$

Thus the bijection respects the Lie bracket and it is in fact a Lie algebra isomorphism, and from here forward we think of the Lie algebra $\mathfrak{h} \ltimes \mathbb{R}^n$ as this collection of $(n+1) \times (n+1)$ matrices. □

It is quite easy to see that if H is a Lie group with Lie algebra $\mathfrak{h}$, then the Lie algebra of $H \ltimes \mathbb{R}^n$ is exactly $\mathfrak{h} \ltimes \mathbb{R}^n$. Taking the exponential of an element $(h, \vec{v})$ of $\mathfrak{h} \ltimes \mathbb{R}^n$, we have

$$\exp\left(t\begin{pmatrix} h & \vec{v} \\ \vec{0}^\intercal & 0 \end{pmatrix}\right) = \begin{pmatrix} e^{th} & h^{-1}(e^{th} - I_n)\vec{v} \\ \vec{0}^\intercal & 1 \end{pmatrix}.$$

By definition, $e^{th} \in H$. Since $h^{-1}(e^{th} - I_n)\vec{v} \in \mathbb{R}^n$, the resulting matrix is guaranteed to be in $H \ltimes \mathbb{R}^n$.

From these we can see that the group $M = \mathrm{GL}(n+1, \mathbb{R}) \ltimes \mathbb{R}^{n+1}$ is a Lie group with Lie algebra

$$\mathfrak{m} = \mathfrak{gl}(n+1, \mathbb{R}) \ltimes \mathbb{R}^{n+1}.$$

Like with the group M, we have a nice matrix representation of $\mathfrak{m}$.

Anker and Zeigler are able to streamline the proof given by Gorini [5] by employing a result about the Lie algebra of G given by Bourbaki [2] that is as follows:

Lemma 5.8. *Let L be a finite-dimensional Lie group with subgroup H. Then there exists a unique differentiable manifold structure on H such that for each $r \in \mathbb{N}$, every function from every C^r-manifold V to H (i.e. every $f : V \to H$ such that f has at least r continuous derivatives) is C^r when considered as a function $V \to H$ if and only if it is C^r as a function $V \to L$. Moreover, this manifold structure makes H into a Lie group such that the inclusion map into L is an immersion. The Lie algebra of H is a Lie subalgebra of the Lie algebra of L.*

Applying Lemma 5.8 to our subgroup G, we see that it is a Lie group with Lie algebra

$$\mathfrak{g} = \{Z \in \mathfrak{gl}(n+1, \mathbb{R}) \ltimes \mathbb{R}^{n+1} \mid e^{tZ} \in G \text{ for all } t \in \mathbb{R}\}.$$

Next we show that this Lie algebra is either equal to

$$\mathfrak{k} \ltimes \mathbb{R}^{n+1} = \mathrm{Lie}(K \ltimes \mathbb{R}^{n+1})$$

or the direct sum of this Lie algebra with $\mathfrak{p}_\sigma$ for some $\sigma \in \mathbb{R} \cup \{\infty\}$, depending on what the intersection in the statement of Theorem 3.3 happens to be. Before we can go any further with the computation, we will need to take the scenic route through some representation theory.

Definition 5.9. A **representation** of a Lie group H (or Lie algebra $\mathfrak{h}$) is a Lie group (or algebra) homomorphism $\Pi : H \to \mathrm{GL}(V)$ (or $\pi : \mathfrak{h} \to \mathfrak{gl}(V)$). It consists of the pair containing the homomorphism Π (or π) and the vector space V.

Essentially, a representation dictates how the Lie group (or algebra) acts on a vector space. We often refer to representations just by their associated vector spaces.

Example 5.10. The **adjoint representation** of a Lie group H with associated Lie algbera $\mathfrak{h}$ is a representation $\mathrm{Ad} : H \to \mathrm{GL}(\mathfrak{h})$. For any $A \in H$, the adjoint representation gives rise to an (invertible!) linear transformation $\mathrm{Ad}_A : \mathfrak{h} \to \mathfrak{h}$ where $\mathrm{Ad}_A(h) = AhA^{-1}$.

Lemma 5.11. *Suppose that X is an invertible matrix. Then for any matrix A, $e^{XAX^{-1}} = Xe^A X^{-1}$.*

Proof. A simple computation shows that $(XAX^{-1})^k = XA^k X^{-1}$ for any integer k. Thus considering the power series definition of the matrix exponential, we have

$$e^{t(XAX^{-1})} = Xe^A X^{-1} \in H$$

as desired. $\square$

Perhaps this whole notion of treating the elements of Lie groups and Lie algebras as matrices seems a little suspicious now that we know about representations. In fact, we have secretly been using a representation the whole time! All of our matrix Lie groups are subgroups of $\mathrm{GL}(n+1, \mathbb{R}) \ltimes \mathbb{R}^{n+1}$ and our matrix Lie algebras are subalgebras of $\mathfrak{gl}(n+1, \mathbb{R}) \ltimes \mathbb{R}^{n+1}$. We can think about our matrix elements as being images of (abstract) group elements under a "standard representation" that is just given by inclusion into the general linear group or algebra, which acts on Euclidean space. This allows us to treat the groups and algebras themselves as representations to get some useful results.

When considered as a vector space, $\mathfrak{m} = \mathfrak{gl}(n+1, \mathbb{R}) \ltimes \mathbb{R}^{n+1}$ can be written as the direct sum of the following four vector spaces:

$$\begin{aligned}
m_1 &= \left\{ \begin{pmatrix} tI_n & \vec{0} \\ \vec{0}^\intercal & s \end{pmatrix} \mid t, s \in \mathbb{R} \right\} \ltimes \mathbb{R}^{n+1} \\
m_2 &= \left\{ \begin{pmatrix} A & \vec{0} \\ \vec{0}^\intercal & 0 \end{pmatrix} \mid A^\intercal = -A \right\} \ltimes \mathbb{R}^{n+1} = \mathfrak{k} \ltimes \mathbb{R}^{n+1} \\
m_3 &= \left\{ \begin{pmatrix} A & \vec{0} \\ \vec{0}^\intercal & 0 \end{pmatrix} \mid A^\intercal = A \right\} \ltimes \mathbb{R}^{n+1} \\
m_4 &= \left\{ \begin{pmatrix} 0_n & \vec{v} \\ \vec{w}^\intercal & 0 \end{pmatrix} \mid v, w \in \mathbb{R}^n \right\} \ltimes \mathbb{R}^{n+1}
\end{aligned}$$

so that $\mathfrak{m} = m_1 \oplus m_2 \oplus m_3 \oplus m_4$. Loosely, this is because any matrix can be written as the sum of a symmetric matrix and an anti-symmetric matrix.

