Time Arrow Spinors: Lorentz Transformations for the Modified Cosmological Model

Jonathan W. Tooker

January 23, 2024

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Contents

§1 Purpose and Scope

In [1], Steane says that there are two ways of thinking about a spinor: a two-component vector with complex entries, or a null 4-vector with real entries. However, we want to examine what happens if we convert to the Euclidean metric with complex 4-vectors. Also, if $\chi^4_\pm$ is oppositely timelike and spacelike in $\Sigma^\pm$, then when we assemble 4-vectors from x^i and either of $x^0, \chi^4_\pm$, there are only null vectors in one variant of $\chi^4_\pm$. Therefore, we need to work out all of the mechanics for the conventions on real and complex 4-vectors. Carroll's treatment of 4-vectors in spacetime in [2] requires real-valued 4-vectors, so we will examine the details.

Spinors have different x-form properties under Lorentz transformations, so we need to examine what they are.

The Lorentz invariant associated with a 4-vector x^μ (where the sum $x^\mu \hat{e}_\mu$ is implied) is the inner product $x^\mu x_\mu$. This is called the "spacetime interval" or the "Minkowski length squared." (CHECK ON SQUARED???) For a 4-vector in $\mathbb{R}^4$,

$$x^\mu = (t, x, y, z) \quad ,$$

the inner product is written

$$\big(x^\mu\big)^2 = x^\mu x_\mu = x^\mu \big(g_{\mu\nu} x^\nu\big) \quad .$$

We have the freedom to write the metric in one of two sign conventions:

$$\eta_{\mu\nu} = \text{diag}(\pm 1, \mp 1, \mp 1, \mp 1)$$

so the inner product becomes

$$\begin{aligned}\big(x^\mu\big)^2 &= (t, x, y, z) \cdot (\pm t, \mp x, \mp y, \mp z) \\ &= \pm \big(t\big)^2 \mp \sum_i \big(x^i\big)^2\end{aligned}$$

As these are oppositely signed, we will have a null interval.

We can mention Zeeman here and the fine topology. Suggest a new real-complex biplane topology $\mathbb{C} \times \mathbb{R}^2$. Need to derive the property or orientability from the topology. Need to relate to ADM.

§1.1 List of Things I Want to Explain

- Why is one direction preferred for angular momentum eigenspinors? If we construct a time-orientable manifold as $\mathbb{C} \times \mathbb{R}^2$, and the time-orientability comes from the uniqueness of t as the imaginary part of $\mathbb{C}$, isn't one spatial direction favored for orientation by being equally unique in $\mathbb{R}$?

- What is the connection between chronological and chirological time states maybe having 2 or three basis states: vectorial and spinorial, or all spinorial? We have an application toward violation of conservation of information when going between bases with unequal cardinality. Could a boost be used make the connection since it identifies a 2D perpendicular to the boost direction.

- Superluminal boosts are imaginary. What convention can we we impose to let superluminally boosted $V^{\mu} \in \mathbb{R}^{1,3}$ always have the right phase?

- When $|\boldsymbol{v}| = c$, $\gamma = \widehat{\infty}$. This also has a distinction about a plane perpendicular to the boost direction not going to the higher level of aleph.

- The ADM theorem requires the cosmological principle, but that is ruled out by modern data. How can we get one direction needed for the symplectic form at infinity? It is weird that the CMB seems divided by the plane of the solar system, so maybe we can associate the angular momentum of the solar system with a favored spatial direction like we do in QM. Uzer had some nice paper where he showed some atomic limit cycle was like an asteroid's gravitational limit cycle, and we should look at what he said about the likeness between the solar system and the atomic system.

  URL:

-

§2 Main Results

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

§3 The Manifold of Special Relativity

A useful result in group theory is that certain groups have an associated manifold. The Lorentz group is such a group, an example of what is called a Lie group[1], and the associated manifold is what physicists call Minkowski space, which is *the manifold of special relativity.* Unlike curved spaces in the general theory of relativity, the special theory is confined to flat Lorentzian 4-space: Minkowski space. The flat Lorentzian metric (called the Minkowski metric) may be chosen in signature $\{+---\}$ or $\{-+++\}$ so that the line element is

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 \quad , \quad \text{or} \qquad ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 \quad .$$

This freedom to choose a sign is reflected in the convention to label the Lorentz group as O(3,1) or O(1,3).[2] The letter O indicates *the orthogonal group*, and the points in the associated manifold are specified with (*i*) three real numbers and one imaginary, or (*ii*) three imaginary and one real. However, the current trend in physics is to label points in Minkowski space (called events) with $x^\mu \in \mathbb{R}^4$, and then obtain the manifold of special relativity by defining distance between special relativistic events with the flat Lorentzian metric

$$\eta_{\mu\nu} = \begin{bmatrix} \pm c^2 & 0 & 0 & 0 \\ 0 & \mp & 0 & 0 \\ 0 & 0 & \mp & 0 \\ 0 & 0 & 0 & \mp 1 \end{bmatrix} \qquad \Longrightarrow \qquad ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu \quad .$$

Since $\mathbb{R}^4$ is the manifold of O(4), the manifold of four orthogonal we must associate vectors in $\mathbb{R}^4$ with the Lorentz group by introducing complexity with a non-Euclidean metric: distance in Lorentzian 4-space is not a positive-definite quantity like it is in Euclidean 3-space, and we call it *the spacetime interval* to highlight this distinction.

The main point of this paper will be to put away $\mathbb{R}^4$ and examine the natural representation of the Lorentz group: three real numbers and one imaginary, or vice versa. For example, the 2D real vector space $\mathbb{R}^2$ and the 1D complex vector space $\mathbb{C}^1$ are isomorphic (the same), but we gain a lot of powerful theorems for holomorphic functions by working in $\mathbb{C}$. Here, we will explore what might be gained similarly with a complex representation of the Lorentz group. In the vector space of the natural representation of the Lorentz group, call it $\mathcal{M}$ with definition

$$\mathcal{M} = \begin{cases} i\mathbb{R} \times \mathbb{R}^3 & \text{if} \quad \text{O(3,1)} \\ \mathbb{R} \times i\mathbb{R}^3 & \text{if} \quad \text{O(1,3)} \end{cases} \qquad ,$$

[1] The reader is referred to Appendix 2 in Roman [3] for an excellent and concise, 30-page statement of everything 90% of physicists will ever need to know about group theory.

[2] One often distinguishes the proper and restricted Lorentz groups from the full Lorentz group so that the freedom to choose a sign in the metric is reflected in a choice between SO(3,1) and SO(1,3) or $SO_+(3,1)$ or $SO_+(3,1)$. The Lorentz subgroups are distinguished in Section XXX.

the metric will be Euclidean, and all of the magic usually assigned to the Lorentzian metric signature must be reassigned elsewhere. As transformations on an other-than-usual vector space, the elements of the Lorentz group, the Lorentz transformations themselves, will not have their usual analytical forms when written as matrices. These matrices will be developed in Section 4.5.

Every finite-dimensional vector space is a manifold, and $\mathbb{R}^4$ and $\mathcal{M}$ are the same vector space. However, the manifolds associated with matrix groups O(3,1) and O(4) are not the same, we need to state the nuance by which we can say that vectors in $\mathbb{R}^4$ and $\mathcal{M}$ are both representations of the Lorentz group. This will require the introduction of (pseudo-)Riemannian manifolds, which are manifolds equipped with a metric that is not necessarily the Euclidean one. First, we will present a synopsis of the basic physics described by 4-vectors in Minkowski space to establish the context for Lorentz transformations. Then we will introduce Riemannian geometry and establish a context in the Modified Cosmological Model.

§3.1 Special Relativity

Since observers in different reference frames ought to be able to agree on the facts of objective phenomena in their joint experience, we are able to obtain tight constraints on what mathematical forms our physical theories might take. For two frames to be non-trivially different—as opposed to the case of two different observers observing things from different places in the same frame—one frame must be in motion *relative* to the other, and we arrive at the concept of *relativity*. Of particular interest is the case in which the origins of coordinates in each reference frame are separated by a time-varying displacement vector: one frame is moving with velocity $\boldsymbol{v}(t)$ relative to the other. To simplify things, we often consider the case when $\boldsymbol{v}$ is constant and the relative orientation of the coordinate axes in each reference frame is fixed; there is no acceleration or rotation between frames: one is simply and steadily moving relative to the other. In the so-called *standard configuration*, we choose things to be such that the two coordinate systems were at the same place at $t = 0$. Since $\boldsymbol{v}$ is a constant, we will define the coordinates of 3-space to be such that $\boldsymbol{v} = v\hat{x}$: the constant velocity is purely in the x-direction. Taking frame S with coordinates (t, x, y, z) as fixed while frame S' with coordinates (t', x', y', z') moves along the x-axis, the coordinates of an event with a time and place observed from S' might be described as

$$t' = t \qquad\qquad x' = x - vt \qquad\qquad y' = y \qquad\qquad z' = z \quad .$$

If a bird is flying away from S with the same speed and direction as S', then its x position at time t is, from Newtonian dynamics, $x(t) = vt$, and we obtain $x'(t') = 0$. The bird remains stationary at the origin in S' for any value of t' because it is co-moving with S'. This formulation reflects what is called *Galilean relativity*. It is a good approximation at low speeds, which are called *non-relativistic*. If the bird becomes a photon, then it will be moving at the speed of light relative to both S and S', and it cannot possibly be stationary in either. Clearly, the Galilean theory of relativity will not comply with the requirement that different observers' theories must agree on objective facts. Therefore, a *special* theory of relativity is needed when the bird flies very quickly. Loosely following Tipler and Llewellyn [4], we may derive the correct relativistic transformation as follows.

Assume the special relativistic transformation of position is

$$x' = \gamma\big(x - vt\big) \quad , \quad \text{where} \qquad \gamma = \gamma(v) \quad .$$

This must reduce to the Galilean expression as $v \to 0$, and it will yield the correct transformation otherwise. Since our theory must be consistent if S' is moving with speed v relative to S or if S is moving with speed $-v$ relative to stationary frame S', we obtain the equivalent expression

$$x = \gamma\big(x' + vt'\big) \quad .$$

By inserting the expression for x' into the one for x, we obtain

$$t' = \gamma\left[\frac{x\big(1-\gamma^2\big)}{\gamma^2 v} + t\right] \quad .$$

Since our bird has become a photon moving with velocity c in both the S and S' frames, irrespective of the magnitude of their relative velocities (*!!!*), the position of the photon will be described as

$$x(t) = ct \quad , \quad \text{and} \qquad x'(t') = ct' \quad .$$

From the former, we obtain the constraint $x/t = c$, and into the latter we insert our expressions for x' and t':

$$\gamma\big(x - vt\big) = c\gamma\left[\frac{x\big(1-\gamma^2\big)}{\gamma^2 v} + t\right] \quad .$$

Multiplying by $\gamma v t^{-1}$ and using $x/t = c$ yields

$$\gamma^2 v\big(c - v\big) = c^2\big(1-\gamma^2\big) + \gamma^2 vc \quad .$$

Canceling like terms and dividing by c^2 yields an expression which is easily solved to obtain

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

Since Lorentz had derived this transformation before Einstein found its application in the physics of simultaneity, the resultant expressions

$$x' = \gamma\big(x - vt\big) \quad , \quad \text{and} \qquad t' = \gamma\left(t - \frac{vx}{c^2}\right) \quad ,$$

reflect what is called "a rotation-free Lorentz transformation." More concisely, such transformations are called *boosts*. Using $\beta = v/c$, such transformations are concisely written as matrices pre-multiplying column 4-vectors:

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \gamma\left(ct - \beta x\right) \\ \gamma\big(x - vt\big) \\ y \\ z \end{bmatrix} \quad .$$

Once spatial rotations are added and we allow arbitrarily directed boost velocities $\boldsymbol{v} = v\hat{n}$, and once we add things like parity and time-reversal, the set of all such matrix transformations forms "the Lorentz group." Rather than studying the group as a whole, the present work regards a particular representation of the group in the MCM: complex 4-vectors $V^\mu \in M$ which will be developed in Section 4.5.

Definition 3.1.1 The Lorentz factor is

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \quad , \quad \text{where} \qquad \beta = \frac{v}{c}$$

Definition 3.1.2 The flat Lorentzian metric is called the **Minkowski metric**. It is a rank-(0,2) tensor

$$\eta_{\mu\nu} = \text{diag}(\pm 1, \mp 1, \mp 1, \mp 1) \quad .$$

The variants $\eta^\pm_{\mu\nu}$ will be used to specify one sign convention or the other with $\pm$ corresponding to the sign of η_{00}, which is the first diagonal entry:

$$\eta^+_{\mu\nu} = \text{diag}(1, -1, -1, -1) \quad , \quad \text{and} \qquad \eta^-_{\mu\nu} = \text{diag}(-1, 1, 1, 1) \quad .$$

Definition 3.1.3 The symbol $g_{\mu\nu}$ denotes an arbitrary metric that may or may not be the Minkowski metric $\eta_{\mu\nu}$.

Definition 3.1.4 The **Minkowski square** of a 4-vector V^μ is its contraction with itself, or its double contraction with the metric:

$$V_M^2 = V^\mu V_\mu = g_{\mu\nu} V^\mu V^\nu \quad .$$

This quantity is also called the **Lorentz norm** and denoted $\|\boldsymbol{V}\|_L$.

Remark 3.1.5 In the usual conventions for special relativity, the 4-vector in question is $V^\mu \in \mathbb{R}^4$, and the metric is the Minkowski metric $\eta_{\mu\nu}$. In that case, the Minkowski square is

$$V_M^2 = \eta_{\mu\nu} V^\mu V^\nu = \pm\big(V^0\big)^2 \mp \Big[\big(V^1\big)^2 + \big(V^2\big)^2 + \big(V^3\big)^2\Big] \quad .$$

However, this will not be the formula for V_M^2 when we represent the Lorentz group with $V^\mu \in \mathcal{M}$ because the metric in that representation is not $\eta_{\mu\nu}$. Such distinctions are made in Section 3.3.

Definition 3.1.6 An **event** is something that happens at a time and place. Events in spacetime are uniquely specified with 4-vectors whose first component specifies a time t and whose other three components specify a position vector $\boldsymbol{x}$ in Euclidean 3-space.

Definition 3.1.7 The **spacetime interval** $(\Delta s)^2$ between two events in Minkowski space is the Minkowski square of their separation vector. What distance is to space, spacetime interval is to spacetime. Given a separation 4-vector $x^\mu = b^\mu - a^\mu$, the spacetime interval between events a^μ and b^μ is the Minkowski square of x^μ:

$$\left(\Delta s\right)^2 = x_M^2 = \eta_{\mu\nu} x^\mu x^\nu \quad .$$

Remark 3.1.8 Spacetime interval is the Minkowski square of position 4-vectors, or the Lorentz invariant associated with a position 4-vector, but other kinds of 4-vectors, e.g.: momentum 4-vectors or force 4-vectors, have Minkowski squares defining their own Lorentz invariants that are not spacetime intervals, as in Table 1. The pervasive presence of Lorentz invariants in every corner of the theory (due to the requirement that different observers must agree on objective facts) is part of the reason why the manifold of special relativity is said to be the manifold of the Lorentz group. Regarding position 4-vectors, relativity replaces the classical invariance of length with a Lorentz invariant spacetime interval.

Definition 3.1.9 If the spacetime interval between two events vanishes, meaning $(\Delta s)^2 = 0$, then the events are said to be **lightlike** separated, separated by a null interval, or separated by a lightlike interval. A general 4-vector, not necessarily a position vector, is lightlike if and only if its Minkowski square vanishes: $x_M^2 = 0$, c.f.: Definition 3.1.7. Lightlike separation between two events is a type of **causal** separation because correlations restricted by the speed of light can exist between them.

Definition 3.1.10 The separation vector between lightlike separated events is said to be on the **light cone**, which is the union of all lightlike separation vectors anchored at one event or the other. The future light cone of event x^μ at time t is the union of all lightlike 4-vectors anchored at x^μ and pointing to events with time $t' > t$, and the past light cone is the union of all such 4-vectors pointing to events with time $t' < t$. This arrangement is shown in Figure 1.

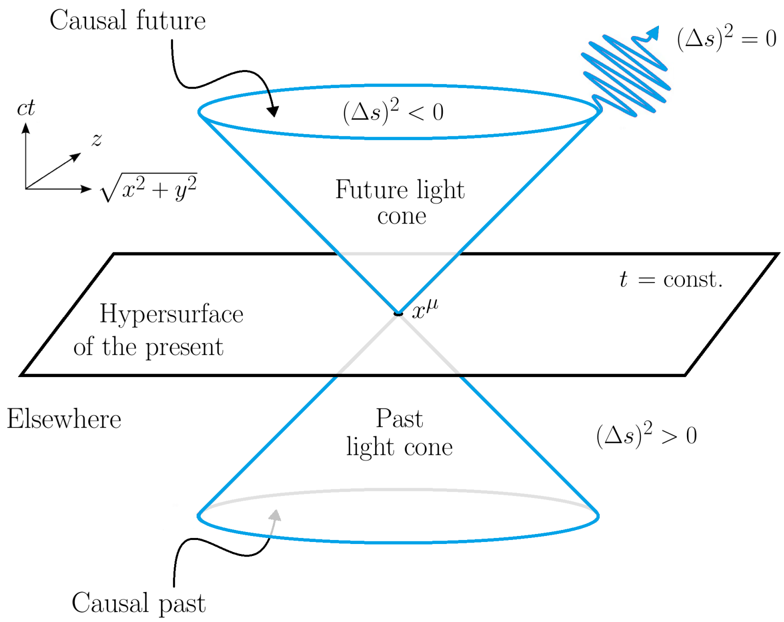


Figure 1: This figure shows the light cone at event x^μ in Minkowski space. When time is measured in units of ct, objects moving at the speed of light cross equal amounts of time and space in a given time (Theorem 3.1.15), so every lightlike interval is a 45 degree line in the x, ct-plane. $(\Delta s)^2 = 0$ along such intervals, regardless of the chosen conventions. Points on the interior of the light cone (past or future) are timelike separated from x^μ. This corresponds to $(\Delta s)^2 < 0$ in Lorentzian metric signature $\{-{+}{+}{+}\}$. If one event (point) is in the past or future light cone of another, they are said to be causally separated because more time (measured as ct) has elapsed than space (measured as x). Communication between timelike separated events does not require faster-than-light signals, and one might have caused the other. If two events are spacelike separated, then they are said to be acausally separated, or to lie elsewhere with respect to one another. Events on a hypersurface of constant proper time are not observable at any other event on that surface because the propagation of a signal at finite speeds must span some interval in the ct-direction. For something to observed at the same time it happened, the signal would have to propagate infinitely fast. Instead, every event that is observable at x^μ lies on the past light cone. This distinction in what it means for something to be happening "now" highlights a main difference between Newtonian dynamics and relativity. Classically, everything observable at time t is happening at proper time $\tau = t$, but in relativity an event must have happened at some earlier time $t < \tau$ if it is observable on the hypersurface of the present. Note that x^μ is an arbitrary event, and every event in spacetime has its own past and future light cones defining (i) a region where things might caused the event: the past light cone and its interior, and (ii) a region where things might be consequences of the event: the future light cone and its interior.