Definition 5.12. A **subrepresentation** of a representation (ρ, V) is a representation $(\rho_{|W}, W)$ where W is a subspace of the vector space V such that $\rho(x)w \in W$ for all $w \in W$ and $\rho_{|W}(x) = \rho(x)_{|W}$ for all x in the group or Lie algebra that is being represented.

Definition 5.13. Let H be a group and let $\{T_i\}_{i=1}^k$ be representations of G acting on $\{V_i\}_{i=1}^k$ respectively. Then the **direct sum** of the representations T_i (which acts on the usual direct sum of the vector spaces V_i) is defined by

$$\left(\bigoplus_{i=1}^m T_i(X) \right) (v_1, \ldots, v_k) = (T_1(X)v_1, \ldots, T_k(X)v_k)$$

for all $X \in H$.

By the intersection hypothesis of Theorem 3.3, we can think of the orthogonal group $\mathrm{O}(n)$ as a subgroup of M. Really, $K \ltimes \mathbb{R}^{n+1} \subseteq G$ has a subgroup that is isomorphic to $\mathrm{O}(n)$: it is the subgroup of matrices in $K \ltimes \mathbb{R}^{n+1}$ whose "center" entry is $+1$ and whose translation component is $\vec{0}$. Consider the adjoint representation $\mathrm{Ad} : M \to \mathrm{GL}(\mathfrak{m})$. We can restrict it to $\mathrm{O}(n) \subset M$ to get a representation $\mathrm{Ad}_{|G} : O(n) \to \mathrm{GL}(\mathfrak{m})$ of $O(n)$.

Proposition 5.14. *For each i, the representation $Ad_i : O(n) \to GL(m_i)$ given by*

$$Ad_i(X) = Ad_{|O(n)}(X)$$

for each $X \in m_i$ forms a subrepresentation of $O(n)$. Moreover,

$$\mathfrak{m} = \bigoplus_{i=1}^{4} m_i$$

as the direct sum of these representations.

Proof. We show the first result for each $i \in \{1, 2, 3, 4\}$. An element $\boldsymbol{\Theta} \in \mathrm{O}(n) \subset M$ has the form

$$\begin{pmatrix} \Theta & \vec{0} & \vec{0} \\ \vec{0}^{\intercal} & 1 & 0 \\ \vec{0}^{\intercal} & 0 & 1 \end{pmatrix}$$

for Θ in the usual $\mathrm{O}(n)$.

(1) Let $g \in m_1$. Then we can write

$$g = \begin{pmatrix} tI_n & \vec{0} & \vec{0} \\ \vec{0}^{\intercal} & s & 0 \\ \vec{0}^{\intercal} & 0 & 0 \end{pmatrix}$$

for $t, s \in \mathbb{R}$. Applying the adjoint representation we can see that

$$\mathrm{Ad}_{\boldsymbol{\Theta}}(g) = \begin{pmatrix} \Theta & \vec{0} & \vec{0} \\ \vec{0}^{\intercal} & 1 & 0 \\ \vec{0}^{\intercal} & 0 & 1 \end{pmatrix} \begin{pmatrix} tI_n & \vec{0} & \vec{0} \\ \vec{0}^{\intercal} & s & 0 \\ \vec{0}^{\intercal} & 0 & 0 \end{pmatrix} \begin{pmatrix} \Theta^{-1} & \vec{0} & \vec{0} \\ \vec{0}^{\intercal} & 1 & 0 \\ \vec{0}^{\intercal} & 0 & 1 \end{pmatrix} = g \in m_1$$

so that m_1 is indeed a subrepresentation $\mathfrak{m}$.

(2) Let $g \in m_2$. Then we can write

$$g = \begin{pmatrix} A & \vec{0} & \vec{0} \\ \vec{0}^{\intercal} & 0 & 0 \\ \vec{0}^{\intercal} & 0 & 0 \end{pmatrix}$$

for antisymmetric A. Applying the adjoint representation we can see that

$$\mathrm{Ad}_{\boldsymbol{\Theta}}(g) = \begin{pmatrix} \Theta & \vec{0} & \vec{0} \\ \vec{0}^{\intercal} & 1 & 0 \\ \vec{0}^{\intercal} & 0 & 1 \end{pmatrix} \begin{pmatrix} A & \vec{0} & \vec{0} \\ \vec{0}^{\intercal} & 0 & 0 \\ \vec{0}^{\intercal} & 0 & 0 \end{pmatrix} \begin{pmatrix} \Theta^{-1} & \vec{0} & \vec{0} \\ \vec{0}^{\intercal} & 1 & 0 \\ \vec{0}^{\intercal} & 0 & 1 \end{pmatrix} = \begin{pmatrix} \Theta A \Theta^{-1} & \vec{0} & \vec{0} \\ \vec{0}^{\intercal} & 0 & 0 \\ \vec{0}^{\intercal} & 0 & 0 \end{pmatrix}.$$

To check whether this is in $m_2 = \mathfrak{k} \ltimes \mathbb{R}^{n+1}$, we can show that $e^{t(\Theta A \Theta^{-1})} \in K \ltimes \mathbb{R}^{n+1}$ for all $t \in \mathbb{R}$. By Lemma 5.11, $e^{t(\Theta A \Theta^{-1})} = \Theta e^{tA} \Theta^{-1}$. Because $\Theta \in \mathrm{O}(n) \subset K$ and $A \in \mathfrak{k}$, the product must be in K. Thus $\boldsymbol{\Theta} g \boldsymbol{\Theta}^{-1} \in m_2$ as desired.

(3) Let $g \in m_2$. Then we can write

$$g = \begin{pmatrix} A & \vec{0} & \vec{0} \\ \vec{0}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}$$

for symmetric A. The application of the adjoint representation looks identical to the application in (2). Simply verifying that

$$(\Theta g \Theta^{-1})^\intercal = (\Theta^{-1})^\intercal A^\intercal \Theta^\intercal = \Theta A \Theta^{-1}$$

by orthogonality is sufficient to show that $\mathrm{Ad}_{\boldsymbol{\Theta}}(g) \in m_3$ for all $g \in m_3$.