Symbol	Definition	Components	Name(s)	Invariant $\lvert V_M^2 \rvert$
x^μ	x^μ	$\big(ct, \boldsymbol{x}\big)$	event, position, interval displacement, separation	$\big(\Delta s\big)^2$
U^μ	$dx^\mu/d\tau$	$\big(\gamma c, \gamma \boldsymbol{v}\big)$	4-velocity	c^2
p^μ	$m_0 U^\mu$	$\big(E/c, \boldsymbol{p}\big)$	4-momentum, energy-momentum	$m_0^2 c^2$
f^μ	$dp^\mu/d\tau$	$\big(\gamma W/c, \gamma \boldsymbol{f}\big)$	4-force, work-force	—
j^μ	$\rho_0 U^\mu$	$\big(c\rho, \boldsymbol{j}\big)$	4-current	$c^2 \rho_0^2$
A^μ	A^μ	$\big(\phi/c, \boldsymbol{A}\big)$	4-potential	—
a^μ	$dU^\mu/d\tau$	$\big(\gamma\dot{\gamma}c, \gamma\dot{\gamma}\boldsymbol{v} + \gamma^2 \boldsymbol{a}\big)$	4-acceleration	a_0^2
k^μ	$\Box\psi$	$\big(\omega/c, \boldsymbol{k}\big)$	4D wave vector	—

Table 1: This table is adapted from Steane [5]. Quantities such as the 4-potential squared $A^\mu A_\mu$ are not Lorentz invariant, and A^μ not called a Lorentz vector. When dealing with such quantities in practice, one constructs other objects such as the electromagnetic field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_n A_\mu$ that does have a Lorentz invariant $F_{\mu\nu}F^{\mu\nu}$.

Definition 3.1.11 If motion from one event to another does not require any speeds greater than c, then the events are said to be **timelike** separated, separated by a timelike interval, or **causally** separated. For two timelike events, one is in the future lightcone of the other, and the former is in the past lightcone of the latter. Events that are not timelike or lightlike separated are **spacelike** separated, separated by a spacelike interval, or **acausally** separated. Such events are said to lie **elsewhere** with respect to one another. Depending on the metric sign convention $\{\pm\mp\mp\mp\}$, a general 4-vector V^μ (not necessarily a position vector) is said to be timelike or spacelike depending on the sign of its Minkowski square:

$$V^\mu \text{ timelike:} \begin{cases} V_M^2 > 0 & \text{if} \quad \eta_{00} = 1 \\ V_M^2 < 0 & \text{if} \quad \eta_{00} = -1 \end{cases}$$

$$V^\mu \text{ spacelike:} \begin{cases} V_M^2 < 0 & \text{if} \quad \eta_{00} = 1 \\ V_M^2 > 0 & \text{if} \quad \eta_{00} = -1 \end{cases} .$$

A vector is said to be timelike if its Minkowski square agrees with the sign of the time part of the metric. If it is not timelike or null, it is spacelike.

Definition 3.1.12 A **Lorentz frame**, or an inertial frame, is a region in which spacetime is assumed to be locally flat: an observer's reference frame is one in which he appears unaccelerated and at rest. In practice, Lorentz frames do not exist because spacetime curvature exists everywhere—even gravity on the surface of the Earth is an acceleration—but special relativity assumes that all of spacetime may be described with a Lorentz frame. For two frames S and S' as in the introduction to this section, both are said to be Lorentz frames and Lorentz transformations relate the quantities that can be measured in either.

Definition 3.1.13 A **dimensional transposing parameter** is a dimensionful scalar constant that changes the units of one quantity to another by multiplication. The speed of light c is the dimensional transposing parameter between time and space:

$$\frac{[\text{length}]}{[\text{time}]} \times [\text{time}] = [\text{length}] \quad .$$

Example 3.1.14 Optional constants in the metric. Dimensionful units are irrelevant to the most general Lorentz invariant $g_{\mu\nu}V^\mu V^\mu$, but addition is only defined in physics among quantities with like dimensionality. If the Minkowski square of a dimensionful, physical 4-vector is going to be a well-defined scalar, the components summed in the Minkowski square must have the same units. In the introduction to this section, we developed a boost in the x-direction and wrote the transformation as a matrix operation:

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \gamma\,(ct - \beta x) \\ \gamma\big(x - vt\big) \\ y \\ z \end{bmatrix} \quad .$$

The form of the matrix clearly depends on the representation of the Lorentz group that we have chosen. Namely, we have taken $V^\mu \in \mathbb{R}^4$ to be a displacement vector, and we have used the dimensional transposing parameter c to affect a convention in which displacement in spacetime, called spacetime interval, has units of meters squared[3]:

$$x^0 = ct \qquad x^1 = x \qquad x^2 = y \qquad x^3 = z \quad .$$

Contraction with the Minkowski metric

$$\eta_{\mu\nu} = \text{diag}(\pm 1, \mp 1, \mp 1, \mp 1) \quad .$$

given in Definition 3.1.2 yields a well-defined Lorentz scalar (Lorentz invariant): the Minkowski square $x^\mu x_\mu$. However, we might not have put the dimensional transposing parameter into the vector and used an equivalent convention

$$x^0 = t \qquad x^1 = x \qquad x^2 = y \qquad x^3 = z \quad .$$

[3]It is also possible to use c^{-1} as the transposing parameter so that spacetime interval is said to have units of time squared.

If we use $\eta_{\mu\nu}$ such that $\eta_{00} = \pm 1$, then x_M^2 will contain a term $\pm t^2$ that does not have units of meters squared and cannot be added to terms such as z^2 that do have those units. To use $x^\mu = (t, x, y, z)$, well-definition of the Minkowski square requires that we put the transposing parameter into the metric

$$\eta'_{\mu\nu} = \text{diag}(\pm c^2, \mp 1, \mp 1, \mp 1) \quad .$$

In this convention, the Lorentz boost is written as

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\frac{v\gamma}{c^2} & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \gamma\left(t - \frac{vx}{c^2}\right) \\ \gamma\big(x - vt\big) \\ y \\ z \end{bmatrix} \quad .$$

We lose the symmetry of the transformation matrix, but spatial rotation matrices are not symmetric, and they belong to the Lorentz group as well, so we have not broken anything unbreakable. Since the $\eta_{\mu\nu}$ is its own inverse:

$$\eta_{\mu\nu}\eta^{\mu\nu} = \begin{bmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \mp 1 & 0 & 0 \\ 0 & 0 & \mp 1 & 0 \\ 0 & 0 & 0 & \mp 1 \end{bmatrix} \begin{bmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \mp 1 & 0 & 0 \\ 0 & 0 & \mp 1 & 0 \\ 0 & 0 & 0 & \mp 1 \end{bmatrix} = \mathbb{1}_4 \quad ,$$

and $\eta'_{\mu\nu}$ is not its own inverse, we usually work in the convention where the transposing parameter appears in the vector.

This example has demonstrated an inherent freedom in special relativity to choose different combinations of metrics and vectors as representations of the Lorentz group. We have seen that c^2 in η_{00} shows up as c when we move it into the 4-vector, and the reader is invited to observe the same quadratic relationship later when we take -1 out of η_{00} and put i into the 4-vector. Since the complex inner product contains a sign-inverting conjugation operation not present in the Minkowski square of a complex 4-vector, and we might associate this with the sign inversion of spinors under full spatial rotations, we will want to closely examine the case for i in the 4-vector as a non-trivial analogue of the present example.

================

imaginary transposing parameter

Theorem 3.1.15 *When time is expressed in units of length via multiplication by the dimensional transposing parameter c, an object moving at the speed of light will cross equal amounts of space and time during any natural number of small time increments.*

Proof. Suppose nt_0 is a natural number of small time increments. For $t = nt_0$, the distance formula $x(t) = vt$ gives $x(nt_0) = cnt_0$. This is the time nt_0 expressed in units of length.

Remark 3.1.16 When x^0 is measured in seconds and x^i is measured in meters, the light cone is very broad. In one second, light travels almost 300 million meters. When x^0 is measured in meters, light conveniently crosses equal amounts of ct and $\sqrt{x^2+y^2+z^2}$ in an appropriately defined Lorentz frame.

Definition 3.1.17 An observer's **proper time** τ is the time measured on a clock stationary in an inertial frame with the observer. If the spacelike interval between two events is timelike, as it always is if an observer was present at both events, then the proper time between events is proportional to the square root of the spacetime interval $(\Delta s)^2$. Proper time is real-valued, so $(\Delta\tau)^2 = (\Delta s)^2$ in the convention where $(\Delta s)^2 > 0$ indicates timelike separation, and $(\Delta\tau)^2 = -(\Delta s)^2$ if timelike separation is indicated by $(\Delta s)^2 < 0$. Proper time is not defined between spacelike separated events: no observer could carry a clock from one event to another without superluminal motions. The proper time is always $\Delta\tau = 0$ between lightlike separated events: clocks stop ticking as one approaches the speed of light.

Definition 3.1.18 If the Minkowski square of a vector is invariant under Lorentz transformations, the scalar is called a **Lorentz scalar** and the vector is called a **Lorentz vector**. If C is a Lorentz scalar and Λ is a Lorentz transformation such that $V^{\mu'} = \Lambda^{\mu'}_{\mu} V^{\mu}$, then

$$V^\mu V_\mu = C \qquad \implies \qquad C = V^{\mu'} V_{\mu'} = \big(\Lambda^{\mu'}_{\mu} V^\mu\big)\big(\Lambda^{\nu}_{\mu'} V_\nu\big) \quad .$$

Definition 3.1.19 The **4-velocity** is

$$U^\mu = \frac{dx^\mu}{d\tau} \quad .$$

Theorem 3.1.20 *The 4-velocity is a Lorentz vector with Lorentz scalar $U^\mu U_\mu = \pm c^2$.*

Proof. We have already established that position x^μ relative stationary frame S will Lorentz transform to a boosted frame S' as

$$x' = \gamma\big(x - vt\big) \quad , \quad \text{and} \qquad t' = \gamma\left(t - \frac{vx}{c^2}\right) \quad ,$$

Since t was measured in the stationary frame, it was the proper time $t = \tau$ of the observer whose *relative* velocity was $\boldsymbol{v} = 0$ (Definition 3.1.17). Per Definition 3.1.19, we have an unboosted 4-velocity

$$U^\mu = \left(\frac{dct}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau}\right) = \left(c\frac{d\tau}{d\tau}, v_x, v_y, v_z\right) = (c, 0, 0, 0) \quad ,$$

and the Minkowski square is

$$U_M^2 = \eta_{\mu\nu} U^\mu U^\nu = \pm c^2 \quad .$$

In the boosted frame S', we have the Lorentz transformed position

$$x^{\mu'} = \begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \Lambda^{\mu'}_{\mu} x^\mu = \begin{bmatrix} \gamma\left(c\tau - \beta x\right) \\ \gamma\big(x - v\tau\big) \\ y \\ z \end{bmatrix} \quad .$$

Observing that $\gamma \neq \gamma(\tau)$ because we have assumed a constant velocity $\boldsymbol{v} = v\hat{e}_x$ between S and S', it follows that U^μ observed from S' will be

$$U^{\mu'} = \begin{bmatrix} \gamma \dfrac{d}{d\tau}\left(c\tau - \beta x\right) \\ \gamma \dfrac{d}{d\tau}\big(x - v\tau\big) \\ \dfrac{d}{d\tau} y \\ \dfrac{d}{d\tau} z \end{bmatrix} = \begin{bmatrix} \gamma c \\ -\gamma v \\ 0 \\ 0 \end{bmatrix} \quad .$$

The Minkowski square of the boosted 4-velocity is

$$U_M'^2 = \eta_{\mu'\nu'} U^{\mu'} U^{\nu'} = \pm\gamma^2 c^2 \mp \gamma^2 v^2 = \pm\gamma^2\big(c^2 - v^2\big) \quad ,$$

where contraction with the metric requires that we prime the *dummy* indices on the metric for consistency in notation. We have

$$\gamma^2 = \frac{c^2}{c^2 - v^2} \quad ,$$

so

$$U_M'^2 = \pm\frac{c^2}{c^2 - v^2}\big(c^2 - v^2\big) = \pm c^2 \qquad \Longrightarrow \qquad U_M'^2 = U_M^2 \quad ,$$

The 4-velocity is a Lorentz vector with invariant $\pm c^2$. This proves the theorem.

Alternatively, U^μ is a Lorentz vector, so given a boost

$$\Lambda^{\mu'}_{\ \mu} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad ,$$

we may write

$$U^{\mu'} = \Lambda^{\mu'}_{\ \mu} U^\mu = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \gamma c \\ -\beta\gamma c \\ 0 \\ 0 \end{bmatrix} \quad ,$$

to compute the Minkowski square directly as

$$U_M'^2 = \begin{bmatrix} \gamma c \\ -\beta\gamma c \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \mp 1 & 0 & 0 \\ 0 & 0 & \mp 1 & 0 \\ 0 & 0 & 0 & \mp 1 \end{bmatrix} \begin{bmatrix} \gamma c \\ -\beta\gamma c \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \gamma c \\ -\beta\gamma c \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} \pm\gamma c \\ \pm\beta\gamma c \\ 0 \\ 0 \end{bmatrix} = \pm c^2 \quad ,$$

where we have use $\gamma^2(1-\beta^2)=1$ at the final step.

Example 3.1.21 Time dilation. It is sometimes useful to think of the 4-velocity as a 1-velocity and a 3-velocity. The 1-velocity measures how quickly time passes, and the 3-velocity is the ordinary mechanical velocity. Since the 4-velocity is normalized such that $U^0 + |\boldsymbol{v}|^2 = \pm c^2$ in any reference frame, the 1-velocity must decrease whenever the 3-velocity increases. This reflects what is called *time dilation*. An observer is always at rest in his own frame ($\boldsymbol{v}=0$), so his time must go slower in any other reference frame. The 1-velocity decreases as the magnitude of the 3-velocity increases. In the standard example of the twin paradox, the twin that goes to another planet and then returns to Earth is younger than his twin that stayed at home because his frame S' was in motion ($\boldsymbol{v} \neq 0$) relative to the Earth frame S. Globally, there are some issues related to acceleration and deceleration that make the twin paradox more complicated, but this is the gist of it.

Example 3.1.22 A spinning bucket of water. In the previous example, we considered S and S' to be related by a rotation-free Lorentz boost, but we might consider S' to be rotating with a constant angular velocity relative to S. Special relativity requires that one frame is as good as another, but consider a bucket of water which is begins rotating in S at $t=0$ and which is co-rotating in S' so that it appears stationary in that frame where the universe appears to rotate in the background. In frame S, the centrifugal acceleration will make the flat surface of the water curve up the sides of the spinning bucket, but the surface of the water will also be observed to curve up the bucket in S' where there is no centrifugal acceleration. If the experiment takes place in deep space, there will not be enough time for the bucket to communicate with the gravitational background so that one might say that the water is dragged up the sides of the bucket by gravity. The bucket experiment exceeds relativity altogether, and makes an appeal to what is called Mach's Principle, described excellently by Woodward in [6].

The MCM offers a new option for explaining the behavior of the water in co-rotating frame S'. Although the water in deep space cannot communicate with the celestial background in the time it takes it to curve up the sides, there exists an immediately proximal quantum vacuum with what is thought to be an infinite vacuum energy density. The usual wrong argument about the water being dragged up the sides of the bucket by the celestial background might be replaced with one referring to the QFT vacuum. The MCM's new algebraic object $\widehat{\infty}$ [7–10] might allow us to quantify an interaction with the vacuum energy like friction so that the water begins to curve as it drags against the non-rotating quantum vacuum. Certainly, the quantum vacuum in and around the bucket is not so far removed that communication becomes impossible.

COMBINE FOLLOWING 2 DEFS
====================

Definition 3.1.23 The **relativistic energy** and **rest energy** of an object are

$$E_{\text{tot}} = \gamma mc^2 \quad , \quad \text{and} \qquad E_0 = mc^2 \quad .$$

The rest energy E_0 is the energy measured in a frame where the object is a rest: $\boldsymbol{v} = 0$ and $\gamma(0) = 1$. The *rest mass* m, also called the proper mass or the *invariant mass*, is the mass of an object in the frame where it is at rest. The relativistic energy E_{tot} is called the total energy. The kinetic energy is the difference between the rest energy and the total energy:

$$E_k = E_{\text{tot}} - E_0 = m_0 c^2 \big(\gamma - 1 \big) \quad .$$

Sometimes one labels the rest mass m_0 and introduces the relativistic mass $m = \gamma m_0$ so that $E_{\text{tot}} = mc^2$, but we will use m to describe the rest mass.

Definition 3.1.24 The relativistic energy and momentum of a massive particle are

$$E = \gamma mc^2 \quad , \quad \text{and} \qquad \boldsymbol{p} = \gamma m \boldsymbol{u} \quad ,$$

where the 3-velocity $\boldsymbol{u}$ of the particle is not to be confused with the 3-velocity $\boldsymbol{v}$ between two Lorentz frames. The kinetic energy is

$$E_k = mc^2 \big(\gamma - 1 \big) \quad .$$

The rest energy is mc^2, which is also called *the invariant mass*, so the kinetic energy is the difference between the total relativistic energy and the rest energy.

Definition 3.1.25 The **dispersion relation**, or the energy-momentum relation, for a massive relativistic particle is

$$E^2 = \big| \boldsymbol{p} c \big|^2 + \big(mc^2 \big)^2 \quad ,$$

This relationship is best understood as the Pythagorean theorem for the Einstein triangle, as in Figure XXXX.

"The energy–momentum relation goes back to Max Planck's article published in 1906. "

`https://en.wikipedia.org/`

For a massless particle, the dispersion relation is

Definition 3.1.26 The **4-momentum** is

$$p^\mu = mU^\mu \quad , \quad \text{with} \qquad p_M^2 = m_0^2 c^2 \quad .$$

Note that Lorentz invariant associated with 4-momentum satisfies $p_M^2 = m_0^2 U_M^2$.

Definition 3.1.27 The **rapidity** ρ is a parameter defined as

$$\beta = \frac{v}{c} = \tanh\rho \quad , \quad \text{and} \qquad \gamma = \cosh\rho \quad .$$

In terms of rapidity, boost matrices may be written with hyperbolic trigonometry functions:

$$\mathbf{\Lambda}(\beta) = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \longrightarrow \quad \mathbf{\Lambda}(\rho) = \begin{bmatrix} \cosh\rho & -\sinh\rho & 0 & 0 \\ -\sinh\rho & \cosh\rho & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad .$$

The components of the 4-momentum satisfy

$$E_{\text{tot}} = mc^2\cosh\rho \quad , \quad \text{and} \qquad |\mathbf{p}| = mc\sinh\rho \quad .$$

====================

For connections to the ordinary formulae in special relativity, it is convenient to define $\beta = \tanh\chi$ (where $\beta = v/c$) so that using $\cosh^2\chi - \sinh^2\chi = 1$ we obtain:

$$\cosh\chi = \frac{1}{\sqrt{1-\beta^2}} \quad , \quad \text{and} \qquad \sinh\chi = \frac{\beta}{\sqrt{1-\beta^2}}$$

Filling these values into the B matrix gives the familiar formulae:

$$x_0' = \frac{x_0 + x_1\beta}{\sqrt{1-\beta^2}} \quad , \quad \text{and} \qquad x_1' = \frac{\beta x_0 + x_1}{\sqrt{1-\beta^2}}$$

WHAT ABOUT BOOSTS WITH $\beta > 0$??? SHOULD MOVE E INTO THE TRANSFINITE REGIME.

Remark 3.1.28 This is why they say a boost is rotation between space and time by a complex angle.

Example 3.1.29 Lorentz boost of a 4-momentum. If we let $\Lambda_\mu^{\mu'}$ be a boost in an arbitrary direction, then the $U^{\mu'}$ obtained in Theorem 3.1.20 becomes

$$U^{\mu'} = (\gamma c, -\gamma v_x, -\gamma v_x, -\gamma v_z) \quad .$$

Changing the dummy index back to μ and substituting the frame velocity $\boldsymbol{v}$ with $-\boldsymbol{u}$ (the velocity of a particle moving oppositely in a stationary frame such that $\gamma(0)=1$), an arbitrary 4-momentum is written

$$p^\mu = \begin{bmatrix} mc \\ mu_x \\ mu_y \\ mu_z \end{bmatrix} = \begin{bmatrix} E/c \\ p_x \\ p_y \\ p_z \end{bmatrix} \quad .$$

The p^0 component is the relativistic energy, and (p^1, p^2, p^3) is the classical momentum $\boldsymbol{p}$. When $\gamma = 1$ ($\boldsymbol{v}=0$), p^0 is the rest energy over c: E_0/c.

A boost in the x-direction (boosted with $\boldsymbol{v}'$ not necessarily equal to $-\boldsymbol{u}$) is written

$$p^{\mu'} = \begin{bmatrix} p^{0'} \\ p^{1'} \\ p^{2'} \\ p^{3'} \end{bmatrix} = \Lambda^{\mu'}_{\mu} p^\mu = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} E/c \\ p_x \\ p_y \\ p_z \end{bmatrix} = \begin{bmatrix} \gamma\left(\frac{E}{c} - \beta p_x\right) \\ \gamma\left(p_x - \frac{\beta E}{c}\right) \\ p_y \\ p_z \end{bmatrix}$$

Since the components of the boosted 4-momentum depend on γ, energy and momentum are not conserved (CHECK E_{tot}!!!) in special relativity. We have Minkowski square

$$p^\mu p_\mu = \begin{bmatrix} \pm E/c \\ \mp p_x \\ \mp p_y \\ \mp p_z \end{bmatrix}^T \begin{bmatrix} E/c \\ p_x \\ p_y \\ p_z \end{bmatrix} = \pm\left[\left(\frac{E}{c}\right)^2 - \left|\boldsymbol{p}\right|^2\right] \quad ,$$

and for a boosted 4-momentum we have

$$\begin{aligned} p^{\mu'} p_{\mu'} &= \begin{bmatrix} \pm\gamma\left(\frac{E}{c} - \beta p_x\right) \\ \mp\gamma\left(p_x - \frac{\beta E}{c}\right) \\ \mp p_y \\ \mp p_z \end{bmatrix}^T \begin{bmatrix} \gamma\left(\frac{E}{c} - \beta p_x\right) \\ \gamma\left(p_x - \frac{\beta E}{c}\right) \\ p_y \\ p_z \end{bmatrix} \\ &= \pm\left[\gamma^2\left(\tfrac{E}{c} - \beta p_x\right)^2 - \gamma^2\left(p_x - \tfrac{\beta E}{c}\right)^2 - p_y^2 - p_z^2\right] \end{aligned}$$

After grouping like terms and using $\gamma^2(1-\beta^2)=1$, we obtain

$$p^{\mu'} p_{\mu'} = \pm\left[\left(\frac{E}{c}\right)^2 - \left|\boldsymbol{p}\right|^2\right] \qquad \Longrightarrow \qquad p_M^2 = p_M'^2$$

Thus, the 4-momentum p^μ is a Lorentz vector. Per Definition 3.1.25, the dispersion relation for a massive relativistic particle is

$$E^2 = \left|\boldsymbol{p}c\right|^2 + \left(mc^2\right)^2 \quad ,$$

so we are able to obtain the Lorentz invariant $p_M^2 = \pm m^2c^2$.
TOTAL REL ENERGY IS CONSERVED???

Remark 3.1.30 This will be useful later in QFT when we can obtain an arbitrary momentum state $|k\rangle$ by applying a boost to the $k = 0$ state $|0\rangle$. Since many of these particles will be fermions, we will need to identify the rules for spinors under LORENTZ.

Theorem 3.1.31 *The Lorentz invariant associated with a displacement 4-vector is SPACETIME INTERVAL.*

$$XXXXXXXXXXX \quad .$$

Theorem 3.1.32 *"In other words, the 4-acceleration of a particle is always orthogonal to its 4-velocity."*
URL:

Remark 3.1.33 On-shell and off-shell for massive and massless particles.
Real and virtual, observable and unobservable.

§3.2 Lorentzian Manifolds

At this point, it will not be our intention to get too deep in the jargon of the mathematical underpinnings of the most general group theory of relativistic Lorentz transformations. However, a *Lie group* such as the Lorentz group has an associated manifold, and we will want to study a group representation $V^\mu \in \mathcal{M}$ other than the usual one $V^\mu \in \mathbb{R}^4$. To demonstrate the well-motivation of this research line, we will need to show that the new manifold $\mathcal{M}$ presented here is the same as the better-studied one in contemporary usage. To present at least the veneer of rigor in this task, we will need some mathematical definitions beyond the bare minimum needed to do physics in special relativity.

Wald describes a manifold as follows [11].

> "An n-dimensional manifold is a set that has the local differential structure of $\mathbb{R}^n$ but not necessarily its global properties."

Local resemblance to $\mathbb{R}^n$ guarantees that we can assume locally inertial coordinates in the manifold: a Lorentz frame if $n = 4$. This will hold in the general theory of relativity when the manifold becomes curved, and presently the identical flatness of Minkowski space in the special theory is such that every Lorentz frame, assumed at any point, in any orientation, and moving with any $|\boldsymbol{v}| < c$, is valid on all of spacetime. However, the caveat about the manifold not necessarily having the global properties of $\mathbb{R}^4$ opens the door to 4-vectors being representations of groups other than O(4): the orthogonal group in four dimensions whose manifold is the vector space $\mathbb{R}^4$ spanned by four orthogonal directions. The global properties of $\mathbb{R}^n$ include distance being defined with the Euclidean metric, but, in a manifold, we may introduce $V^\mu \in \mathbb{R}^4$ equipped with the scalar product

$$V^\mu W_\mu = g_{\mu\nu} V^\mu W^\nu \quad ,$$

where $g_{\mu\nu}$ is an arbitrary metric tensor. We say V^μ belongs to $\mathbb{R}^4$ because it is a -tuple of four real numbers, and because the four components transform as a vector in the vector space $\mathbb{R}^4$, but we may further specify $V^\mu \in (\mathbb{R}^4, \eta_{\mu\nu})$ where $(\mathbb{R}^4, \eta_{\mu\nu})$ is a manifold equipped with the Minkowski metric. Up to some technical details, a manifold with a metric tensor is a Riemannian manifold.

Every finite-dimensional vector space is automatically a manifold, but in Euclidean manifolds distance is more properly attributed to the Pythagorean theorem than it is to tensor contraction with the trivial Euclidean metric $\delta_{\mu\nu}$. Riemannian manifolds that specify a metric explicitly open the door non-Euclidean geometries in which the Pythagorean theorem doesn't hold (curved spacetime) and distance isn't positive-definite (spacelike, timelike, and null spacetime intervals). With these tools, once we know $(\mathbb{R}^4, \eta_{\mu\nu})$ is the manifold of the Lorentz group, a Lorentzian manifold which is a certain type of pseudo-Riemannian manifold, so is any other manifold that is diffeomorphic to $(\mathbb{R}^4, \eta_{\mu\nu})$. We need to show (CHECK SAMENESS OF MANIFOLDS??) that $(\mathcal{M}, \delta_{\mu\nu})$ is diffeomorphic to $(\mathbb{R}^4, \eta_{\mu\nu})$.

Lee writes the following about a manifold M [12].

> "We write local coordinates on any open subset $U \subset M$ as $(x^1, \ldots, x^n)$, x^i, or x, depending on context. Although, formally speaking, coordinates constitute a map from U to $\mathbb{R}^n$, it is more common to use a coordinate chart to identify U with its image in $\mathbb{R}^n$, and to identify a point in U with its coordinate representation (x^i) in $\mathbb{R}^n$."

The arrangement described by Lee is shown in Figure 2. For special relativity, the manifold M is not curved, and we may take U to be the whole of M. The $V^\mu \in \mathbb{R}^4$ representation of the Lorentz group *identifies* the vector in the manifold with its image in $\mathbb{R}^n$, as is "more common." $\mathbb{R}^n$ itself satisfies the requirement that the manifold looks locally like $\mathbb{R}^n$. However, the program here will be to carefully separate complex vectors in the manifold from the coordinate representation, and we will need to construct a complex manifold that looks locally like $\mathbb{C}^n$ rather than $\mathbb{R}^n$. Depending on the convention for

$$\mathcal{M} = i\mathbb{R} \times \mathbb{R}^3 \quad , \quad \text{or} \qquad \mathcal{M} = \mathbb{R} \times i\mathbb{R}^3 \quad ,$$

the map from $U \subset M$ to its image $x^\mu(U)$ in $\mathcal{M} \subset \mathbb{C}^4$ with have some factors of i in it.

A first separation of a general manifold from the Euclidean manifold $\mathbb{R}^n$ is the Riemannian manifold: a set equipped with a positive-definite inner

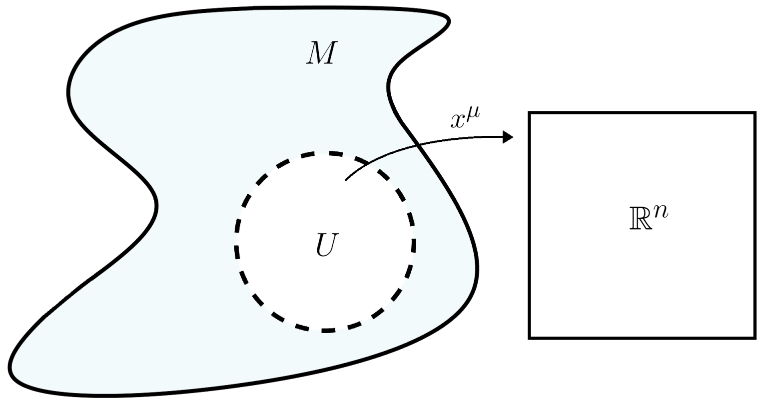


Figure 2: If M is spacetime, U is a local Lorentz frame. MOAR MOAR MOAR MOAR MOAR MOAR M is the realm of abstraction, and $\mathbb{R}^n$ is the realm observable quantities. Since time, left-right position, front-back position, and up-down position are measured as real-valued quantities of seconds or meters, the theory of special relativity must connect with $\mathbb{R}^4$ on some level: the space of tuples of four real numbers. Because the theory is pretty simple, one can select the manifold to be $\mathbb{R}^4$ equipped as a vector space, which is the Euclidean 4-space $\mathbb{E}^4$. By letting M be a pseudo-Riemannian manifold, which is a manifold equipped with a metric, we are able to obtain the manifold of special relativity.

product. This does not suffice for the manifold of special relativity because we want the inner products of our 4-vectors to be spacelike, null, and timelike corresponding to positive, negative, and vanishing scalar products. Thus, we arrive at the pseudo-Riemannian manifold. The restriction that the inner product should be positive-definite is relaxed, and instead we only require that the metric is non-degenerate. Such things will be defined in this section.

Definition 3.2.1 Two sets S_1 and S_2 are **homeomorphic** if there exists a continuous bijection $f : S_1 \to S_2$ with a continuous inverse $f^{-1} : S_2 \to S_1$. Such a function is called a **homeomorphism**.

Definition 3.2.2 A set S is **locally Euclidean** if every point $p \in S$ has a neighborhood $U \in S$ that is homeomorphic to an open subset $U' \in \mathbb{R}^n$ for some fixed n.

Remark 3.2.3 Boothby writes the following [13].

> "It follows from the homeomorphism of U and U' that locally Euclidean is equivalent to the requirement that each point p have a neighborhood U homeomorphic to an n-ball in $\mathbf{R}^n$. Thus a manifold of dimension 1 is locally homeomorphic to an open interval, a manifold of dimension 2 is locally homeomorphic to an open disk, and so on."

Definition 3.2.4 A **topology** on a set S is a collection $\mathcal{T}$ of open subsets of S with the following properties.

- $\mathcal{T}$ contains S and $\varnothing$.
- $\mathcal{T}$ contains the union of any of the elements of $\mathcal{T}$, namely $\bigcup \tau_k \in \mathcal{T}$.
- $\mathcal{T}$ contains the finite intersection of the elements of $\mathcal{T}$. If $n < \infty$, then $\bigcap\limits_{k=1}^{n} \tau_k \in \mathcal{T}$.

The open sets in $\mathcal{T}$ are called the **basis** of the topology. The topology is the set of all unions of the sets in its basis. Together, the pair $(S, \mathcal{T})$ is called a **topological space**. One says S is equipped with the topology $\mathcal{T}$, that $\mathcal{T}$ is a datum of S, or that $\mathcal{T}$ is in the data of S.

Definition 3.2.5 The **usual basis** $\mathcal{B}_0$ for the **usual topology** on $\mathbb{R}^n$ is the collection of all open balls centered on every point in $\mathbb{R}^n$. The topology generated by $\mathcal{B}_0$ is called $\mathcal{T}_0$.

Definition 3.2.6 A topological space $(S, \mathcal{T})$ is **Hausdorff** if any two points $p, q \in S$ belong to open sets $O_p, O_q \in \mathcal{T}$ such that $O_p \cap O_q = \varnothing$. In general, a Hausdorff space is one in which every point has an open neighborhood that is also in the space. A topological space is Hausdorff if any two distinct points have non-intersecting open neighborhoods, or, in other words, if any two distinct points are separated by an open set in the topology. FIGURE

Definition 3.2.7 A **manifold** of dimension n, or an n-manifold, is a locally Euclidean topological space that is also Hausdorff and has a second-countable basis.

Remark 3.2.8 The details of second-countability may be found in Appendix A of Wald [11], for example, or Chapter 2 of Lee [14]. By restricting the open sets in the basis for $\mathbb{R}^n$ to be the set of all open balls with rational radius $r \in \mathbb{Q}$ centered on points with rational coordinates $(q^1, \ldots, q^n)$ with $q^k \in \mathbb{Q}$, we will have a second-countable basis for the usual topology. In the present context, it suffices that a second-countable basis will guarantee that a manifold admits a Riemannian metric. We will not study the topological properties of manifolds here, however, and the specifications of the open sets in a given topology will be mostly out of scope. Instead, we will be concerned with the differentiable properties of manifolds even as all differentiable manifolds are automatically topological manifolds. In general, from the perspective of a physicist, a manifold is a space that can be described with coordinates, and a metric defines distance between points. Lee writes the following [14].

> “A physicist would say that an n-dimensional manifold is an object with n ‘degrees of freedom.’”

========================

FROM: https://math.stackexchange.com/questions/472998/why-do-we-need-hausdorff-ness-in-definition-of-topological-manifold

"One author who explicitly avoids assuming the Hausdorff property is Serge Lang, "Fundamentals of differential geometry", Springer 1999, 2001. In Chapters II and II, he does not assume Hausdorff. He introduces the condition at the beginning of Chapter IV. He says: 'We see no reason to assume that X is Hausdorff. If we wanted X to be Hausdorff, we would have to place a separation condition on the covering. This plays no role in the formal development in Chapters II and III.'"

====================

Example 3.2.9 An n-manifold need not be $\mathbb{R}^n$. A manifold must be locally Euclidean, but it does not have to be Euclidean altogether. As an example of a manifold that is not globally homeomorphic to $\mathbb{R}^n$, Boothby describes the 2-sphere $\mathbb{S}^2$ in Chapter 1 of [13]. The 2-sphere is the set of all points at fixed distance from the origin of $\mathbb{R}^3$; suppression of one degree of freedom by the fixed spherical radius in $\mathbb{R}^3$ leaves two degrees of freedom characterizable as the azimuth and zenith angles. The neighborhood of any point $p \in \mathbb{S}^2$ has an open neighborhood U that can be projected in one-to-one correspondence onto a plane passing through the origin of $\mathbb{R}^3$. This projection is onto, clearly, as in Figure XXX, so it is bijective, and its inverse exists. The plane of projection is necessarily $\mathbb{R}^2$, so $\mathbb{S}^2$ is *locally Euclidean*. By a well-known stereographic projection, however, a Euclidean plane is homeomorphic to a 2-sphere missing a point: as points become infinitely far from the origin of a plane, they become infinitely close to a pole on $\mathbb{S}^2$. The projection is not one-to-one and not bijective. So, $\mathbb{S}^2$ is not *homeomorphic* to $\mathbb{R}^n$ globally.

Wald writes the following [11].