(4) Let $g \in m_4$. Then we can write

$$g = \begin{pmatrix} 0_n & \vec{v} & \vec{0} \\ \vec{w}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}.$$

Applying the adjoint representation we can see that

$$\mathrm{Ad}_{\boldsymbol{\Theta}}(g) = \begin{pmatrix} \Theta & \vec{0} & \vec{0} \\ \vec{0}^\intercal & 1 & 0 \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} \begin{pmatrix} 0_n & \vec{v} & \vec{0} \\ \vec{w}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix} \begin{pmatrix} \Theta^{-1} & \vec{0} & \vec{0} \\ \vec{0}^\intercal & 1 & 0 \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0_n & \Theta\vec{v} & \vec{0} \\ \vec{w}^\intercal \Theta^{-1} & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}$$

which is clearly in m_4.

Thus each of the m_i are subrepresentations of $\mathfrak{m}$ under the adjoint representation. Then we can write $\mathrm{Ad} : \mathrm{O}(n) \to \mathrm{GL}(\mathfrak{m})$ as the direct sum

$$\bigoplus_{i=1}^{4} \mathrm{Ad} : \mathrm{O}(n) \to \mathrm{GL}(m_i)$$

by simple application of the definition. □

Definition 5.15. A representation of a Lie group or Lie algebra is **irreducible** if it has no nontrivial subrepresentations.

Definition 5.16. A representation of a Lie algebra is **completely reducible** if it can be written as the direct sum of finitely many irreducible representations.

Lemma 5.17. *Let V be a completely reducible representation of the Lie group H. Then every invariant subspace of V is also completely reducible. Moreover, any such invariant subspace can be written as the direct sum of its intersections with the irreducible representations that add to the "parent" representation.* [6]

We do have that $\mathfrak{g}$ is an invariant subspace of $\mathfrak{m}$ under the adjoint representation of the orthogonal group $\mathrm{O}(n)$ when considered as a subgroup of $K \ltimes \mathbb{R}^{n+1} \subset G \subset M$. As above, elements of this subgroup take the form

$$\boldsymbol{\Theta} = \begin{pmatrix} \Theta & \vec{0} & \vec{0} \\ \vec{0}^\intercal & 1 & 0 \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix}$$

for $\Theta \in \mathrm{O}(n)$ while elements of $\mathfrak{g}$ take the form

$$g = \begin{pmatrix} A & \vec{b} & \vec{v}_- \\ \vec{c}^\intercal & d & v_{n+1} \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}$$

with some as-yet-unknown restrictions on A, $\vec{b}$, $\vec{c}$, and d. The vector $\vec{v}_-$ is just $\vec{v}$ from our construction of the semidirect product without the last component v_{n+1}. This notation is used so that the block matrices are of equal dimension.

The adjoint representation $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ takes an element of G to some invertible linear transformation of its Lie algebra $\mathfrak{g}$. We wish to show that the linear transformation given by our $\mathbf{\Theta}$ leaves $\mathfrak{g}$ invariant. That is, that $\mathrm{Ad}_{\mathbf{\Theta}}(g) \in \mathfrak{g}$ for any $g \in \mathfrak{g}$. Applying the adjoint representation to our matrix $\mathbf{\Theta}$ shows that

$$\mathrm{Ad}_{\mathbf{\Theta}}(g) = \mathbf{\Theta} g \mathbf{\Theta}^{-1}.$$

By definition, this will be in $\mathfrak{g}$ if and only if $e^{t(\mathbf{\Theta} g \mathbf{\Theta}^{-1})}$ is in G for all $t \in \mathbb{R}$. By Lemma 5.11,

$$e^{t(\mathbf{\Theta} g \mathbf{\Theta}^{-1})} = \mathbf{\Theta} e^{tg} \mathbf{\Theta}^{-1}.$$

Since $g \in \mathfrak{g}$, by definition $e^{tg} \in G$ for all $t \in \mathbb{R}$. Since $\mathbf{\Theta}$ is in G, the whole product is therefore in G and $\mathfrak{g}$ is an invariant subspace.

After applying Lemma 5.17 to $\mathfrak{g} \subset \mathfrak{m}$, we can rewrite $\mathfrak{g}$ as

$$\mathfrak{g} = \bigoplus_{i=1}^{4} (m_i \cap \mathfrak{g}).$$

By properties of the intersection, we have from the hypothesis of Theorem 3.3 that $K \ltimes \mathbb{R}^{n+1} \subseteq G$. Because passing to the Lie algebra preserves (not necessarily proper) subgroup relations in the Lie subalgebra, we know that $\mathfrak{k} \ltimes \mathbb{R}^{n+1} \subseteq \mathfrak{g}$. But K is isomorphic to two copies of $\mathrm{O}(n)$ (that is, $\mathrm{O}(n) \times \{\pm 1\}$), so its Lie algebra as the tangent space to the identity (which is contained only in the $+1$ copy of $\mathrm{O}(n)$) is isomorphic to $\mathfrak{o}(n)$. This Lie algbera consists exactly of the skew-symmetric matrices. Consequently, $m_2 = \mathfrak{k}$.

Now suppose that some $a \neq 0_{n+2}$ is contained in the intersection of $\mathfrak{g}$ and m_1. Then the exponential $e^{ta} \in G$ for all $t \in \mathbb{R}$ so that in particular,

$$\exp\left(\begin{pmatrix} I_n & \vec{0} & \vec{0} \\ \vec{0}^\intercal & 1 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}\right) = \begin{pmatrix} e_n & \vec{0} & \vec{0} \\ \vec{0}^\intercal & e & 0 \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} \in G.$$

But this matrix is also in the intersection of G with M, so by definition it is in $K \ltimes \mathbb{R}^{n+1}$. However, $e \neq \pm 1$ despite the fact that if this matrix were in K, the "middle" entry would have to be equal to ± 1. Therefore there can be no nonzero element of m_1 in $\mathfrak{g}$, so $m_1 \cap \mathfrak{g} = \{0_{n+2}\}$.

In the case of m_3, note that no nonzero symmetric matrix is skew-symmetric. The exponential of any nonzero element of m_3 is also in M, which can be seen from a similar computation to the above:

$$\exp\left(\begin{pmatrix} A & \vec{0} & \vec{0} \\ \vec{0}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}\right) = \begin{pmatrix} e^A & \vec{0} & \vec{0} \\ \vec{0}^\intercal & 1 & 0 \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} \in G.$$

But the exponential of a nonzero symmetric matrix cannot be an orthogonal matrix because the Lie algbera of $\mathrm{O}(n)$ consists of just the antisymmetric matrices. Therefore there can be no nonzero element of m_3 in $\mathfrak{g}$, so $m_3 \cap \mathfrak{g} = \{0_{n+2}\}$.

Finally we can show that m_4 behaves exactly as some $\mathfrak{p}_\sigma$. That is, the two vectors $\vec{v}$ and $\vec{w}$ must be scalar multiples of each other. The matrix $\mathcal{O} = \vec{v}\vec{w}^\intercal - \vec{w}\vec{v}^\intercal$

is antisymmetric, and we can therefore choose it to be the "top left" part of an element of $m_2 = \mathfrak{k} \ltimes \mathbb{R}^{n+1}$. Then computing the following Lie bracket, we have

$$\begin{aligned}\left[\begin{pmatrix} 0_n & \vec{v} & \vec{0} \\ \vec{w}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}, \left[\begin{pmatrix} 0_n & \vec{v} & \vec{0} \\ \vec{w}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}, \begin{pmatrix} \mathcal{O} & \vec{0} & \vec{0} \\ \vec{0}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}\right]\right] &= \left[\begin{pmatrix} 0_n & \vec{v} & \vec{0} \\ \vec{w}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0_n & -\mathcal{O}\vec{v} & \vec{0} \\ \vec{w}^\intercal\mathcal{O} & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}\right] \\ &= \begin{pmatrix} \vec{v}\vec{w}^\intercal\mathcal{O} + \mathcal{O}\vec{v}\vec{w}^\intercal & \vec{0} & \vec{0} \\ \vec{0}^\intercal & -2(\vec{w}^\intercal\mathcal{O}\vec{v}) & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}.\end{aligned}$$

Because $\mathfrak{g}$ is a Lie algebra, it is closed under its bracket. Moreover, $m_4 \subseteq \mathfrak{g}$, so this matrix must be in $\mathfrak{g}$. Our lovely computations show that $\mathfrak{g} \subseteq (m_2 \oplus m_4)$, so by definition of m_2 and m_4 the entry $-2(\vec{w}^\intercal\mathcal{O}\vec{v})$ must be zero. We can rewrite this entry as $2(||\vec{v}||^2||\vec{w}||^2 - \langle\vec{v},\vec{w}\rangle^2)$, which is zero exactly when equality is satisfied for the Cauchy-Schwarz inequality. This condition is met when there exists $\sigma \in \mathbb{R}$ for which $\vec{v} = \sigma\vec{w}$.

Now suppose that $\mathfrak{g}$ contains the matrices

$$\begin{pmatrix} 0 & \vec{v} & \vec{0} \\ \alpha\vec{v}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & \vec{v} & \vec{0} \\ \beta\vec{v}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}$$

for real numbers $\alpha \neq \beta$. Then for any $A \in \mathrm{O}(n)$, we can apply the adjoint representation to the first one to find

$$\begin{pmatrix} A & \vec{0} & \vec{0} \\ \vec{0}^\intercal & 1 & 0 \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & \vec{v} & \vec{0} \\ \alpha\vec{v}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} \begin{pmatrix} A^{-1} & \vec{0} & \vec{0} \\ \vec{0}^\intercal & 1 & 0 \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & A\vec{v} & \vec{0} \\ \alpha\vec{v}^\intercal A^{-1} & 0 & 0 \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} \in \mathfrak{g}$$

again by invariance under the adjoint representation. Then any linear combination

$$x\begin{pmatrix} 0 & A\vec{v} & \vec{0} \\ \alpha\vec{v}^\intercal A^{-1} & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & \vec{v}\vec{0} & \\ \beta\vec{v}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}$$

is contained in $\mathfrak{g}$. But $\mathrm{O}(n)$ contains all of the rotation matrices, so we can choose A to be any one of these to give rise to any element

$$\begin{pmatrix} 0 & \vec{u} & \vec{0} \\ \vec{w}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix} \in m_4.$$

Thus $m_4 \subseteq \mathfrak{g}$ so that the direct sum is exactly $\mathfrak{g} = m_2 \oplus m_4$. But this is not possible because this sum does not form a Lie algebra as it is not closed under the bracket:

$$\left[\begin{pmatrix} 0 & \vec{v} & \vec{0} \\ \vec{w}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & \vec{u}\vec{0} & \\ \vec{z}^\intercal & 0 & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}\right] = \begin{pmatrix} \vec{v}\vec{z}^\intercal - \vec{u}\vec{w}^\intercal & \vec{0} & \vec{0} \\ \vec{0}^\intercal & \vec{w}^\intercal\vec{u} - \vec{z}^\intercal\vec{v} & 0 \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix} \notin \mathfrak{g}$$

for $w^\intercal\vec{u} - \vec{z}^\intercal\vec{v} \neq 0$. Therefore there can be only one value of σ allowed, so

$$m_4 \cap \mathfrak{g} = \mathfrak{p}_\sigma \ltimes \mathbb{R}^{n+1}$$

for that value of σ.

It is important to note that depending on what form the group G actually takes, $\mathfrak{g}$ may not include any of the elements of $\mathfrak{p}_\sigma \ltimes \mathbb{R}^{n+1}$. All we have shown is that if it does, it will look like this. When there is no $\mathfrak{p}_\sigma$ factor to consider, $\mathfrak{g} = \mathfrak{k} \ltimes \mathbb{R}^{n+1}$ exactly.

Now we need a notion of the normalizer of $\mathfrak{g}$ for Lie algebras of our Lie groups. It consists of elements of $\mathrm{GL}(n+1, \mathbb{R}) \ltimes \mathbb{R}^{n+1}$ that leave $\mathfrak{g}$ invariant under conjugation. We use the normalizer to find G because as it turns out, G is contained in the normalizer of $\mathfrak{g}$.