> "The entire 2-sphere S^2 cannot be mapped into $\mathbb{R}^2$ in a continuous, 1–1 manner, but 'pieces' of S^2 can, and these can be 'smoothly sewn together.' For example, if we define six hemispherical open sets $O_i^\pm$ for $i = 1, 2, 3$ by
>
> $$O_i^\pm = \left\{(x^1, x^2, x^3) \in S^2 \;\middle|\; \pm x^i > 0\right\} \quad ,$$
>
> then $\{O_i^\pm\}$ covers S^2. Furthermore, each $O_i^\pm$ can be mapped homeomorphically into the open disk $D = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 < 1\}$ in the plane via the 'projective maps' $f_1^+ : O_1^+ \to D$, $f_1^- : O_1^- \to D$, etc, defined by $f_1^+(x^1, x^2, x^3) = (x^2, x^3)$, etc. The overlap functions $f_i^\pm \circ (f_j^\pm)^{-1}$ can be checked to be C^∞ in their domain of definition[.] Thus, S^2 is a two-dimensional manifold."

Jumping ahead to Riemannian manifolds, which are manifolds with metrics, one might equip $\mathbb{S}^2$ so that the distance between $p, q \in \mathbb{S}^2$ is the arclength on the sphere rather than the straight-line distance between $p, q \in \mathbb{R}^3$. On a small enough patch of $\mathbb{S}^2$, the curvature becomes negligible, and the arclength approaches the Euclidean distance, but these metric functions will not agree on distances in general, so we would have two globally unequal Riemannian manifolds.

Definition 3.2.10 A **metric space** is a set S equipped with a **metric function** $d : S \times S \to \mathbb{R}$. The metric function has the following properties.

- For every $x, y \in S$, we have $d(x, y) \geq 0$.
- We have $d(x, y) = 0$ if and only if $x = y$.
- For every $x, y, z \in S$, $d(x, z) \leq d(x, y) + d(y, z)$, and
- $d(x, y) = d(y, x)$.

NICE NOTES ON METRIC SPACE: URL:

Example 3.2.11 A hierarchy of increasing data. Before we can ask if S is *locally Euclidean*, we must ensure that S contains the open sets required by Definition 3.2.2. Therefore, a locally Euclidean set is already equipped as a topological space: it has a basis of open sets, as in Definition 3.2.4. In the absence of any words to the contrary, we should assume that the topology is the usual one, meaning that S is already equipped as a metric space as well because open balls centered on points are defined with distances calculated by the *metric function* given in Definition 3.2.10.

Remark 3.2.12 In the remainder of this section, we will continue to build the hierarchy of data by adding a metric tensor $g_{\mu\nu}$ to S. Then will establish conditions on the metric tensor, which we call *the metric*, until we arrive at the Lorentzian manifold most commonly used to study special or general relativity. In the next two sections, we will build the tangent spaces to the manifold, and in the third following section we will examine the kinds of inner products we can equip to our tangent spaces and develop the requirements for the new, other-than-Lorentzian manifold that we want to study. In Section 3.4, we will continue building the hierarchy of data until we find that the XXXX manifold is the manifold we are looking for.

Example 3.2.13 Differentiable Manifolds. In [13], Boothby introduces differentiable manifolds as follows (depicted independently in Figure 3).

> "Each pair U, φ, where U is an open set of M and φ is a homeomorphism of U to an open subset of $\mathbf{R}^n$, is called a *coordinate neighborhood*: to $q \in U$ we assign the n *coordinates* $x^1(q), \ldots, x^n(q)$ of its image $\phi(q)$ in $\mathbf{R}^n$—each $x^i(q)$ is a real-valued function on U,

> the ith coordinate function. If q lies also in a second coordinate neighborhood V, ψ, then it has coordinates $y^1(q), \dots, y^n(q)$ in this neighborhood. Since ϕ and ψ are homeomorphisms, this defines a homeomorphism
>
> $$\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V),$$
>
> the domain and range being the two open subsets of $\mathbf{R}^n$ which correspond to the points of $U \cap V$ by the two coordinate maps φ, ψ, respectively."

Let U and V be two Lorentz frames in Minkowski space: S and S' respectively with coordinates x^μ and $x^{\mu'}$. These frames might be boosted, rotated, boosted and rotated, or merely displaced, but we will assume that $x^\mu \neq x^{\mu'}$ so that $\varphi = x^\mu$ and $\psi = x^{\mu'}$ are two different coordinate charts. (Compare to Figure 2 in which the coordinate chart was labeled x^μ, as per Lee [12].) If the past light cones of observers in S and S' overlap, then there will be some region of spacetime containing events $p \in M$ jointly described by the x^μ and $x^{\mu'}$ coordinates. This is the region $U \cap V$ in Figure 3: the intersection of U and V. The images of $U \cap V$ under coordinate charts ψ and φ are disjoint in Figure 3 but the images are not necessarily disjoint; the same tuple of four real numbers may describe two different events in S and S', but the figure highlights that these are two different images: $\varphi(U \cap V)$ and $\psi(U \cap V)$. Boothby's notation $\psi \circ \varphi^{-1}$ means "ψ acting on the inverse of ϕ," so, referring to Figure 3, ϕ^{-1} takes an image in $\mathbb{R}^n$ back to M, and then ψ takes that back to a different image in $\mathbb{R}^n$. The composition $\psi \circ \varphi^{-1}$ is a coordinate transformation from S to S', and the inverse $\varphi \circ \psi^{-1}$ transforms the coordinates of S' into those of S. The composition of two homeomorphisms is another homeomorphism, so the Lorentz transformations between different coordinates in different inertial frames in spacetime are automatically homeomorphisms. Boothby continues as follows [13].

> "In coordinates, $\psi \circ \varphi^{-1}$ is given by continuous functions
>
> $$y^i = h^i(x^1, \dots, x^n), \quad i = 1, \dots, n,$$
>
> giving the y-coordinates of each $q \in U \cap V$ in terms of its x coordinates. Similarly, $\varphi \circ \psi^{-1}$ gives the inverse mapping which expresses the x-coordinates as functions of the y-coordinates
>
> $$x^i = g^i(y^1, \dots, y^n), \quad i = 1, \dots, n.$$
>
> The fact that $\varphi \circ \psi^{-1}$ and $\psi \circ \phi^{-1}$ are homeomorphisms and are inverse to each other is equivalent to the continuity of $h^i(x)$ and $g^j(y)$, $i, j = 1, \dots, n$ together with the identities
>
> $$h^i(g^1(y), \dots, g^n(y)), \quad i = 1, \dots, n,$$

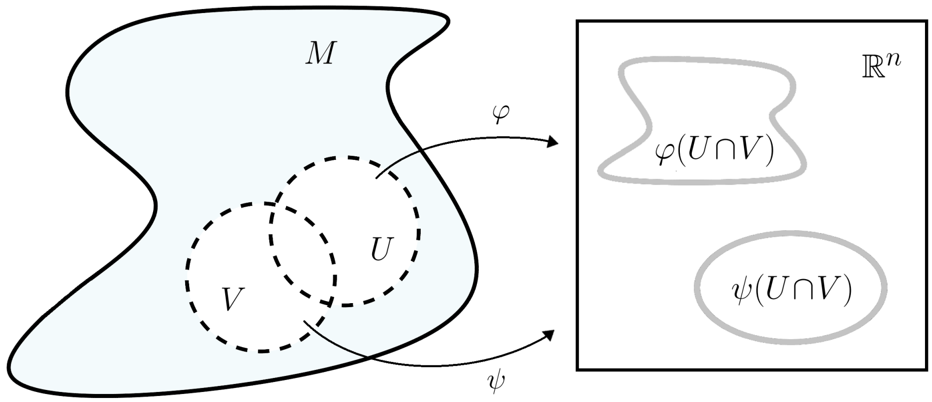


Figure 3: OBSERVABLE QUANTITIES ARE REAL!!! It is also possible to think of φ and ψ as mapping to two different copies of $\mathbb{R}^n$.

> and
> $$g^j(h^1(x),\ldots,h^n(x)),\quad j=1,\ldots,n.$$
> Thus every point of a topological manifold M lies in a very large collection of coordinate neighborhoods, but whenever two neighborhoods overlap we have the formulas just given for a change of coordinates. The basic idea that leads to differentiable manifolds is to try to select a family or subcollection of neighborhoods so that the change of coordinates is always given by differentiable functions."

Boothby's words easily interpretted in the context of special relativity. Let $S=U$ and $S'=V$ be offset by an unboosted rotation in the xy-plane. The primes coordinates $x^{\mu'}$ in S' are what Boothby calls y^i, which are separate from the $x^2=y$ used to specify the plane of rotation. The rotation matrix is

$$\mathbf{\Lambda}=\begin{bmatrix}1&0&0&0\\0&\cos\theta&\sin\theta&0\\0&-\sin\theta&\cos\theta&0\\0&0&0&1\end{bmatrix}\quad,$$

so

$$\begin{bmatrix}x^{0'}\\x^{1'}\\x^{2'}\\x^{3'}\end{bmatrix}=\begin{bmatrix}1&0&0&0\\0&\mathrm{C}&\mathrm{S}&0\\0&-\mathrm{S}&\mathrm{C}&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}x^0\\x^1\\x^2\\x^3\end{bmatrix}=\begin{bmatrix}x^0\\x^1\mathrm{C}+x^2\mathrm{S}\\x^2\mathrm{C}-x^1\mathrm{S}\\x^3\end{bmatrix}\quad.$$

Matrix algebra is a concise way of dealing with systems of equations, and the continuous transformation functions cited by Boothby are

$$x^{0'}=h^0(x^0,x^1,x^2,x^3)=x^0$$

$$x^{1'} = h^1(x^0,x^1,x^2,x^3) = x^1\cos\theta + x^2\sin\theta$$
$$x^{2'} = h^2(x^0,x^1,x^2,x^3) = x^2\cos\theta - x^1\sin\theta$$
$$x^{3'} = h^3(x^0,x^1,x^2,x^3) = x^3 \quad .$$

Each of the transformed coordinates is specified by a function that may depend on one or all of the untransformed coordinates. To obtain the unprimed coordinates in terms of the primed coordinates, we must rotate by $-\theta$. Using $\cos(-\theta) = C$ and $\sin(-\theta) = -S$, we have

$$\begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathrm{C} & -\mathrm{S} & 0 \\ 0 & \mathrm{S} & \mathrm{C} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{bmatrix} = \begin{bmatrix} x^{0'} \\ x^{1'}\mathrm{C} - x^{2'}\mathrm{S} \\ x^{2'}\mathrm{C} + x^{1'}\mathrm{S} \\ x^{3'} \end{bmatrix} \quad ,$$

so

$$x^0 = g^0(x^{0'},x^{1'},x^{2'},x^{3'}) = x^{0'}$$
$$x^1 = g^1(x^{0'},x^{1'},x^{2'},x^{3'}) = x^{1'}\cos\theta - x^{2'}\sin\theta$$
$$x^2 = g^2(x^{0'},x^{1'},x^{2'},x^{3'}) = x^{2'}\cos\theta + x^{1'}\sin\theta$$
$$x^3 = g^3(x^{0'},x^{1'},x^{2'},x^{3'}) = x^{3'} \quad .$$

To verify that everything works, we plug the expressions for the primed coordinates back into the unprimed coordinates:

$$x^0 = x^0$$
$$x^1 = \big(x^1\mathrm{C} + x^2\mathrm{S}\big)\mathrm{C} - \big(x^2\mathrm{C} - x^1\mathrm{S}\big)\mathrm{S} = x^1\big(\mathrm{C}^2 + \mathrm{S}^2\big) = x^1$$
$$x^2 = \big(x^2\mathrm{C} - x^1\mathrm{S}\big)\mathrm{C} + \big(x^1\mathrm{C} + x^2\mathrm{S}\big)\mathrm{S} = x^2\big(\mathrm{C}^2 + \mathrm{S}^2\big) = x^2$$
$$x^3 = x^3 \quad .$$

Everything works.

For physics, every point in the manifold of special relativity lies in a very large number of Lorentz frames, not only ones separated by an unboosted rotation. Since physics requires calculus, and meaning that we will want to study motion, which is the smooth variation of coordinates, we require that the manifold of special relativity should be a *differentiable manifold* in which the h^i and g^i functions given by the Lorentz transformation matrices are always differentiable functions.

Remark 3.2.14 At this point, it is reasonable to ask, “If the coordinates are in $\mathbb{R}^n$, then what is in M?” This is the issue described by Lee when he says it is common to *identify* U with its image in $\mathbb{R}^n$. In special relativity, a event is a point $p \in M$ and its coordinates are a tuple (t, x, y, z) defined with a coordinate chart. Usually the tuple is four real numbers in $\mathbb{R}^4$, but we are going to study the case when the 4-tuple contains one real number or three with the other imaginary.

Definition 3.2.15 If U is an open subset of n-manifold M, a **coordinate map** is a homeomorphism $\varphi : M \to \mathbb{R}^n$. For $p \in U$, the n components of $\varphi(p) = (x^1(p), \ldots, x^n(p))$, are called **coordinate functions**. As functions, they are $x^\mu : M \to \mathbb{R}$. The pair (U, φ) is called a **coordinate chart**, and $\varphi^{-1} : \mathbb{R}^n \to U$ is called a **parameterization** of U. MAKE FIG

Remark 3.2.16 Following a standard practice in physics, we will refer to vector $x = \sum x^\mu \hat{e}_\mu$ by its components x^μ, which intermingles the concept of a coordinate map that takes $p \in M$ to $\mathbb{R}^n$ and the individual coordinate functions which take $p \in M$ to $\mathbb{R}$. So, the coordinate map is the vector of the values of the coordinate functions.

Wald (p24) [11] uses "abstract index notation" such that we would distinguish x^a with a Latin index as referring to the vector while x^μ with the Greek index is a single component. The single upper index on x^a reminds us that a vector is a (1,0) tensor. It returns a scalar by acting on a dual vector. In this case, x^a would be the coordinate map called φ in Definition 3.2.15, and x^μ would be the components of the map, called the coordinate functions.

NEED TO STANDARDIZE? MAYBE SF FONT? x^a x^{a}

The components usually refer to the basis assigned to a given Lorentz frame, and this tends to be a preferred basis for talking about physics. This is why we have written x^μ and $x^{\mu'}$ referring to frames S and S' but often mathematicians like to talk about the basis-independent object.

Definition 3.2.17 A union of open subsets of M is called a **cover** of M if

$$M = \bigcup_i U_i \quad .$$

Definition 3.2.18 An **atlas** is a collection of coordinate charts $\{(U_i, \varphi_i)\}$ on manifold M such that $\{U_i\}$ is a cover of M. To avoid defining a new manifold every time one assigns a new coordinate map to $\mathbb{R}^n$, the standard convention is that a manifold comes equipped with the maximal atlas of all of its coordinate charts. The atlas $\{(U_i, \varphi_i)\}$ is said to be a datum of the manifold or in the data of the manifold.

Definition 3.2.19 For any two coordinate charts $(U_j, \varphi_j), (U_k, \varphi_k)$ in an atlas $\{(U_i, \varphi_i)\}$, the compositions $\varphi_j \circ \varphi_k^{-1}$ are called **transition functions**. In physics, they are usually called coordinate transformations. The action of two transition functions is shown in Figure XXX.

Definition 3.2.20 A function is smooth if it is infinitely differentiable. If X and Y are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, then Lee states in Chapter 1 of [15] that $f : X \to Y$ is **smooth**, "if each of its component functions has continuous partial derivatives of all orders." Smooth functions are called C^∞ functions.

Remark 3.2.21 Sometimes a function is called smooth if it has continuous partial derivatives to order $k > 0$. Such functions are called C^k, but we will follow the convention of Lee [15] to identify smoothness with C^∞.

Definition 3.2.22 Two sets S_1 and S_2 are **diffeomorphic** if there exists a C^∞ bijection $f : S_1 \to S_2$ with a C^∞ inverse $f^{-1} : S_2 \to S_1$. Such a function is called a **diffeomorphism**. Every diffeomorphism is a homeomorphism, so every pair of diffeomorphic sets are homeomorphic.

Definition 3.2.23 For two open subsets U_1, U_2 of manifold M, the coordinate charts $(U_1, \varphi_i), (U_2, \varphi_2)$ are **C^∞-compatible** if their transition functions are diffeomorphisms. Since one is the inverse of the other (Figure XXX), they are both diffeomorphisms if either of them are.

Theorem 3.2.24 *If a transition function $\varphi_j \circ \varphi_k^{-1}$ is smooth, then the maps φ^j and φ^k are smooth.*

Proof. XXXXXXXXXXXXXXXXXX

Definition 3.2.25 A **smooth structure** on a manifold M is subset of its atlas with the following properties.

- The U_i domains in $\mathcal{U}$ cover M.
- Every pair of charts $(U_1, \varphi_i), (U_2, \varphi_2) \in \mathcal{U}$ are C^∞-compatible.
- Any chart that is C^∞-compatible with every chart in $\mathcal{U}$ is also in $\mathcal{U}$.

Definition 3.2.26 A **smooth manifold**, also called a differentiable manifold, is manifold equipped with a smooth structure.

Definition 3.2.27 To each point $p \in M$ we associate a vector space spanned by the derivatives with respect to the coordinate functions at p. This space is called the **tangent space** to M at p and it is denoted T_pM. The **tangent bundle** is the union of all the tangent spaces to a manifold, and it is denoted TM.

Penrose Vol 1: Nice comment on bundles p333.

Definition 3.2.28 A **metric tensor** is a bilinear form on a vector space.

A metric is not a metric tensor!!

A metric tensor is **positive-definite** is

A metric tensor is **non-degenerate** if

A metric can always be represented as a diagonal matrix??? A **metric signature**

"By Sylvester's law of inertia these numbers do not depend on the choice of basis and thus can be used to classify the metric. "

Definition 3.2.29 A **Riemannian manifold** is a smooth manifold equipped with a positive-definite metric tensor field on its tangent bundle. If the positive-definite condition is relaxed to allow a non-degenerate metric, then the manifold is **pseudo-Riemannian**.

QUOTE: A **Lorentzian manifold** is an important special case of a pseudo-Riemannian manifold in which the signature of the metric is $(1, n-1)$ (equivalently, $(n-1, 1)$; see Sign convention).

Definition 3.2.30 A metric tensor is a **Euclidean metric** if it may be written as

$$g_{\mu\nu} = \text{diag}(\underbrace{1, \ldots, 1}_{n \text{ 1's}}) \quad .$$

It is a **pseudo-Euclidean metric** if it may be written as

$$g_{\mu\nu} = \text{diag}(\underbrace{1, \ldots, 1, -1, \ldots, -1}_{p \text{ 1's and } q - 1\text{'s}}) \quad , \quad \text{where} \qquad p + q = n \quad .$$

If $p = 1$ or $q = 1$, then $g_{\mu\nu}$ is a **flat Lorentzian metric**. If $p \geq 2$ and $q \geq 2$, the it is a **pseudo-Lorentzian metric**. Every Lorentzian or pseudo-Lorentzian metric is pseudo-Euclidean.