Definition 5.18. The **normalizer** of a Lie algebra $\mathfrak{h}$ associated to a Lie group $H \subset \mathrm{GL}(n, \mathbb{R})$ is given by

$$N(\mathfrak{h}) = \{X \in \mathrm{GL}(n, \mathbb{R}) \mid X\mathfrak{h}X^{-1} \subset \mathfrak{h}\}.$$

Lemma 5.19. *Let $H \subset GL(n, \mathbb{R})$ be a Lie group with associated Lie algebra $\mathfrak{h}$. Then H is contained in the normalizer $N(\mathfrak{h})$.*

Proof. Let $X \in H$. We wish to show that for any $h \in \mathfrak{h}$, $XhX^{-1} \in \mathfrak{h}$. Since the Lie algebra $\mathfrak{h}$ consists exactly of those matrices $Y \in \mathrm{GL}(n, \mathbb{R})$ for which $e^{tY} \in H$ for all $t \in \mathbb{R}$, we need only exponentiate the product and show that it is indeed in H for all $t \in \mathbb{R}$. This follows immediately from Lemma 5.11. $\square$

For two of the cases in Theorem 3.3 ($\sigma = 0$ and $\sigma = \infty$) we do not even need to compute the normalizer! First consider the case $\sigma = 0$, and note that by definition of $\mathfrak{p}_\sigma$ the proof for the $\sigma = \infty$ case is essentially identical, just transposed. Then

$$\mathfrak{p}_0 = \left\{ \begin{pmatrix} 0_n & \vec{v} \\ \vec{0}^\intercal & 0 \end{pmatrix} \mid \vec{v} \in \mathbb{R}^n \right\}$$

so that elements of $\mathfrak{g} = (\mathfrak{k} \oplus \mathfrak{p}_0) \ltimes \mathbb{R}^{n+1}$ take the form

$$g = \begin{pmatrix} B & \vec{b} & \vec{y}_- \\ \vec{0}^\intercal & 0 & y_{n+1} \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}.$$

Suppose that $B \neq 0_n$ and $\vec{b} \neq \vec{0}$. Now choose

$$a = \begin{pmatrix} A & \vec{u} & \vec{a}_- \\ \vec{w}^\intercal & x & a_{n+1} \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} \in N(\mathfrak{g}).$$

Then by invariance, each $g \in \mathfrak{g}$ has an associated $g' \in \mathfrak{g}$ such that $ag = g'a$. That is,

$$\begin{pmatrix} A & \vec{u} & \vec{a}_- \\ \vec{w}^\intercal & x & a_{n+1} \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} \begin{pmatrix} B & \vec{b} & \vec{y}_- \\ \vec{0}^\intercal & 0 & y_{n+1} \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix} = \begin{pmatrix} C & \vec{c} & \vec{z}_- \\ \vec{0}^\intercal & 0 & z_{n+1} \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix} \begin{pmatrix} A & \vec{u} & \vec{a}_- \\ \vec{w}^\intercal & x & a_{n+1} \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} AB & A\vec{b} & \vec{v}_- \\ \vec{w}^\intercal B & \vec{w}^\intercal \vec{b} & v_{n+1} \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix} = \begin{pmatrix} CA + \vec{c}\vec{w}^\intercal & C\vec{u} + x\vec{c} & \vec{v}'_- \\ \vec{0}^\intercal & 0 & v'_{n+1} \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix}.$$

Because we assumed that B and $\vec{b}$ are nonzero, it must be that $\vec{w} = \vec{0}$ for all $a \in N(\mathfrak{g})$. Since the normalizer contains G, every element of G takes the form

$$\begin{pmatrix} A & \vec{u} & \vec{v}_- \\ \vec{0}^\intercal & t & v_{n+1} \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} = \begin{pmatrix} A & \vec{0} & \vec{x}_- \\ \vec{0}^\intercal & t & x_{n+1} \\ \vec{0}^\intercal & 0 & 1 \end{pmatrix} \exp\left(\begin{pmatrix} 0_n & A^{-1}\vec{u} & \vec{y}_- \\ \vec{0}^\intercal & 0 & y_{n+1} \\ \vec{0}^\intercal & 0 & 0 \end{pmatrix} \right).$$

Because we have chosen $\sigma = 0$, the matrices corresponding to the second factor are certainly in G, and because this must be closed under matrix multiplication, the matrices corresponding to the first factor must also be in G. Thus by hypothesis these matrices are in $K \ltimes \mathbb{R}^{n+1}$. Thus we may write

$$\begin{aligned} G &= (K \ltimes \mathbb{R}^{n+1})(\exp(\mathfrak{p}_0 \ltimes \mathbb{R}^{n+1})) \\ &= (K \ltimes \mathbb{R}^{n+1})(\exp(\mathfrak{p}_0) \ltimes \mathbb{R}^{n+1}) \\ &= (K \exp(\mathfrak{p}_0)) \ltimes \mathbb{R}^{n+1} \end{aligned}$$

with the last equality from simple matrix multiplication.

For the other cases ($\sigma \neq 0, \infty$), we do have to compute the normalizer. In order to do this, Anker and Ziegler define two "inner products" on $\mathbb{R}^{n+1}$. Instead of computing the normalizer of $\mathfrak{g}$, we will compute the normalizer of $\mathfrak{g}/\mathbb{R}^{n+1}$, the quotient of $\mathfrak{g}$ by $\mathbb{R}^{n+1}$. Essentially what this means is we "chop off" the rightmost and bottommost columns of any matrix in $\mathfrak{g}$ and say that any two matrices that look the same after being shaved down are actually the same. Our new normalizer is therefore

$$N(\mathfrak{g}/\mathbb{R}^{n+1}) = \{X \in \mathrm{GL}(n+1, \mathbb{R}) \mid XgX^{-1} \in \mathfrak{g}/\mathbb{R}^{n+1} \text{ for all } g \in \mathfrak{g}/\mathbb{R}^{n+1}\}.$$

Proposition 5.20. *Let $H \ltimes \mathbb{R}^n \subseteq GL(n, \mathbb{R}) \ltimes \mathbb{R}^n$ be a Lie group with Lie algebra $\mathfrak{h} \ltimes \mathbb{R}^n$. Then $N(\mathfrak{h}) \ltimes \mathbb{R}^n$ is isomorphic to $N(\mathfrak{h} \ltimes \mathbb{R}^n)$.*

Proof. Let $X \in N(\mathfrak{h} \ltimes \mathbb{R}^n)$. Then we can write

$$X = \begin{pmatrix} X_0 & \vec{v} \\ \vec{0}^\intercal & 1 \end{pmatrix} = (X_0, \vec{v})$$

for $v \in \mathbb{R}^n$ and some $n \times n$ matrix X_0. What should this matrix look like? By definition, $X \in N(\mathfrak{h} \ltimes \mathbb{R}^n)$ if and only if $XhX^{-1} \in \mathfrak{h} \ltimes \mathbb{R}^n$ for all $h \in \mathfrak{h} \ltimes \mathbb{R}^n$. This means that

$$\begin{pmatrix} X_0 & \vec{v} \\ \vec{0}^\intercal & 1 \end{pmatrix} \begin{pmatrix} h_0 & \vec{x} \\ \vec{0}^\intercal & 0 \end{pmatrix} \begin{pmatrix} X_0^{-1} & -X_0^{-1}\vec{v} \\ \vec{0}^\intercal & 1 \end{pmatrix} = \begin{pmatrix} X_0 h_0 X_0^{-1} & -X_0 h_0 X_0^{-1}\vec{v} + X_0\vec{x} \\ \vec{0}^\intercal & 0 \end{pmatrix} \in \mathfrak{h} \ltimes \mathbb{R}^n$$

for all $h = (h_0, \vec{x}) \in \mathfrak{h} \ltimes \mathbb{R}^n$. That is, $X_0 h_0 X_0^{-1} \in \mathfrak{h}$ for all $h \in \mathfrak{h}$. This is the exact condition for $X_0 \in N(\mathfrak{h})$. Thus the map given by

$$(X_0, \vec{v}) \in N(\mathfrak{h} \ltimes \mathbb{R}^n) \mapsto (Y_0, \vec{w}) \in N(\mathfrak{h}) \ltimes \mathbb{R}^n$$

is a bijection. Moreover, it is an isomorphism because both groups have matrix multiplication for an operation. $\square$

The first inner product is denoted $\langle \cdot, \cdot \rangle_+$ and is defined by

$$\langle \vec{x}, \vec{y} \rangle_+ = \vec{x}^\intercal \eta_+ \vec{y}, \text{ where } \eta_+ = \begin{pmatrix} -\sigma I_n & \vec{0} \\ \vec{0}^\intercal & 1 \end{pmatrix}$$

for $x, y \in \mathbb{R}^{n+1} \times \{0\}$.

Remark 5.21. When $\sigma > 0$, the bilinear operator $\langle \cdot, \cdot \rangle_+$ (resp. $\langle \cdot, \cdot \rangle_+$ when $\sigma < 0$) is not really an inner product as it is not positive definite. In Section 6 we discuss the interpretation of $\langle \cdot, \cdot \rangle_+$ as a sort of "signed metric" on Minkowski spacetime.

The second inner product $\langle \cdot, \cdot \rangle_-$ is similar:

$$\langle \vec{x}, \vec{y} \rangle_- = \vec{x}^\intercal \eta_- \vec{y}, \text{ where } \eta_- = \begin{pmatrix} \sigma I_n & \vec{0} \\ \vec{0}^\intercal & 1 \end{pmatrix}.$$

There are certain restrictions on the spaces on which these kinds of metrics take positive values.

Lemma 5.22. *Let $\{v_i\}_{i=1}^{n+1}$ be an orthogonal basis of $\mathbb{R}^{n+1}$. Then for $\sigma > 0$, (resp. $\sigma < 0$), the value of $\langle v_i, v_i \rangle_+$ (resp. $\langle v_i, v_i \rangle_-$) is greater than zero for exactly one of the v_i.* [8]

By definition, the adjoint of a transformation $z \in \mathfrak{gl}(n+1, \mathbb{R})$ under each of these inner products (denoted z^+ and z^- respectively) is given by $z^{\pm} = \eta_{\pm}^{-1} z^{\intercal} \eta_{\pm}$ so that for

$$z = \begin{pmatrix} A & \vec{b} \\ \vec{c}^{\intercal} & d \end{pmatrix}, \quad z^+ = \begin{pmatrix} A^{\intercal} & -\sigma^{-1}\vec{c} \\ -\sigma\vec{b}^{\intercal} & d \end{pmatrix} \text{ and } z^- = \begin{pmatrix} A^{\intercal} & \sigma^{-1}\vec{c} \\ \sigma\vec{b}^{\intercal} & d \end{pmatrix}.$$

The adjoints under the inner products give a sort of new notion of symmetry and anti-symmetry, so that we rewrite our Lie algebras as sets of matrices whose adjoints satisfy these “symmetry” requirements:

$$\begin{aligned} \mathfrak{g}/\mathbb{R}^{n+1} &= \{z \in \mathfrak{gl}(n+1, \mathbb{R}) \mid z^+ = -z\} \\ \mathfrak{p}_\sigma &= \{z \in \mathfrak{gl}(n+1, \mathbb{R}) \mid z^+ = -z, z^- = z\} \\ \mathfrak{k} &= \{z \in \mathfrak{gl}(n+1, \mathbb{R}) \mid z^{\pm} = -z\} \end{aligned}$$

and the group K as

$$K = \{k \in \mathfrak{gl}(n+1, \mathbb{R}) \mid k^{\pm} k = I_{n+1}\}.$$

Note that the condition on the k requires that they are also invertible, so K is in fact contained in $\mathrm{GL}(n+1, \mathbb{R})$.

Our first computation of the normalizer is to show that

$$N(\mathfrak{g}/\mathbb{R}^{n+1}) = \{A \in \mathrm{GL}(n+1, \mathbb{R}) \mid A^+ A = \lambda I_{n+1} \text{ for some } \lambda \in \mathbb{R}_{>0}\}.$$

Using the new expression for $\mathfrak{g}/\mathbb{R}^{n+1}$ and the usual rules about adjoints of products, we can rewrite $N(\mathfrak{g}/\mathbb{R}^{n+1})$ as

$$\begin{aligned} N(\mathfrak{g}/\mathbb{R}^{n+1}) &= \{A \in \mathrm{GL}(n+1, \mathbb{R}) \mid AzA^{-1} \in \mathfrak{g}/\mathbb{R}^{n+1} \text{ for all } z \in \mathfrak{g}/\mathbb{R}^{n+1}\} \\ &= \{A \in \mathrm{GL}(n+1, \mathbb{R}) \mid (AzA^{-1})^+ = -AzA^{-1} \text{ for all } z \in \mathfrak{g}/\mathbb{R}^{n+1}\} \\ &= \{A \in \mathrm{GL}(n+1, \mathbb{R}) \mid (A^+)^{-1} z A^+ = AzA^{-1} \text{ for all } z \in \mathfrak{g}/\mathbb{R}^{n+1}\} \\ &= \{A \in \mathrm{GL}(n+1, \mathbb{R}) \mid [A^+ A, z] = 0 \text{ for all } z \in \mathfrak{g}/\mathbb{R}^{n+1}\} \end{aligned}$$