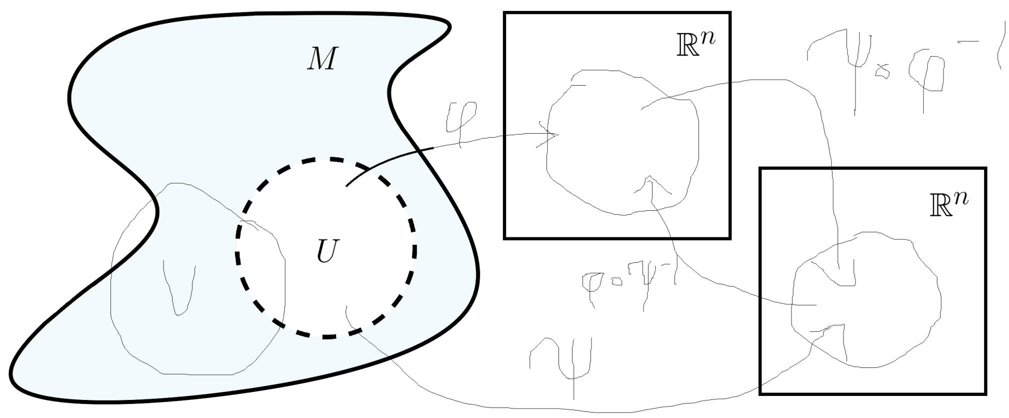


Figure 4: XXX

KKKKKKKKKKKKKKKKKKK

Theorem 3.2.31 *The Minkowski metric $\eta_{\mu\nu}$ is non-degenerate but not positive-definite.*

Remark 3.2.32 We have to get rid of the positive-definite to have topologically separated lightlike, spacelike, and timelike regions.

Definition 3.2.33 The *pseudo-Euclidean metric* is

$$\delta_{\mu\nu} = \text{diag}(\pm 1, \pm 1, \pm 1, \pm 1) \quad .$$

We will also use the symbol $\delta^{\pm}_{\mu\nu}$ so that $\delta^{+}_{\mu\nu}$ is the Euclidean metric, and $\delta^{-}_{\mu\nu} = -\delta^{+}_{\mu\nu}$.

================= NOTES =================
================= NOTES =================
================= NOTES =================
URLS in .tex comments:

Theorem 3.2.34 *It is possible to associate a vector space with the negative Euclidean metric so that vectors in the resultant manifold form a representation of the Lorentz group.*

§3.3 Tangent Space and Beyond

A metric is a scalar product on a manifold's tangent bundle. In this section, we will construct the tangent bundle as the union of a manifold's tangent spaces and then examine the kinds of scalar products we can define for its tangent vectors.

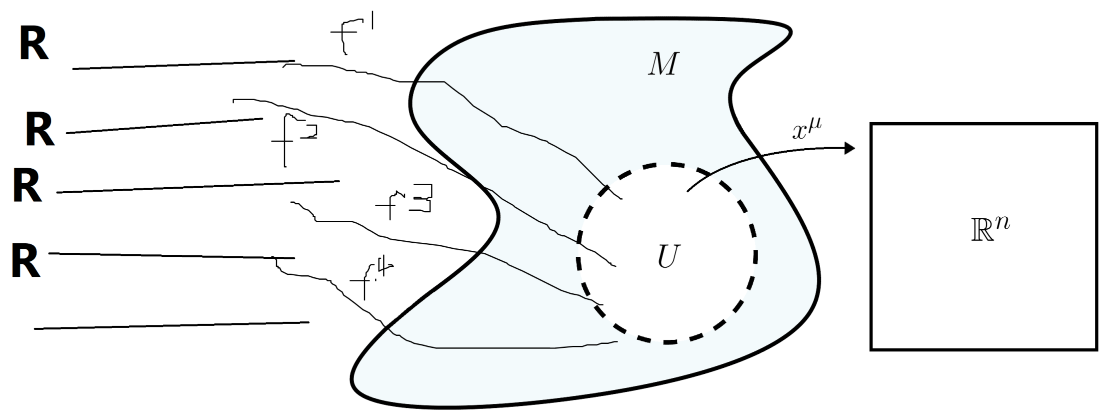


Figure 5: XXX

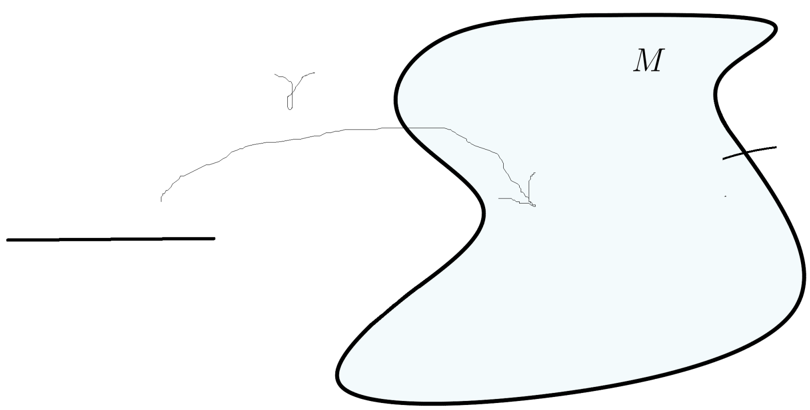


Figure 6: A curve γ is map from a interval $I \subset \mathbb{R}$ to the manifold M.

In constructing the manifold of the previous section we mostly used the mathematical language of Lee [12, 14, 15], which is well consistent with similar language in Boothby [13] and Isham [16] that we have also used. WALD??? In this section, we will continue to do so, and labor to make broader connections with the language physicists use for describing spacetime manifolds, most specifically following Penrose and Rindler [17] for an alternative way of describing things.

Penrose Vol 1, p1: Minkowski vector space, tetrads.

BRING BACK STUFF FROM APPENDIX

======================

Wald writes the following about the tangent space [11].

> "The concept of a vector space is undoubtedly familiar to most readers. In pre-relativity physics it is assumed that space has the natural structure of a three-dimensional vector space once one has designated a point to serve as the origin; the natural rules for adding and scalar multiplying spatial displacements satisfy the vector space axioms[...] In special relativity, spacetime similarly has the natural structure of a four-dimensional vector space. However, when one considers curved geometries [...], this vector space structure is lost. For example, there is no natural notion of how to 'add' two points on a sphere and end up with a third point on the sphere."

Position vectors are only inside the manifold because it is flat. Really, they are in the tangent space which overlaps with the manifold's topological space at every $p \in M$.

REWRITE BETTER FOLLOW ON:

Although physical quantities such as momentum and acceleration are not positions, they are linearly dependent on the derivatives of position, and the importance of position vectors in special relativity cannot be understated. In $\mathbb{R}^3$, two positions can be added to yield a third position in $\mathbb{R}^3$, and, since Minkowski space is an *identically flat* Lorentzian manifold, it necessarily overlaps with its tangent space at every point, and two events in the manifold of special relativity can be added to yield a third event in it. When the metric becomes non-flat, however, as will be required in certain sectors of the MCM unit cell, the sum of two vectors pointing from the origin to spacetime events x^μ and y^μ may not point to a third event z^μ. In general, it usually won't if the manifold is non-flat because every vector belongs to a Euclidean line, and non-flat means Euclidean lines won't always lie within the manifold. (One might consider curvature in a single direction, e.g.: a cylinder, as a counterexample where some lines remain in the manifold.) When the manifold is curved, there will be a unique tangent space associated with every point in the manifold: the space of straight Euclidean lines passing through the point, which is the space spanned by the tangent vectors anchored at that point. The union of a manifold's tangent spaces at every $p \in M$ will be called the tangent bundle, and it is of foremost importance in our study because the metric that makes a given manifold Lorentzian acts on the tangent spaces rather than the manifold itself. Special relativity is simply a limiting case where there is only one tangent space, or where every tangent space is the same space: the tangent space is the same topological space as the manifold itself.

==============================

Distinguish vector inner product from the manifold inner product. These things are the same in the $\mathbb{R}^4$ representation of the Lorentz group, but the two inner products will be sign inverted when we use the $\mathcal{H}$ representation. The inner product using the complex conjugate is natural to quantum theory but not really relativity, so we open the door toward new ways of describing spinors and also new tools applications toward quantum gravity. The main difference between vectors and spinors is that rotation by 2π, usually an identity operation, becomes a sign inversion operation: a spinor is an eigenfunction of OPERATOR with eigenvalue -1.

Maybe mention $M = \mathbb{C} \times \mathbb{R}^2$ as favoring the spatial dimension associated spin eigenstate representations.

WALD TALKS ABOUT CREATING PRODUCT MANIFOLDS ON p13

====================

The distance function vs the metric tensor. Since the metric space defines the topology, we must still be able to talk about 4-distance between (ct_1, x_1, y_1, z_1) and (ct_2, x_2, y_2, z_2). This is separate from the spacetime interval, but both refer to the concept of a distance between two events.

If the tangent space to a manifold at a point is a vector space, then the standard complex inner product is not equal to the Minkowski square. Is the complex inner product an irrelevant operation for physics like the distance function is? Is that in the data of the manifold that we can ignore?

Th tangent space to a manifold is vector space!

==========================

In the case of a real-valued 4-vector $x^\mu \in \mathbb{R}^4$, the inner product of x^μ with itself is a Lorentz scalar. In the case of $x^\mu \in \mathcal{H}$, the inner product, which is the dot product with the complex conjugate, or the matrix product with the conjugate transpose, the inner product is no longer a Lorentz scalar. Rather it is the contraction of $x^\mu, x^\nu \in \mathcal{H}$ with the metric which is invariant under Lorentz transformations. The key point here is that $x^\mu \in \mathbb{R}^4$ implies that the inner product is identical to contraction with the metric, but we will need to be careful to make this distinction when $x^\mu \in \mathbb{R}^4$, and we should study what will be the implications of the inner product with itself not being a Lorentz scalar.

==============================

We can assume that the manifold is embedded in a higher dimensional space because the manifold of spacial relativity is identically flat. Otherwise it would be the manifold of general relativity. The intrinsic way to do it is in Appendix A. We go over the popular ways, show they are the same, and we

propose a minor change to show a new way how to do it.
======================

Definition 3.3.1 A **smooth curve** is a coordinate-independent, continuous C^∞ map from a real interval $I \subset \mathbb{R}$ to a manifold M:

$$\gamma : I \to M \quad .$$

MAYBE MAKE A REMARK ABOUT WHAT PENROSE CALLS "AFFINE MINKOWSKI SPACE" and MINKOWSKI VECTOR SPACE

Theorem 3.3.2 *Given a smooth manifold M and a smooth curve γ with the property $\gamma(\lambda_0) = p \in M$, there exists an $\varepsilon > 0$ such that*

$$\gamma : [\lambda_0 - \varepsilon, \lambda_0 + \varepsilon] \to U \quad ,$$

where U has at least one corresponding chart (U, φ) in the atlas of M.

Proof. Definition 3.2.18 requires that the $\{U_i\}$ in the atlas of M are a cover of M, so it is guaranteed that $\gamma_k(\lambda_0) = p$ belongs to some U_i. Definition 3.3.1 grants that γ is continuous on I, so

$$\lim_{\lambda \to \lambda_0} \gamma(\lambda) = p \quad .$$

This requires $\gamma(\lambda_0 \pm \varepsilon) \to p$ as $\varepsilon \to 0$. If we suppose that $\gamma(\lambda_0 \pm \varepsilon)$ is not in U for any $\varepsilon > 0$, then U must be a one-point set because the points $\gamma(\lambda_0 \pm \varepsilon) = q \neq p$ can be made arbitrarily close to p. M is Hausdorff (Definition 3.2.6) and every one-point set in M is strictly closed, but every U in the atlas $\{(U_i, \varphi_i)\}$ is open. By contractions, therefore, the theorem is proven (and demonstrated in Figure 7.)

Definition 3.3.3 Two curves

$$\gamma_1 : I_1 \to M \quad , \qquad \text{and} \qquad \gamma_2 : I_2 \to M \quad ,$$

are **tangent** at $p \in M$ if there exists $\lambda_0 \in I_1 \cap I_2$ such that

$$\gamma_1(\lambda_0) = p \quad , \qquad \text{and} \qquad \gamma_2(\lambda_0) = p \quad ,$$

and the derivatives of the images of γ under some atlased coordinate map φ around p are equal at λ_0:

$$\frac{d}{d\lambda}\big(\varphi \circ \gamma_1\big)\bigg|_{\gamma_1(\lambda_0)} = \frac{d}{d\lambda}\big(\varphi \circ \gamma_2\big)\bigg|_{\gamma_2(\lambda_0)} \quad .$$

Theorem 3.3.4 *If two curves γ_1, γ_2 are tangent at $p \in M$, then the derivatives of their images in $\mathbb{R}^n$ are equal for any C^∞ coordinate map $\psi : M \to \mathbb{R}^n$.*

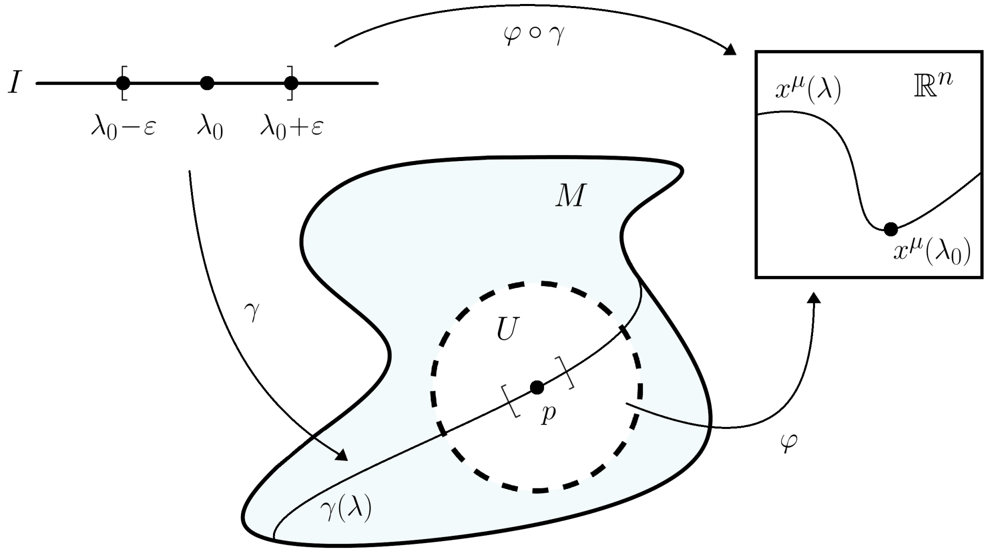


Figure 7: Theorem 3.3.2 proves the existence of the arrangement shown in this figure. Since U is an open subset of M and γ is continuous on I with the property $\gamma(\lambda_0 = p)$, we can always choose a neighborhood of I on which the image of γ in M belongs to the same U as p. This guarantees that we can take the image of γ around p to $\mathbb{R}^n$ via the map in some chart (U, φ).

Proof. Given the images of γ_1, γ_2 under φ, we obtain their images another map ψ by the transition function $\psi \circ \varphi^-1$. The derivatives of the images under ψ are

$$
\begin{aligned}
\frac{d}{d\lambda}\big(\psi \circ \varphi^{-1}\big) \circ \big(\varphi \circ \gamma_1\big) &= \frac{\partial\big(\psi \circ \varphi^{-1}\big)}{\partial\big(\varphi \circ \gamma_1\big)} \frac{d\big(\varphi \circ \gamma_1\big)}{d\lambda} \\
\frac{d}{d\lambda}\big(\psi \circ \varphi^{-1}\big) \circ \big(\varphi \circ \gamma_2\big) &= \frac{\partial\big(\psi \circ \varphi^{-1}\big)}{\partial\big(\varphi \circ \gamma_2\big)} \frac{d\big(\varphi \circ \gamma_2\big)}{d\lambda} \quad .
\end{aligned}
$$

The definition of tangency guarantees that the derivatives with respect to λ on the right are equal at p, so it remains to prove

$$
\frac{\partial\big(\psi \circ \varphi^{-1}\big)}{\partial\big(\varphi \circ \gamma_1\big)}\bigg|_p = \frac{\partial\big(\psi \circ \varphi^{-1}\big)}{\partial\big(\varphi \circ \gamma_2\big)}\bigg|_p \quad .
$$

Since $\gamma_1(\lambda_0) = \gamma_2(\lambda_0) = p$, the two derivatives will be equal at p. Since $\psi \circ \varphi^{-1}$ was an arbitrary transition function, and the C^∞ condition guarantees the well definition of the relevant derivatives, γ_1, γ_2 at p will satisfy the definition of tangency under any C^∞ coordinate map. 🍃

======================
The tangent vectors to all curves tangent at p are linearly dependent.

Applying a map to a curve gives us a real-valued function of a real variable, and we can take its derivative.

=====================

Definition 3.3.5 The tangent space at to M at p is denoted T_pM and it is the set of all tangent vectors to smooth curves passing through p:

$$T_pM = \left\{ \frac{dc}{dt}\bigg|_{t=0} \big| c \text{ a smooth curve in } M, c(0) = p \right\} \quad .$$

EXCELLENT, BRIEF DEF OF TANGENT SPACE: GOOD NOTES URL:

KKKKKKKKKKKKKKKKKKKKKKKKKK
KKKKKKKKKKKKKKKKKKKKKK
KKKKKKKKKKKKKKKKKKKKKKKKKK

Definition 3.3.6 A vector $\boldsymbol{V} \in \mathbb{R}^n$ is written

$$\boldsymbol{V} = (V^1, \dots, V^n) \quad ,$$

where the tuple notation abbreviates pairwise multiplication with a unit vector basis spanning $\mathbb{R}^n$:

$$(V^1, \dots, V^n) \equiv V^1 \hat{e}_1 + \cdots + V^n \hat{e}_n \quad .$$

CHECK BASIS VECTOR INDEX NUANCE!!!!

Remark 3.3.7 Coordinate map φ vs coordinate functions x^μ

Definition 3.3.8 An n-dimensional **vector space** $(\mathcal{V}, +, \times)$ has the following properties.

- Vectors are added component-wise in the usual way:

  $$\forall \boldsymbol{V}, \boldsymbol{W} \in \mathcal{V} \quad \exists \boldsymbol{V} + \boldsymbol{W} = (V^1 + W^1, \dots, V^n + W^n) \in \mathcal{V} \quad .$$

  The sum of two vectors in $\mathcal{V}$ is another vector in $\mathcal{V}$. The vector space is closed under addition, which is commutative and associative. The additive identity is the zero vector $\mathbf{0}$. $\mathcal{V}$ contains the zero vector and an additive inverse for all of its vectors:

  $$\forall \boldsymbol{V} \in \mathcal{V} \quad \exists \boldsymbol{V}^{-1} = -\boldsymbol{V} \in \mathcal{V} \qquad \text{s.t.} \qquad \boldsymbol{V} + \big(-\boldsymbol{V} \big) = \mathbf{0} \in \mathcal{V} \quad .$$

- The $\times$ operator is scalar multiplication. The $\times$ operator is defined on a vector paired with a scalar, and it is associative and distributive over addition.