By definition, the condition that some bracket equals zero is equivalent to requiring that the two matrices commute with each other. Since scalar multiples of the identity matrix commute with every matrix, we definitely have that

$$\{A \in \mathrm{GL}(n+1, \mathbb{R}) \mid A^+ A = \lambda I_{n+1} \text{ for some } \lambda \in \mathbb{R}_{>0}\} \subseteq N(\mathfrak{g}/\mathbb{R}^{n+1})$$

in its latest iteration. Now suppose that $A^+ A \in \mathrm{GL}(n+1, \mathbb{R})$ commutes with every $z \in \mathfrak{p}_\sigma$:

$$\begin{aligned} \left[\begin{pmatrix} B & \vec{u} \\ \vec{w}^{\intercal} & x \end{pmatrix}, \begin{pmatrix} 0_n & \vec{b} \\ \sigma\vec{b}^{\intercal} & 0 \end{pmatrix} \right] &= \begin{pmatrix} B & \vec{u} \\ \vec{w}^{\intercal} & x \end{pmatrix} \begin{pmatrix} 0_n & \vec{b} \\ \sigma\vec{b}^{\intercal} & 0 \end{pmatrix} - \begin{pmatrix} 0_n & \vec{b} \\ \sigma\vec{b}^{\intercal} & 0 \end{pmatrix} \begin{pmatrix} B & \vec{u} \\ \vec{w}^{\intercal} & x \end{pmatrix} \\ &= \begin{pmatrix} \sigma\vec{b}^{\intercal}\vec{u} - \vec{b}\vec{w}^{\intercal} & B\vec{b} - x\vec{b} \\ \sigma x\vec{b}^{\intercal} - \sigma\vec{b}^{\intercal} B & \vec{w}^{\intercal}\vec{b} - \sigma\vec{b}^{\intercal}\vec{u} \end{pmatrix} = \begin{pmatrix} 0_n & \vec{0} \\ \vec{0}^{\intercal} & 0 \end{pmatrix} \end{aligned}$$

for all vectors $\vec{b} \in \mathbb{R}^n$. By analyzing either the top right or bottom left entry, we can see that B must be some scalar multiple of the $n \times n$ identity matrix. In fact, $B = xI_n$.From the bottom right entry, we must have $\vec{w} = \sigma\vec{u}$. We can substitute this last result into the top left entry to see that $\vec{u}\vec{b}^\intercal - \vec{b}\vec{u}^\intercal = 0$ for all $b \in \mathbb{R}^n$. The only vector for which this is true is $\vec{0}$, so both $\vec{w}$ and $\vec{v}$ must be zero. Thus $A^+A = xI_{n+1}$. We also have that $x > 0$: in the case that $\sigma < 0$, $\langle\cdot,\cdot\rangle_+$ is a true inner product, so

$$\langle A\vec{v}, A\vec{v}\rangle_+ = \langle \vec{v}, A^+A\vec{v}\rangle_+ = x(-\sigma\vec{v}_-^2 + v_{n+1}^2) > 0$$

and therefore $x > 0$ when $\vec{v} \neq 0$. Now let $\sigma > 0$ and suppose that $x < 0$. Then every element of the subspace $\{\vec{r} = (\vec{r}_-, 0)^\intercal \mid \vec{r}_- \in \mathbb{R}^{n+1}\} \subset \mathbb{R}^{n+1}$ has positive squared "norm" under a transformation $A \in N(\mathfrak{g}/\mathbb{R}^{n+1})$:

$$\langle A\vec{r}, A\vec{r}\rangle_+ = \langle \vec{r}, A^+A\vec{r}\rangle_+ = -\sigma\lambda\langle\vec{r},\vec{r}\rangle \geq 0$$

where the last inner product is the usual Euclidean one. The signs of σ and λ guarantee the inequality. This space has dimension n because all transformations A are invertible, but by Lemma 5.22, a subspace on which the squared "norm" given by $\langle\cdot,\cdot\rangle_+$ is positive can have at most dimension 1.

This is sufficient to show that

$$N(\mathfrak{g}/\mathbb{R}^{n+1}) \subseteq \{A \in \mathrm{GL}(n+1,\mathbb{R}) \mid A^+A = \lambda I_{n+1} \text{ for some } \lambda \in \mathbb{R}_{>0}\}$$

so that the two sets are actually equal because if some matrix A is such that A^+A does commute with every element of $\mathfrak{g}/\mathbb{R}^{n+1}$, it must commute with those elements that are contained in $\mathfrak{p}_\sigma$ and therefore must take the diagonal form described above.

The cases $\sigma > 0$ and $\sigma < 0$ are similar, so we prove Theorem 3.3 in the case $\sigma > 0$ because it happens to be the case we observe in the universe. We will show that we can rewrite the normalizer yet again as

$$N(\mathfrak{g}/\mathbb{R}^{n+1}) = \{A \in \mathrm{GL}(n+1,\mathbb{R}) \mid A = \sqrt{\lambda}Be^z \text{ for some } \lambda \in \mathbb{R}_{>0}, B \in K, z \in \mathfrak{p}_\sigma\}.$$

This is sufficient to prove Theorem 3.3: suppose that some $A = \lambda Be^z$ as in the set described above is also contained in $G/\mathbb{R}^{n+1}$. Then because $K\exp(\mathfrak{p}_\sigma) \subseteq G/\mathbb{R}^{n+1}$, we also have that $Ae^{-z}B^{-1} = \sqrt{\lambda}I_{n+1}$ is in $G/\mathbb{R}^{n+1}$. Thus by hypothesis of Theorem 3.3, considering the bottom-right entry of a matrix in $G/\mathbb{R}^{n+1} \cap \mathrm{GL}(n+1,\mathbb{R})$ yields $\lambda = 1$ so that the other inclusion holds:

$$G/\mathbb{R}^{n+1} \subseteq K\exp(\mathfrak{p}_\sigma)$$

and re-attaching the semidirect product gives the desired result.

From our expressions of the Lie algebras in terms of adjoints of the inner products $\langle\cdot,\cdot\rangle_\pm$ and the fact that the adjoint of an exponential is the exponential of the adjoint, we have that for a matrix $A = \sqrt{\lambda}Be^z$, for $B \in K$ and $z \in \mathfrak{p}_\sigma$,

$$\begin{aligned} A^-A &= (\sqrt{\lambda}Be^z)^-(\sqrt{\lambda}Be^z) \\ &= \lambda e^{z^-}B^-Be^z \\ &= \lambda e^{2z}. \end{aligned}$$

Then solving for z and B we have

$$z = \frac{1}{2}\ln\left(\frac{1}{\lambda}A^-A\right) \text{ and } B = \frac{1}{\sqrt{\lambda}}Ae^{-z}.$$

Suppose that we can write every element A of our most recent confirmed expression of the normalizer like this. That is,

$$N(\mathfrak{g}/\mathbb{R}^{n+1}) \subseteq \{A \in \mathrm{GL}(n+1,\mathbb{R}) \mid A = \sqrt{\lambda} B e^z \text{ for some } \lambda \in \mathbb{R}_{>0}, B \in K, z \in \mathfrak{p}_\sigma\}.$$

We must check whether this z is well defined. In fact it is: the inner product $\langle \cdot, \cdot \rangle_-$ is a true inner product (it is positive definite) and therefore the matrix $p = \frac{1}{\lambda} A^- A$ is also positive definite because

$$\langle A^- A\vec{v}, \vec{v}\rangle_- = \langle A\vec{v}, A\vec{v}\rangle_- > 0$$

for nonzero $\vec{v} \in \mathbb{R}^{n+1}$. Thus the logarithm is well-defined and also positive definite.

Now we compute the product of the adjoint

$$\begin{aligned} p^+ p &= \left(\frac{1}{\lambda} A^- A\right)^+ \left(\frac{1}{\lambda} A^- A\right) \\ &= \frac{1}{\lambda^2} A^+ (A^-)^+ A^- A \\ &= \frac{1}{\lambda^2} A^+ (\lambda I_{n+1}) A \\ &= \frac{1}{\lambda} A^+ A \\ &= I_{n+1} \end{aligned}$$

so that $p^+ = p^{-1}$. From this we can see that in fact $z \in \mathfrak{p}_\sigma$: because z is positive definite, it is self-adjoint under the “true” inner product $\langle \cdot, \cdot \rangle_-$, so $z^- = z$ and the first condition is satisfied. To see that $z^+ = -z$, note that by definition, $z = \frac{1}{2}\ln(p)$. Then applying the + adjoint and the usual logarithm rules,

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

Now to see that $B \in K$, we compute

$$B^+ B = \left(\frac{1}{\sqrt{\lambda}} A e^{-z}\right)^+ \left(\frac{1}{\sqrt{\lambda}} A e^{-z}\right) = \frac{1}{\lambda} e^{-z^+} A^+ A e^{-z} = e^z e^{-z} = I_{n+1}$$

$$B^- B = \left(\frac{1}{\sqrt{\lambda}} A e^{-z}\right)^- \left(\frac{1}{\sqrt{\lambda}} A e^{-z}\right) = \frac{1}{\lambda} e^{-z^-} A^- A e^{-z} = e^{-z} e^{2z} e^{-z} = I_{n+1}.$$

Together, this all shows that

$$N(\mathfrak{g}/\mathbb{R}^{n+1}) \subseteq \{A \in \mathrm{GL}(n+1,\mathbb{R}) \mid A = \sqrt{\lambda} B e^z \text{ for some } \lambda \in \mathbb{R}_{>0}, B \in K, z \in \mathfrak{p}_\sigma\}.$$

The other inclusion is clear: if $A = \sqrt{\lambda} B e^z$ as described, then

$$A^+ A = \left(\sqrt{\lambda} B e^z\right)^+ \left(\sqrt{\lambda} B e^z\right) = \lambda e^{z^+} B^+ B e^z = \lambda e^{-z} e^z = \lambda I_{n+1}.$$

6. Discussion of Result

First, we will restrict ourselves to the case $n = 3$ because we live in three spatial dimensions. Theorem 3.3 then affords five possibilities for the symmetry group of Minkowski spacetime: $G = K \ltimes \mathbb{R}^4$ or $G = K \exp(\mathfrak{p}_\sigma) \ltimes \mathbb{R}^4$ for one choice of $\sigma = 0$, $\sigma = \infty$, $\sigma < 0$ or $\sigma > 0$. The one that agrees with experiment is the case $\sigma > 0$, corresponding to the observation that the speed of light c is positive but finite. Moreover, we proved that there can be only one value of σ, which implies that the speed of light is an invariant for all observers.

The form $\langle \cdot, \cdot \rangle_+$ for $\sigma > 0$ corresponds to the Minkowski metric of spacetime. If we let the speed of light $c = \frac{1}{\sqrt{\sigma}}$, we can measure the spacetime interval Δs^2 between an event (t, x, y, z) and the origin (usually the location of an observer) by

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2.$$

The sign of this interval is important. If it is negative, we say that the event has **spacelike** separation from the origin. That is, not even light can "get to" that point before the event occurs. If the interval is zero, the event has **lightlike** separation from the origin. This means that light leaving the origin at time will arrive at the spatial point (x, y, z) at exactly time t. Finally, if the interval is positive, the event has **timelike** separation from the origin. This means that light leaving the origin will arrive at the location before the event occurs, and therefore the origin and the event are **causally connected**. The transformations in the Poincaré group are exactly those that *preserve* the interval Δs^2. One interesting consquence of this is that the order of two events x and y that are causally connected (i.e. Δs^2 for $\vec{x} - \vec{y} > 0$) will stay the same for every observer, while two events that are not causally connected may switch orderings in time depending on who is observing them.

Acknowledgments

First and foremost, I would like to thank my mentor, Thomas Hameister, for helping me to trudge through the physics and weave it in with the mathematics. Thank you to my parents, Nancy Poletti and Robert Horowitz, for supporting my mathematical interests from the very beginning and for letting me practice presentations and proofs on them. Finally, thank you to the Ryerson Refugees (Geoffrey Baring, Elle Bower, Alex Ghorbani, John Hopper, Chen Lin, Christian Liu, John Naughton, Matthew Niemiro, Eha Srivastava, and Tanner Struck) for making virtual lectures feel real and turning a cold classroom into a fun place to freeze!

References

[1] Anker, Jean-Phillippe & Ziegler, François (2020) *Relativity without light: A new proof of Ignatowski's theorem.*

[2] Bourbaki, Nicholas. *Groupes et algèbres de Lie. Chapitre III: Groupes de Lie.* Hermann, Paris 1972

[3] Einstein, Albert (1905) *Zur Elektrodynamik bewegter Körper.* Annalen der Physik, 17, 891.

[4] Gallier, Jean (2015) UPenn CIS 610. Chapter 9: Lie Groups, Lie Algebras and the Exponential Map. 2018

[5] Gorini, Vittorio (1971) *Linear kinematical groups.* Comm. Math. Phys **21** 150-163

[6] Hall, Brian C. *Lie Groups, Lie Algebras, and Representations: An elementary introduction.* Springer 2015.

[7] Ignatowski, Wladimir Sergejewistch (1910) *Das Relativitätsprinzip.* Archiv der Mathematik und Physik, 17, Band (1910), S. 1-24.

[8] Lang, Serge. *Algebra.* Springer 2002