- If the scalars are real-valued, then we say the vector space is over the field of real number. If they are complex, then it is over the field of complex numbers.

Remark 3.3.9 Due to exciting new developments in fractional distance analysis [9], it is not correct to follow the historical convention in which $\mathbb{R}$ and $\mathbb{C}$ are axiomatized as number fields. Instead, only the real and complex neighborhoods of the origins satisfy the number field axioms while the respective neighborhoods of infinity do not. Therefore, for the present purposes, we should restrict *scalar* to refer to the neighborhood of the origin. A real scalar is less than some natural number, and the real and imaginary parts of a complex scalar are less than two natural numbers. However, the existence of fractional distance analysis suggests extensions to vector analysis beyond what will be considered here.

SHOW THE CURVE DERIVATIVES ARE A VECTOR SPACE!!

KKKKKKKKKKKKKKKKKKKKKKKKKK

Definition 3.3.10 A vector $\boldsymbol{V} \in \mathbb{R}^n$ is written

$$\boldsymbol{V} = (V^1, \ldots, V^n) \quad ,$$

where the tuple notation abbreviates pairwise multiplication with a unit vector basis spanning $\mathbb{R}^n$:

$$(V^1, \ldots, V^n) \equiv V^1\hat{e}_1 + \cdots + V^n\hat{e}_n \quad .$$

Remark 3.3.11 The Lorentzian manifold $M = (\mathbb{R}^4, \eta_{\mu\nu})$ is usual manifold for special relativity. $\mathbb{R}^4$ automatically meets the definition of a manifold. It is locally Euclidean, meaning there exist homeomorphisms from the manifold's open subsets to $\mathbb{R}^4$, but we equip it with a non-degenerate, non-positive-definite metric $\eta_{\mu\nu}$ so that it becomes a pseudo-Riemannian manifold whose properties are not globally equivalent to $\mathbb{R}^4$ where distance is determined by the Euclidean metric, which exists as a consequence of the Pythagorean theorem. Since our non-degenerate metric is Lorentzian, we say the manifold of special relativity is a Lorentzian manifold. For $p \in U \subset M$, the homeomorphism that assigns p as an event in spacetime having a time and place is

$$x^\mu(p) = (ct, x, y, z) \quad .$$

However, the purpose of the present work is introduce new homeomorphisms so that

$$\widetilde{x}^\mu(p) = (ict, x, y, z) \quad , \qquad \text{or} \qquad \widetilde{x}^\mu(p) = (ct, ix, iy, iz) \quad ,$$

are coordinate maps on a different manifold $\mathcal{M}$. At first glance, and referring to Definition 3.2.15, we see that the existing coordinate functions are

$$x^0(p) = ct \qquad x^1(p) = x \qquad x^2(p) = y \qquad x^3(p) = z \quad ,$$

so the new functions will be

$$\widetilde{x}^0(p) = ix^0(p) \qquad \widetilde{x}^1(p) = x^1(p) \qquad \widetilde{x}^2(p) = x^2(p) \qquad \widetilde{x}^3(p) = x^3(p) \quad ,$$

or

$$\widetilde{x}^0(p) = x^0(p) \qquad \widetilde{x}^1(p) = ix^1(p) \qquad \widetilde{x}^2(p) = ix^2(p) \qquad \widetilde{x}^3(p) = ix^3(p) \quad .$$

This extension is trivial. However, even while a homeomorphism $f : S_1 \to S_2$ is not restricted to be a real-valued function (Definition 3.2.1), the coordinate functions x^μ in chart (U, x^μ), where U is a particular Lorentz frame in spacetime (which can cover the whole manifold due to the lack of curvature in special relativity), are restricted as $x^\mu : M \to \mathbb{R}^n$ (Definition 3.2.15). At this point, we have to add more data to the manifold to sort things out, or we have to use another sort of manifold.

A complex manifold is a direct extension of a real manifold, which is what we have defined as a *manifold* in the previous section. Rather than being locally Euclidean by the existence of homeomorphisms (coordinate maps) to subsets of $\mathbb{R}^n$, the homeomorphisms map to $\mathbb{C}^n$. A complex vector $\boldsymbol{z} \in \mathbb{C}^4$ is uniquely identified with eight real numbers, four real parts and four imaginary parts, but the $\widetilde{x}^\mu$ are describable with just four real numbers, a time and three positions, so the most general theory of complex manifolds will probably be more complicated than what is needed. However, with due rigor, a $\widetilde{x}^\mu$ is not a coordinate function on a Lorentzian manifold due to the presence of the imaginary number i in

$$\widetilde{x}^\mu : \mathcal{M} \to i\mathbb{R} \times \mathbb{R}^3 \quad , \quad \text{or} \qquad \widetilde{x}^\mu : \mathcal{M} \to \mathbb{R} \times i\mathbb{R}^3 \quad .$$

Furthermore, to preserve the structure of special relativity with timelike, spacelike, and null 4-vectors, the metric tensor in the data of $\mathcal{M}$ can no longer be the Lorentzian one:

$$\widetilde{x}^2_M = \eta_{\mu\nu}\widetilde{x}^\mu\widetilde{x}^\nu = \pm\left[\left(ct\right)^2 + x^2 + y^2 + z^2\right] \quad .$$

The Minkowski square is positive-definite or negative-definite. There are no lightlike vectors in $(\mathcal{M}, \eta_{\mu\nu})$, and vectors are either all spacelike or all timelike. Clearly, this cannot be the manifold of special relativity. In this section, we will spell out the main differences between the real and complex coordinate functions, and then we will find the appropriate manifold in the following section.

KKKKKKKKKKKKKKKKKKKKKKKKKKKKKK

Definition 3.3.12 A **scalar product space** is a vector space equipped with an scalar product $\langle \boldsymbol{V}, \boldsymbol{W} \rangle$. The inner product is also called a scalar product because the inner product of two vectors is a scalar. A scalar product space is a vector space $\mathcal{V}$ equipped with its dual space $\mathcal{V}^*$ containing all linear maps from $\mathcal{V}$ to a given scalar field. If $\boldsymbol{V}$ is a vector $\boldsymbol{V} \in \mathcal{V}$, and if $\boldsymbol{W}(\boldsymbol{V}) = \langle \boldsymbol{V}, \boldsymbol{W} \rangle$ is a scalar product, then $\boldsymbol{W}$ is called a **dual vector** and it has the property $\boldsymbol{W} \in \mathcal{V}^*$.

INNER PRODUCT IS POSITIVE DEFINITE!! $\langle \boldsymbol{V}, \boldsymbol{V} \rangle \geq 0$ guarantees vectors have positive length

DIFFERENTIATE FROM SCALAR PRODUCT!!

Inner product space: URL

Definition 3.3.13 The **dot product** is denoted with $\{\cdot\}$. For two vectors $\boldsymbol{V}, \boldsymbol{W} \in (\mathcal{V}, +, \times, \cdot)$, we have

$$\big\langle \boldsymbol{V}, \boldsymbol{W} \big\rangle = \boldsymbol{V} \cdot \boldsymbol{W} = (V^1, \ldots, V^n) \cdot (W^1, \ldots, W^n) = \sum_{k=1}^{n} V_k W_k \quad .$$

Example 3.3.14 An inner product defines a metric tensor. As a bilinear form, a metric $g_{\mu\nu}$ is a map from two vectors to real numbers, so the contraction of two vectors with the metric is automatically an inner product:

$$g_{\mu\nu}V^\mu W^\nu = V^\mu W_\mu \quad \equiv \quad \big\langle \boldsymbol{V}, \boldsymbol{W} \big\rangle = \boldsymbol{W}(\boldsymbol{V}) \ .$$

By inspection, the dot product defines $g_{\mu\nu} = \text{diag}(1,1,1,1)$, which is the Euclidean metric. The dot product is the metric inner product in Euclidean space. In this case, there is no difference between a vector with an upper index or one with a lower index. Similarly, the dot product allows us to represent $\boldsymbol{V}$ and $\boldsymbol{W}$ as column vectors or row vectors:

$$\big\langle \boldsymbol{V}, \boldsymbol{W} \big\rangle = (V^1, \ldots, V^n) \cdot (W^1, \ldots, W^n) = \begin{pmatrix} V^1 \\ V^2 \\ V^3 \\ V^4 \end{pmatrix} \cdot \begin{pmatrix} W^1 \\ W^2 \\ W^3 \\ W^4 \end{pmatrix} \ .$$

If we want to write the dot product in terms of matrix multiplication, however, then one must be a column matrix and the other a row matrix:

$$\big\langle \boldsymbol{V}, \boldsymbol{W} \big\rangle = \begin{bmatrix} W^1 & W^2 & W^3 & W^4 \end{bmatrix} \begin{bmatrix} V^1 \\ V^2 \\ V^3 \\ V^4 \end{bmatrix} \ ,$$

The usual convention for tensor notation is that the vector V^μ is a column matrix and the dual vector W_μ is the row. Since the metric is a symmetric matrix, the tensor expression $g_{\mu\nu}V^\mu W^\nu$ is such that $V^\mu W_\mu = V_\nu W^\nu$: we can raise or lower either index, which is equivalent to writing $\boldsymbol{W}(\boldsymbol{V}) = \boldsymbol{V}(\boldsymbol{W})$. In the Euclidean manifold implied by the dot product (every finite dimensional vector space is a manifold), the components of a vector and its dual are the same, but this is the only case in which will be true.

Definition 3.3.15 The metric inner product is

$$\big\langle \boldsymbol{V}, \boldsymbol{W} \big\rangle_M = ??????????$$

MINKOWSKI SQUARE

Remark 3.3.16 We call the metric scalar product the Minkowski product even when the space is not Minkowski space. The locally Euclidean property of manifolds means that space looks like Minkowski space on small scales.

COMPLEX SCALAR PRODUCT WITHOUT CONJUGATION IN WILLIAMS NOTES [18]

Example 3.3.17 Hilbert space is an inner product space.

Example 3.3.18 Some kind of matrix inner product space.

Example 3.3.19 The partial derivatives with respect to the coordinate functions span the vector space.

BANACH SPACE

Remark 3.3.20 MORE THAN ONE KIND OF INNER PRODUCT IS POSSIBLE!!! OFTEN ONE SPEAKS OF **THE** INNER PRODUCT, SO DISTINCTIONS MUST BE MADE. WE WILL CALL THE PRODUCT THAT DETERMINES THE MINKOWSKI SQUARE, "THE MINKOWSKI PRODUCT"

Definition 3.3.21 In tensor analysis, the components of a vector are strictly written with an upper index; the entries of the n-vector $\vec{x} = x^k \hat{e}_k$ are x^k. As with the parenthesis notation, by convention we omit the basis vectors and focus on the components x^k. The inner product of $x^k, y^j \in \mathbb{R}^4$ is their contraction with the Euclidean metric γ_{ij}:

$$\langle x^k, y^j \rangle \equiv x^k y^j \gamma_{kj} = x^1 k^1 + x^2 k^2 + \cdots + x^n y^n = \sum_{k=1}^{n} x_k y_k \quad .$$

This is exactly equal to the dot product. Such vectors live in the tangent bundle to a Riemannian manifold.

Example 3.3.22 The dot product of two 4-vectors in not a Lorentz scalar.

Lorentz scalars are uniquely associated with the Lorentzian metric.

Definition 3.3.23 The inner product of complex vectors $\vec{u}, \vec{v} \in \mathbb{C}$ is defined as

$$\langle \vec{u}, \vec{v} \rangle \equiv x^k y^j \gamma_{kj} = x^1 k^1 + x^2 k^2 + \cdots + x^n y^n \quad .$$

The inner product for $x^\mu \in \mathcal{H}$ is defined as

$$\left(x^\mu\right)^2 \equiv |x^\mu|^2 = \left(x^\mu\right)^* x_\mu = \left(x^\mu\right)^* \eta_{\mu\nu} x^\nu$$

ETA/GAMMA?????

Remark 3.3.24 Inner product is positive definite. Since we want timelike, null, and spacelike vectors, this will not suffice. We need another product.

GENERAL BILINEAR FORM

COPY CARROLL p23

If $x^\mu \in \mathbb{R}^4$, then the action of the inner product on two vectors is the inner product. The inner product of two complex vectors $x^\mu \in \mathcal{H}$ will be positive definite when it is defined in the usual way

$$\big(x^\mu\big)^2=\sum_\mu\big(x^\mu\big)^*x^\mu$$

Definition 3.3.25 An **axial vector** V^μ is invariant under the parity operation: $\boldsymbol{P}V^\mu=V^\mu$. Axial vectors are also called pseudovectors.

Definition 3.3.26 A **polar vector** V^μ acquires a sign change under parity: $\boldsymbol{P}V^\mu=-V^\mu$.

Example 3.3.27 Polar and axial vectors. Position is polar, ang mom is axial. Show the reason.

KKKKKKKKKKKKKKKKKKKK

FOLLOW CARROLL: Maps between manifolds. There exists a function of a complex variable that takes $\widetilde{x}^\mu\in\mathbb{R}^4$ to $x^\mu\in\mathcal{H}$.

KKKKKKKKKKKKKKKKKKKKKKKK

Article 3.3.28 Now we will consider a complex 4-vector as

$$x^\mu=(it,x,y,z)\quad,\quad\text{or}\qquad x^\mu=(t,ix,iy,iz)$$

The Lorentz invariant scalar product still

$$\big(x^\mu\big)^2=x^\mu x_\mu=x^\mu\big(\eta_{\mu\nu}x^\nu\big)\quad,$$

but now the metric is the 4D Euclidean metric

$$g_{\mu\nu}=\gamma_{\mu\nu}=\text{diag}(\pm1,\pm1,\pm1,\pm1)\quad.$$

For x^μ with imaginary time part, the inner product is computed as

$$\begin{aligned}x^\mu\big(\gamma_{\mu\nu}x^\nu\big)&=(it,x,y,z)\cdot(\pm it,\pm x,\pm y,\pm z)\\&=\mp t^2\pm\big(x^2+y^2+z^2\big)\end{aligned}$$

There exists a null interval in the convention for complex 4-vectors with imaginary time parts. For complex 4-vectors with imaginary space parts, we have

$$\begin{aligned}x^\mu\big(\gamma_{\mu\nu}x^\nu\big)&=(t,ix,iy,iz)\cdot(\pm t,\pm ix,\pm iy,\pm iz)\\&=\pm t^2\mp\big(x^2+y^2+z^2\big)\end{aligned}$$

The freedom to choose the space part or the time part as imaginary reflects the freedom to choose the metric signature from $\{\pm\mp\mp\mp\}$ in the previous article. However, since we still have freedom to choose the sign convention of the Euclidean metric from $\{\pm\pm\pm\pm\}$, there should be some new physical freedom associated. We will investigate this.

Remark 3.3.29 In the conventions of Carroll [2], any metric with all the same signs is "Euclidean," and any metric with one sign different from the others "Lorentzian." However, the flat metric with signature $\{----\}$ is not positive-definite, and the Pseudo-Riemannian manifold $(\mathbb{R}^4, \text{diag}(-1,-1,-1,-1))$ is not identically $\mathbb{R}^4$. For lack of a better word, we will call this metric *pseudo-Euclidean.*

Now 4-vectors' first component specifies a time, and the other three specify a position vector $\boldsymbol{x}$ in either Euclidean 3-space or pseudo-Euclidean 3-space.

Remark 3.3.30 MAYBE DEFINE A 4-VECTOR DOT PRODUCT

Remark 3.3.31 • We will use $\langle , \rangle$ to denote the standard scalar product:

$$\langle x, y \rangle = \sum_{\mu=0}^{3} x_\mu y_\mu$$

• We will use $\langle , \rangle_M$ to denote the "Minkowski" scalar product:

$$\langle x, y \rangle_M = \langle x, gy \rangle = \sum_{\mu,v=0}^{3} x_\mu g_{\mu\nu} y_\nu = x_0 y_0 - \vec{x} \cdot \vec{y}$$

Note that this is the matrix version of lower and index.

=====================

FORMS: A dual vector is a one-form. I need a symplectic 2-form at space-like infinity, and maybe I can get one with $\mathbb{R}^{1,3} = \mathbb{C} \times \mathbb{R}^2$.

SESQUILINEAR FORMS: URL

"The characteristic quadratic form on M" is $(\Delta s)^2$ from ZEEMAN

Show that the usual topology already has a metric, but we add a metric tensor as a different thing, and we want to use the complex vectors to use the same metric in the topology

"he modulus of the complex number a + bi is $|a + bi| = \sqrt{a2 + b2}$. This is the distance between the origin (0, 0) and the point (a, b) in the complex plane"

So, since have the modulus, this isn't the Euclidean metric.

Distinguish metric from metric tensor.

"The real-complex biplane topology $\mathbb{C} \times \mathbb{R}$." (my words)

KKKKKKKKKKKKKKKKKKKKKKKKKKKKKK

Remark 3.3.32 GREAT NOTES URL :

Remark 3.3.33 WALD p364: "If T(M) is trivial, we say M is **parallelizable**."

Remark 3.3.34 "We are concerned with the case where VU is a Lorentz manifold, i.e., a differentiable manifold of some class CP, 2 ¡ p ¡ a, carrying a second order symmetric covariant tensor field of signature (2 - n) and class C2."

"V is said to be Lorentz orientable if the bundle of time-like tangent vectors is disconnected."

IMPORTANT: !!!!!!!!!!! https://www.ncbi.nlm.nih.gov/pmc/articles/PMC285022/pdf/pnas0020
0124.pdf

Remark 3.3.35 ======================== BOOTHBY: p181

Riemannian manifolds make the tangent space to any point $p \in M$ into a Euclidean space.

========================

Every tangent space of a pseudo-Riemannian manifold is a pseudo-Euclidean vector space.

========================

BOOTHBY Thm V.3.1

Remark 3.3.36 A vector manifold is a set of vectors that generate tangent spaces to the manifold.

A vector field is a function that assigns a tangent vector to each point of the manifold.

A vector space can be regarded as a manifold, but not every manifold is a vector space.

the dot product on a vector space implies the Euclidean metric, but matrix spaces can also be vector spaces

`https://www.bing.com/search?q=is+every+vector+space+a+manifold`

Definition 3.3.37 WALD: A tensor T of type (k, l) over $\mathcal{V}$ is a multilinear map

$$T : \underbrace{\mathcal{V}^* \times \cdots \times \mathcal{V}^*}_{k \text{ copies}} \times \underbrace{\mathcal{V} \times \cdots \times \mathcal{V}}_{l \text{ copies}} \rightarrow \mathbb{R} \quad .$$

p21: The collection of all T(k,l) tensors has the structure of an n-dimensional vector space. The dimension of the space is n^{k+l}.

For two tensors, there exists an outer product $T \otimes T'$

Definition 3.3.38 **Contravariant** means
Covariant means

Definition 3.3.39 An assigment of a tensor over V_p for every $p \in M$ is called a **tensor field**.

Definition 3.3.40 WALD Metric is supposed to tell us "infinitesimal squared distance" associated with an "infinitesimal squared displacement." Wald has already associated tangent vectors with infinitesimal displacements, so the metric tensor should depend on a product of two tangent vectors. The metric is symmetric and non-degenerate. Metric g on manifold M is a symmetric, non-degenerate tensor field of type $(0, 2)$. WALD p23: expand the metric in a coord basis.

Definition 3.3.41 There is always an **orthonormal basis** in which the metric at p may be written as the flat metric: $g(v_\mu, v_\nu) = \pm 1$. The number of plus and minus signs is called the signature of the metric.

Remark 3.3.42 Since the metric is one-to-one and onto, and non-degenerate with an inverse, we can establish a one-to-one correspondence between vectors and dual vectors.

Article 3.3.43 Penrose and Rindler [17] have a nice description of bundles (p332). We call the manifold M "affine" because "curves" have affine parameters. Penrose and Rindler have to specify that M has an origin chosen because an origin is automatically chosen under some coordinate map. The explanation how the graph becomes geometric sets up for a nice statement about how we have defined our derivatives of curves with the Leibniz rule. By following Carroll [2], (CHECK WALD?), we can show that the derivatives are the directional derivatives, but there is this other layer where there is geometry rather than numbers. In computing one of the derivatives, we can quote Zeeman [19] saying that there was already another metric due to the topology, and then show that SC was wrong when he said $q - p$ is undefined.

Definition 3.3.44 A **spanning basis** for a vector space is...

Remark 3.3.45 In the analysis of $\mathbb{R}^3$, one often speaks of an "orthogonal triad" $\{\hat{x}, \hat{y}, \hat{z}\}$. This is a spanning basis for $\mathbb{R}^3$. The cylindrical and spherical polar triads $\{\hat{r}, \hat{\theta}, \hat{z}\}$ and $\{\hat{r}, \hat{\theta}, \hat{\phi}\}$ are also spanning bases for $\mathbb{R}^4$. In Minkowski space, the three-element orthogonal triad becomes a four-element tetrad, in which orthogonality is included in the definition of a tetrad. We say a tetrad is a Minkowski tetrad if

$$\boldsymbol{g}_\mu \cdot \boldsymbol{g}_\nu = \eta_{\mu\nu} \quad .$$

§3.4 Symplectic, Hermitian, and Kähler Manifolds

Remark 3.4.1 xx

> "As as a rule of thumb, proofs that depend only on the invertibility of the metric tensor, such as existence and uniqueness of the Riemannian connection and geodesics, work fine in the pseudo-Riemannian setting, while proofs that use positivity in an essential way, such as those involving distance-minimizing properties of geodesics, do not."

So, if anything more than an uncountable number of sign errors related to terms like $e^{ik^\mu \cdot x^\mu}$ and $\Box^2\psi$ is lost in going to the $\mathcal{H}$ representation from $\mathbb{R}^4$ (unlikely), we know that somethings are gained.

> "Since holomorphic functions are much more rigid than smooth functions, the theories of smooth and complex manifolds have very different flavors: compact complex manifolds are much closer to algebraic varieties than to differentiable manifolds."
> `https://en.wikipedia.org/wiki/Complex_manifold`

Remark 3.4.2 "One can define an analogue of a Riemannian metric for complex manifolds, called a Hermitian metric. Like a Riemannian metric, a Hermitian metric consists of a smoothly varying, positive definite inner product on the tangent bundle, which is Hermitian with respect to the complex structure on the tangent space at each point. As in the Riemannian case, such metrics always exist in abundance on any complex manifold. If the skew symmetric part of such a metric is symplectic, i.e. closed and non-degenerate, then the metric is called Kähler. Kähler structures are much more difficult to come by and are much more rigid."
`https://en.wikipedia.org/wiki/Complex_manifold`

Since Kahler manifold is a symplectic manifold, that will let me get around the positive-definiteness constraint in the ADM theorem.

> "A Hermitian metric h on an (almost) complex manifold M defines a Riemannian metric g on the underlying smooth manifold. The metric g is defined to be the real part of h:
>
> $$g = \frac{1}{2}\big(h + \bar{h}\big)$$
>
> `https://en.wikipedia.org/wiki/Hermitian_manifold`

This is already how I have the metric in $\mathcal{H}$ defined as a sum of contributions from $\Sigma^{\pm}$. We will want to show that the $\{+-\underline{-}-\pm\}$ discrepancy is associated with h and its complex conjugate $\bar{h}$.

==============================

fundamental to Lebesgue integration, which relies on the ability to express a function as a difference of two positive functions $f = f^+ - f^-$ where f^+ denotes the positive part of f and $-f^-$ the negative part.

39: Bourbaki 2004, ch2, p48

================================

> "In mathematics, particularly in complex analysis, a Riemann surface is a one-dimensional complex manifold.
>
> Loosely speaking, this means that any Riemann surface is formed by gluing together open subsets of the complex plane C using holomorphic gluing maps.
>
> Examples of Riemann surfaces include graphs of multivalued functions like $\sqrt{z}$ or $\log(z)$, e.g. the subset of pairs $(z, w) \in \mathbb{C}^2$ with $w = \log(z)$.
>
> Every Riemann surface is a surface: a two-dimensional real manifold, but it contains more structure (specifically a complex structure). Conversely, a two-dimensional real manifold can be turned into a Riemann surface (usually in several inequivalent ways) if and only if it is orientable and metrizable. So the sphere and torus admit complex structures, but the Möbius strip, Klein bottle and real projective plane do not.
>
> Every compact Riemann surface is a complex algebraic curve by Chow's theorem and the Riemann–Roch theorem.
>
> Riemann surfaces were first studied by and are named after Bernhard Riemann."
>
> `https://en.wikipedia.org/wiki/Riemann_surface`

A Riemann surface is complex 1D, so $\mathcal{M}$ cannot be a Riemann surface. However, if we define $\mathcal{M}$ as the product of a Riemann surface with the Cartesian plane $\mathbb{R}^2$, then we have picked out one special direction in space such as the one picked out when a basis quantum mechanical eigenspinors. Furthermore, the asymmetry of the fermionic wavefunction, like the anomalous rotation properties of spinors, depends on nothing more than a minus sign such as the one which differentiate the complex inner product $|x^2| = \mathsf{X}^\dagger \mathsf{X}$ from the Minkowski square $x_M^2 = \mathsf{X}^T \boldsymbol{g} \mathsf{X}$. In fact, a complex inner product space is a Hilbert space, the realm of quantum mechanics, so we should explore what possible correspondences might arise.

"If (M,g) is a time-orientable smooth Lorentzian manifold"

What can we say about orientability and the cosmological principle when one spatial dimension is picked out as belonging to $\mathbb{C}^4$?

> "The general form of an inner product on $\mathbb{C}^n$ is known as the Hermitian form and is given by
>
> $$\langle x, y \rangle = y^\dagger \boldsymbol{M} x = \overline{x^\dagger} \boldsymbol{M} y$$

where M is any Hermitian positive-definite matrix and $y^\dagger$ is the conjugate transpose of y. For the real case, this corresponds to the dot product of the results of directionally-different scaling of the two vectors, with positive scale factors and orthogonal directions of scaling. It is a weighted-sum version of the dot product with positive weights—up to an orthogonal transformation."

`https://en.wikipedia.org/wiki/Inner_product_space`

===================================

Remark 3.4.3 POINCARE CONJECTURE: This was the foundation for much of the program in the MCM anyways.

Remark 3.4.4 Finite extinction time for the solutions to the Ricci flow on certain three-manifolds by Grisha Perelman

https://arxiv.org/pdf/math/0307245.pdf

IMPORTANT FOR ∅???

[A-G] S.Altschuler, M.Grayson Shortening space curves and flow through singularities. Jour. Diff. Geom. 35 (1992), 283-298.

§3.5 Algebraic Topology and the Fundamental Group of a Manifold

Remark 3.5.1 "We are concerned with the case where VU is a Lorentz manifold, i.e., a differentiable manifold of some class CP, 2 ¡ p ¡ a, carrying a second order symmetric covariant tensor field of signature (2 - n) and class C2."

"V is said to be Lorentz orientable if the bundle of time-like tangent vectors is disconnected."

IMPORTANT: !!!!!!!!!!! https://www.ncbi.nlm.nih.gov/pmc/articles/PMC285022/pdf/pnas0020
0124.pdf

Remark 3.5.2 "ON THE CATEGORICAL AND TOPOLOGICAL STRUCTURE OF TIMELIKE AND CAUSAL HOMOTOPY CLASSES OF PATHS IN SMOOTH SPACETIMES"

"Principle of Topological Censorship (PTC). Every causal curve whose initial and final endpoints belong to I is fixed endpoint homotopic to a curve on I. "

URL: https://arxiv.org/pdf/2008.00265.pdf

To attach group theory to the study of manifolds that we have developed, we will use some algebraic topology.

Definition 3.5.3 A **group** is

Definition 3.5.4 A **path homotopy** is

Definition 3.5.5 The fundamental group of a topological space is the homotopy class of loops at basepoint x_0 in X with binary operation $[f][g] = [f * g]$. This group is denote $\pi_1(X, x_0)$.

URL:

Paths in one of our manifolds depend on the line element. We should show that paths in the Euclidean line element are not homotopic to paths in the Minkowski line element. Then we should show the homotopy of paths in $(\mathbb{R}^4, \eta_{\mu\nu})$ and $(\mathcal{M}, \delta_{\mu\nu})$.

Remark 3.5.6 There is only one homotopy class in any convex manifold!

Remark 3.5.7 Causality Implies the Lorentz Group

URL:

THE TOPOLOGY OF MINKOWSKI SPACE

URL:

A new topology for curved space–time which incorporates the causal, differential, and conformal structures

URL:

FUNDAMENTAL GROUP AND FINE TOPOLOGY ON MINKOWSKI SPACE

URL:

Article 3.5.8 MAYBE STICK THE COMPLEX-ADAPTED ZEEMAN (FINE) TOPOLOGY IN HERE.

The topology is on the manifold, not the tangent space, so we will want to go to a complex manifold rather than fudging it with complex coordinate maps.

§3.6 Remarks on Pure Mathematics

Now that all of the preceding material is established with mathematics, it would have sufficed for physics to point out that the manifolds $(\mathbb{R}^4, \eta_{\mu\nu})$ and $(\mathcal{M}, \delta_{\mu\nu})$ are the same because they are both 4D and their line elements are the same. It have been noted briefly that the fundamental representation of O(4) is a vector in $\mathbb{R}^4$: the space spanned by four orthogonal, real-valued directions, and the O(1,3) notation means the fundamental representation is a vector in $\mathbb{R}^{1,3}$: the space spanned by three orthogonal, real-valued directions,and a fourth orthogonal imaginary one. Clearly something has gotten out of order when we're using $\mathbb{R}^4$ with a convoluted metric as the group representation when we could be using the fundamental representation with the Euclidean metric.

FUNDAMENTAL GROUP IS UNRELATED TO FUNDAMENTAL REP OF GROUP!!

For the flavor of many details that we will not be getting into, consider the following from ALGTOP HANDBOOK p6

"LONG QUOTE"

What's this about Kahler manifolds?

DEF 3.8 CHAIN COMPLEXES: When the paths across the unit cell are elliptic curves, possibly transfinite ones, the task of classifying the path homotopy classes seems rather daunting.

HHANBOOK p7: The program is too hard and it didn't get worked out yet.

Still, Two things joined at a point is the usual example in algebraic topology and we want to examine $\Sigma^{\pm}$ joined on $\varnothing$ where dimensional analysis and comparison with string theory suggests that the two topological spaces should be joined on a point, and there does not exist any open set separating them in the fifth direction.

============================

Massey has an entire chapter devoted to the fundamental group of the union of two spaces. He says that the thing we want is the "free-est possible group we can use to promote one diagram to a more complicated one." Lacking the specific language of algebraic topology at the time, this was what was suggested more or less with the two stars attached to the other two stars [20,21]. The Poincare group is know to be icosahedral, which is an appropriate product of a pentagram and a hexagram. To "promote" the diagram to a more complicated one, we added to septagrams with obtuse and acute angles. In the development of the MCM that followed, we don't want to attach $\Sigma^{\pm}$ only at $\mathcal{H}$, but also at $\varnothing$, and there needs to be a loop running between the two, so one septagram contains the other. Furthermore, one $\Sigma^{\pm}$ is a manifold and the other is a manifold with a boundary, we can view the acute septagram as the boundary around the obtuse one. Rather than drawn as a commutative diagram, the figure was meant as a chart with an implied adjacency matrix, and then once that is established we might begin to develop homotopy classes for paths between nodes. Whatever Massey means by "free-est possible," it was suggested that the similar thing for expanding the Poincare group in two different ways would be two different sevens stacked on top of a four or a five. Luckily, Massey didn't get kicked out of college by people who thought he was stupid before he could learn the advanced material relevant to his interests, and he was able to offer a concrete definition of "free-est." From this writer's perspective, the exact words needed to describe the idea remain elusive to some degree. The order of a group BLAH BLAH BLAH, so we expand the group of Poincare transformations (inhomogeneous Lorentz transformations) in spacetime to the full MCM until cell where holographic duality BLAH BLAH. They didn't find the AdS/CFT correspondence until the 1990's, that's still not fully understood, and it only refers to one of $\Sigma^{\pm}$ because only one is AdS. BLAH BLAH BLAH.

CITE the similar language of using numbers to define group representations.
==========================

So, perhaps we might suggest new research! Instead of looking at algebraic topology, or even modern algebra, we will consider a smallish set of matrices.

§4 Lorentz Transformations of Spinors

Lorentz developed a set of transformations to describe the symmetries of the electromagnetic theory. Soon after, Einstein showed with special relativity that absolute notions of length and simultaneity must be set aside if Maxwell's equations are to be respected. To wit, if one takes the curl of the Maxwell–Faraday equation

$$\boldsymbol{\nabla}\times\boldsymbol{E}=-\frac{\partial\boldsymbol{B}}{\partial t}\quad ,$$

with the identity

$$\boldsymbol{\nabla}\times\big(\boldsymbol{\nabla}\times\boldsymbol{V}\big)=\boldsymbol{\nabla}\big(\boldsymbol{\nabla}\cdot\boldsymbol{V}\big)-\nabla^2\boldsymbol{V}\quad ,$$

where

$$\nabla^2\boldsymbol{V}=(\nabla^2V_x,\nabla^2V_y,\nabla^2V_z)\quad ,$$

and the Laplacian ∇^2 is the divergence of the gradient, we obtain

$$\boldsymbol{\nabla}\big(\boldsymbol{\nabla}\cdot\boldsymbol{E}\big)-\nabla^2\boldsymbol{E}=-\frac{\partial\boldsymbol{\nabla}\times\boldsymbol{B}}{\partial t}\quad .$$

If we further substitute into this curl expression Gauss' law and the Ampere–Maxwell law, namely

$$\boldsymbol{\nabla}\cdot\boldsymbol{E}=\frac{\rho}{\varepsilon_0}\quad ,\qquad\text{and}\qquad \boldsymbol{\nabla}\times\boldsymbol{B}=\mu_0\left(\boldsymbol{J}+\varepsilon_0\frac{\partial\boldsymbol{E}}{\partial_t}\right)\quad ,$$

and if we impose the free space conditions $\rho=0$ and $\boldsymbol{J}=0$, then we find

$$\nabla^2\boldsymbol{E}=\varepsilon_0\mu_0\frac{\partial\boldsymbol{E}}{\partial t}\quad .$$

This is the wave equation for a wave with velocity such that $|\boldsymbol{v}|=\sqrt{\varepsilon_0\mu_0}$, which is the speed of light c. Since Maxwellian EM waves in free space propagate at this speed in every reference frame, Einstein showed that special relativity is a necessary consequence, and much to-do has been had since then.

In special relativity, Lorentz transformations preserve objects' Minkowski length. Often one associates 4-vector $x^\mu\in\mathbb{R}^4$ with a spacetime event, but

$$\forall x^\mu=\begin{bmatrix}t\\x\\y\\z\end{bmatrix}\qquad\exists\widehat{x}=\begin{bmatrix}t+z & x-iy\\ x+iy & t-z\end{bmatrix}\qquad\text{s.t.}\qquad x_M^2=\det\widehat{x}\quad .$$

Thus, we may speak of vectors x^μ or matrices $\widehat{x}$ as representations of the Lorentz group. Linear transformations in the form

$$x^\mu\to x^{\mu'}={\Lambda^{\mu'}}_{\mu}x^\mu\quad ,\quad\text{or}\qquad \widehat{x}\to\widehat{x}'=\boldsymbol{\Lambda}\widehat{x}\quad ,$$

will both be said to be Lorentz transformations because they preserve the Minkowski length:

$$x_M^2 = \left(x'\right)_M^2 \quad , \qquad \text{or} \qquad \det \widehat{x} = \det \widehat{x}' \quad .$$

However, the elements of the Lorentz group (the Lorentz transformations) will take different forms in one representation or another. $\mathbf{\Lambda}$ acting a 4-vector is necessarily a 4×4 matrix, and it is necessarily a 2×2 matrix when it acts on $\widehat{x}$. One says there is a *realization* of the group corresponding to each *representation*: real-valued 4-vectors discussed in Section 4.2, 2×2 matrices that are the spinor representation discussed in Section 4.4, or the complex-valued 4-vector that we will examine in Section 4.5. As an example of the differences among the realizations, consider that Lorentz transformations on $\widehat{x}$ preserve the determinant of a 2×2 matrix, but the determinant is not defined at all for a 4-vector x^μ. After examining a few specific realizations, we will study the universal, representation-independent properties of Lorentz transformations as an abstract algebraic group. The material in Sections 4.1, 4.2, and 4.4 mostly follows Jaffe [22] and Steane [5]. When we opt for tensor notation rather than matrix notation, we will also follow a convention from Carroll [2] that primed indices represent Lorentz transformed quantities. Material in Section 4.5 follows Roman [3].

§4.1 Fundamentals of Lorentz Transformations

CHECK THIS OTHER PAPER TOO:
URL:
========================

Definition 4.1.1 A **linear transformation** is one that can be represented as a matrix multiplication operation on a vector. Given a pair of vector spaces $\mathcal{V}, \mathcal{W}$, $T : \mathcal{V} \to \mathcal{W}$ is a linear transformation if the following properties are satisfied.

- T preserves additivity: $T(\boldsymbol{V} + \boldsymbol{W}) = T(\boldsymbol{V}) + T(\boldsymbol{W})$.
- T preserves scalar multiplication: $T(c\boldsymbol{V}) = cT(\boldsymbol{V})$. The important case of $c = 0$ shows that the zero vector is mapped to the zero vector: $T(\mathbf{0}) = \mathbf{0}$.

Definition 4.1.2 An **endomorphism** on a vector space $\mathcal{V}$ is a map $f : \mathcal{V} \to \mathcal{V}$. Endomorphisms map the vectors in a vector space to other vectors in the vector space. If f is invertible such that there exists $f^{-1} : \mathcal{V} \to \mathcal{V}$ such that $f \circ f^{-1} = f^{-1} \circ f$ is the identity, then f is called an **automorphism**.

Definition 4.1.3 A **Lorentz transformation** is a matrix $\mathbf{\Lambda}$ that is an endomorphism in the tangent bundle of a spacetime manifold of special relativity:

$$\mathbf{\Lambda} : T_pM \to T_pM \quad .$$

The set of all $\mathbf{\Lambda}$ is denoted $\mathcal{L}$. We use the notation $\mathcal{L}(\mathbb{R})$ and $\mathcal{L}(\mathbb{C})$ to distinguish the cases in which the matrix elements $\mathbf{\Lambda}_{\mu\nu}$ are real or complex. $\mathcal{L}$ is the set of all matrices that preserve the Minkowski square, or the metric scalar product. If X is the column matrix representation of a 4-vector x^μ observed in frame S, and X' is the column matrix representation of the same 4-vector Lorentz transformed to another Lorentz frame S' (called $x^{\mu'}$), then we write

$$\mathsf{X}' = \mathbf{\Lambda}\mathsf{X} \quad , \quad \text{or} \qquad x^{\mu'} = \Lambda^{\mu'}{}_{\mu} x^\mu \quad .$$

Preservation of the Minkowski length is expressed as

$$\mathsf{X}_M^2 = \left(\mathsf{X}'\right)_M^2 \quad , \quad \text{and} \qquad x_M^2 = x_M'^2 \quad .$$

Any $\mathbf{\Lambda}$ that preserves the Minkowski product of two vectors is a Lorentz transformation:

$$\mathbf{\Lambda} \in \mathcal{L} \qquad \Longrightarrow \qquad \big\langle x, y \big\rangle_M = \big\langle \mathbf{\Lambda} x, \mathbf{\Lambda} y \big\rangle_M \quad .$$

Remark 4.1.4 The tangent space is the same at every point in the manifold, so Lorentz transformations have a special relationship to special relativity. In curved space, each point in manifold has it's own attached group of Lorentz transformations. Since these are the transformations at a given p, the necessarily related two observers at the same p. To dislocate, we need the Poincare group, which is called the inhomogeneous Lorentz group.

ROTATIONS OF TETRADS:

Remark 4.1.5 Lorentz transformations are matrices, not tensors. Although it is normal to write expressions such as

$$x^{\mu'} = \Lambda^{\mu'}{}_{\mu} x^\mu \quad ,$$

for the Lorentz transformation of a 4-vector, there is no tensorial significance given to the placement of the upper and lower indices on $\Lambda^{\mu'}{}_{\mu}$. Rather, the first index μ' is the row of the $\mathbf{\Lambda}$ matrix, and the second is column. The summation of the repeated index μ reflects the contraction of the rows of Λ with the single column of x. Matrix multiplication of an $m \times n$ matrix from the left into an $r \times s$ matrix is only a defined operation if $m = s$, so μ is guaranteed to run over the same number of elements in the matrix and the vector.

In matrix notation, the tensor expression $x^\mu y_\mu = y_\mu x^\mu$ must be written with the dual vector first:

$$y_\mu x^\mu \quad \longrightarrow \quad \mathsf{YX} = \begin{bmatrix} y_0 & y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \quad ,$$

where Y is a 1×4 matrix and X is a 4×1 matrix, and the entries in each are the corresponding components of y_μ and x^μ. (As written, the matrix product

XY is a 4×4 matrix, not a scalar.) In spacetime, however, the entries of y_μ are generally not the entries of y^μ because the metric is required for raising and lowering the indices. Given the Minkowski metric in signature $\{+---\}$, we have

$$y^\mu = (y^0, y^1, y^2, y^3) \quad \Longrightarrow \quad y_\mu = (y^0, -y^1, -y^2, -y^3) \enspace ,$$

so plainly $y^\mu \neq y_\mu$. We cannot use convenient notation in which X, Y are one-column matrices corresponding to tensorial 4-vectors such that $y_\mu x^\mu \to \mathsf{Y}^T\mathsf{X}$. Rather, upper index vectors x^μ are written X, and lower index dual vectors x_μ are written $(\boldsymbol{\eta}\mathsf{X})^T$. (Later, when we introduce the convention for complex 4-vectors in $\mathbb{R}^{1,3}$, we will have the option to use the Euclidean metric in which the dual is just the transpose.) Recalling that the transpose of a product is the reversed product of transposes, matrix notation is such that

$$\big\langle \mathsf{X}, \mathsf{Y} \big\rangle_M = \big\langle \mathsf{X}, \boldsymbol{\eta}\mathsf{Y} \big\rangle = \big(\boldsymbol{\eta}\mathsf{Y}\big)^T \mathsf{X} = \mathsf{Y}^T \boldsymbol{\eta}^T \mathsf{X} \enspace .$$

Since the metric is always symmetric, we have

$$\big\langle \mathsf{X}, \mathsf{Y} \big\rangle_M = \mathsf{Y}^T \boldsymbol{\eta} \mathsf{X} \enspace .$$

Sometimes in the literature the symbol for the metric stripped of its indices is taken to mean the determinant of its matrix representation. The expression $\sqrt{-g}$ for a metric $g_{\mu\nu}$ often means the square root of the sign-inverted determinant, but we will not use that convention here. Instead, we will drop the indices and use the bold typeface to indicate that the metric is being treated as a matrix, and we will always use det to indicate a determinant.

Example 4.1.6 The Minkowski square in matrix notation. Recall that the Minkowski square is the metric inner product of a 4-vector with itself in flat spacetime (Definition XXXX). In tensor notation, we have

$$\begin{aligned} x^\mu = (ct, x, y, z) \quad \Longrightarrow \quad x_M^2 &= \big\langle x, x \big\rangle_g = \eta_{\mu\nu} x^\mu x^\nu \\ &= (\pm ct, \mp x, \mp y, \mp z) \cdot (ct, x, y, z) \\ &= \pm \left(c^2 t^2 - x^2 - y^2 - z^2 \right) \enspace . \end{aligned}$$

In matrix notation, we have

$$\begin{aligned} \mathsf{X} = \begin{bmatrix} ct & x & y & z \end{bmatrix}^T \quad \Longrightarrow \quad \mathsf{X}_M^2 &= \big\langle \mathsf{X}, \mathsf{X} \big\rangle_M = \mathsf{X}^T \boldsymbol{\eta} \mathsf{X} \\ &= \begin{bmatrix} ct & x & y & z \end{bmatrix} \begin{bmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \mp 1 & 0 & 0 \\ 0 & 0 & \mp 1 & 0 \\ 0 & 0 & 0 & \mp 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} ct & x & y & z \end{bmatrix} \begin{bmatrix} \pm ct \\ \mp x \\ \mp y \\ \mp z \end{bmatrix}$$
$$= \pm \left(c^2 t^2 - x^2 - y^2 - z^2 \right) \quad .$$

Observe that $x_M^2 = \mathsf{X}_M^2$.

Example 4.1.7 Lorentz transformations of dual vectors. The invariance of the Minkowski square under Lorentz transformations is the defining property of such transformations:

$$\mathsf{X}^T \boldsymbol{\eta} \mathsf{X} = \big(\mathsf{X}'\big)^T \boldsymbol{\eta} \mathsf{X}' = \big(\boldsymbol{\Lambda}\mathsf{X}\big)^T \boldsymbol{\eta} \big(\boldsymbol{\Lambda}\mathsf{X}\big) = \big(\mathsf{X}^T \boldsymbol{\Lambda}^T\big) \boldsymbol{\eta} \big(\boldsymbol{\Lambda}\mathsf{X}\big) \quad .$$

Thus, the transformations of vectors and dual vectors correspond to multiplication by $\boldsymbol{\Lambda}$ or $\boldsymbol{\Lambda}^T$ from the left and right respectively:

$$\mathsf{X} \to \mathsf{X}' = \boldsymbol{\Lambda}\mathsf{X} = \sum_k \boldsymbol{\Lambda}_{jk} \mathsf{X}_k \quad ,$$

and

$$\mathsf{X}^T \to \big(\mathsf{X}'\big)^T = \mathsf{X}^T \boldsymbol{\Lambda}^T = \sum_k \big(\mathsf{X}^T\big)_k \big(\boldsymbol{\Lambda}^T\big)_{jk} = \sum_k \big(\mathsf{X}^T\big)_k \big(\boldsymbol{\Lambda}\big)_{kj} \quad .$$

Theorem 4.1.8 *The Minkowski metric satisfies* $\boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda} = \boldsymbol{\eta}$.

Proof. Per Definition 4.1.3, the Lorentz invariance of X_M^2 requires

$$\mathsf{X}_M^2 = \big(\mathsf{X}'\big)_M^2 \quad \iff \quad \mathsf{X}^T \boldsymbol{\eta} \mathsf{X} = \big(\boldsymbol{\Lambda}\mathsf{X}\big)^T \boldsymbol{\eta} \big(\boldsymbol{\Lambda}\mathsf{X}\big) \quad .$$

The transpose of a product is the product of the transposes in reverse order, so we may write

$$\mathsf{X}^T \boldsymbol{\eta} \mathsf{X} = \big(\mathsf{X}^T \boldsymbol{\Lambda}^T\big) \boldsymbol{\eta} \big(\boldsymbol{\Lambda}\mathsf{X}\big) \quad .$$

Matrix multiplication operations may be carried out in any order, so introduce $\boldsymbol{\eta}' = \Lambda^T \boldsymbol{\eta} \widetilde{\Lambda}$ to write

$$\mathsf{X}^T \boldsymbol{\eta} \mathsf{X} = \mathsf{X}^T \big(\widetilde{\Lambda}^T \boldsymbol{\eta} \widetilde{\Lambda}\big) \mathsf{X} = \mathsf{X}^T \boldsymbol{\eta}' \mathsf{X} \quad .$$

This equation is only satisfied if $\boldsymbol{\eta}$ is equal to $\boldsymbol{\eta}'$. Since we have defined $\boldsymbol{\eta}' = \Lambda^T \boldsymbol{\eta} \widetilde{\Lambda}$, the theorem is proven.

Written out component-wise in matrix index notation, we have

$$\boldsymbol{\Lambda} \in \mathcal{L} \quad \implies \quad \big\langle \mathsf{X}, \boldsymbol{\eta}\mathsf{X} \big\rangle = \big\langle \boldsymbol{\Lambda}\mathsf{X}, \boldsymbol{\eta}\boldsymbol{\Lambda}\mathsf{X} \big\rangle \quad ,$$

expressed as

$$\langle \mathsf{X}, \boldsymbol{\eta}\mathsf{X} \rangle = \sum_{\mu,v=0}^{3} \mathsf{X}_\mu \boldsymbol{\eta}_{\mu\nu} \mathsf{X}_\nu \quad , \quad \text{and} \qquad \langle \boldsymbol{\Lambda}\mathsf{X}, \boldsymbol{\eta}\boldsymbol{\Lambda}\mathsf{X} \rangle = \sum_{\mu,v=0}^{3} \big(\boldsymbol{\Lambda}\mathsf{X}\big)_\mu \boldsymbol{\eta}_{\mu\nu} \big(\boldsymbol{\Lambda}\mathsf{X}\big)_\nu \quad .$$

If we insert

$$\big(\boldsymbol{\Lambda}\mathsf{X}\big)_\nu = \sum_{\lambda=0}^{3} \boldsymbol{\Lambda}_{\nu\lambda} \mathsf{X}_\lambda \quad ,$$

into the expression on the right, we obtain

$$\begin{aligned} \langle \boldsymbol{\Lambda}\mathsf{X}, \boldsymbol{\eta}\boldsymbol{\Lambda}\mathsf{X} \rangle &= \sum_{\mu,\nu=0}^{3} \left[\left(\sum_{\sigma=0}^{3} \boldsymbol{\Lambda}_{\mu\sigma} \mathsf{X}_\sigma \right)_\mu \boldsymbol{\eta}_{\mu\nu} \left(\sum_{\lambda=0}^{3} \boldsymbol{\Lambda}_{\nu\lambda} \mathsf{X}_\lambda \right)_\nu \right] \\ &= \sum_{\sigma,\lambda=0}^{3} \mathsf{X}_\sigma \left(\sum_{\mu,\nu=0}^{3} \boldsymbol{\Lambda}_{\mu\sigma} \boldsymbol{\eta}_{\mu\nu} \boldsymbol{\Lambda}_{\nu\rho} \right) \mathsf{X}_\rho \end{aligned}$$

where the parenthetical quantity must be equal to $\boldsymbol{\eta}_{\sigma\rho}$ since we have defined $\boldsymbol{\Lambda}$ not to the change the metric inner product

$$\langle \boldsymbol{\Lambda}\mathsf{X}, \boldsymbol{\eta}\boldsymbol{\Lambda}\mathsf{X} \rangle = \langle \mathsf{X}, \boldsymbol{\eta}\mathsf{X} \rangle = \sum_{\sigma,\lambda=0}^{3} \mathsf{X}_\sigma \boldsymbol{\eta}_{\sigma\lambda} \mathsf{X}_\lambda \quad .$$

Reversing the indices in matrix notation is the transpose operation, so we can use $\boldsymbol{\Lambda}_{\mu\sigma} = \boldsymbol{\Lambda}^T_{\sigma\mu}$ in the above parenthetical sum over μ, ν, and we arrive at the proof of the theorem:

$$\boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda} = \sum_{\mu,\nu=0}^{3} \boldsymbol{\Lambda}_{\sigma\mu} \boldsymbol{\eta}_{\mu\nu} \boldsymbol{\Lambda}_{\nu\rho} \quad \Longrightarrow \quad \boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda} = \boldsymbol{\eta} \quad .$$

Remark 4.1.9 Since this property of the metric is in "if and only if" correspondence with the invariance of the Minkowski length, the equation $\boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda} = \boldsymbol{\eta}$ can also be taken as the defining property of $\boldsymbol{\Lambda} \in \mathcal{L}$, supplementing Definition 4.1.3.

Corollary 4.1.10 *Every $\boldsymbol{\Lambda} \in \mathcal{L}$ is uniquely determined by six parameters.*

Proof. Theorem 4.1.8 grants $\boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda} = \boldsymbol{\eta}$, which is represents a system of 16 linear equations with matrix algebra: one equation for each $(\boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda})_{\mu\nu} = \boldsymbol{\eta}_{\mu\nu}$. However, there are only ten unique equations: four for the diagonal entries $(\boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda})_{\mu\mu}$ and six for the off diagonal entries. It is easily verified that the constraint equation associated with each off diagonal element is equal to it symmetric counterpart, i.e.: $(\boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda})_{\mu\nu} = (\boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda})_{\nu\mu}$. Among the 12 constraint equations associated with the 12 zeros in $\boldsymbol{\eta}_{\mu\nu}$, therefore there are only six unique equations. The ten unique equations determine ten of the sixteen parameters in $\mathbf{\Lambda}$ leaving six free parameters to specify a given Lorentz transformations. The corollary is proven.

Remark 4.1.11 Later we will show that $\mathbf{\Lambda}$'s six free parameters are the three components of a boost velocity $\boldsymbol{v}$ and a rotation angle about each of the three spatial axes: the Euler angles.

Theorem 4.1.12 *The product of two Lorentz transformations is another Lorentz transformation.*

Proof. Given $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2 \in \mathcal{L}$, we know from Theorem 4.1.8 that

$$\mathbf{\Lambda}_1^T \boldsymbol{\eta} \mathbf{\Lambda}_1 = \boldsymbol{\eta} \quad , \quad \text{and} \qquad \mathbf{\Lambda}_2^T \boldsymbol{\eta} \mathbf{\Lambda}_2 = \boldsymbol{\eta} \quad .$$

If $(\mathbf{\Lambda}_1\mathbf{\Lambda}_2) \in \mathcal{L}$, then $(\mathbf{\Lambda}_1\mathbf{\Lambda}_2)^T \boldsymbol{\eta} (\mathbf{\Lambda}_1\mathbf{\Lambda}_2) = \boldsymbol{\eta}$, which is easily verified:

$$\big(\mathbf{\Lambda}_1\mathbf{\Lambda}_2\big)^T \boldsymbol{\eta} \big(\mathbf{\Lambda}_1\mathbf{\Lambda}_2\big) = \mathbf{\Lambda}_2^T \big(\mathbf{\Lambda}_1^T \boldsymbol{\eta} \mathbf{\Lambda}_1\big) \mathbf{\Lambda}_2 = \mathbf{\Lambda}_2^T \big(\boldsymbol{\eta}\big) \mathbf{\Lambda}_2 = \boldsymbol{\eta} \quad .$$

While this suffices to prove the theorem, the stated requirement for $(\mathbf{\Lambda}_1\mathbf{\Lambda}_2) \in \mathcal{L}$ is that it should leave the Minkowski square invariant (Definition 4.1.3):

$$\big\langle \mathsf{X}, \mathsf{Y} \big\rangle_M = \big\langle \mathbf{\Lambda}_1\mathbf{\Lambda}_2\mathsf{X}, \mathbf{\Lambda}_1\mathbf{\Lambda}_2\mathsf{Y} \big\rangle_M \quad .$$

This is also easily verified:

$$\begin{aligned}
\big\langle \mathbf{\Lambda}_1\mathbf{\Lambda}_2\mathsf{X}, \mathbf{\Lambda}_1\mathbf{\Lambda}_2\mathsf{Y} \big\rangle_M &= \big\langle \mathbf{\Lambda}_1\mathbf{\Lambda}_2\mathsf{X}, \boldsymbol{\eta}\mathbf{\Lambda}_1\mathbf{\Lambda}_2\mathsf{Y} \big\rangle \\
&= \big(\boldsymbol{\eta}\mathbf{\Lambda}_1\mathbf{\Lambda}_2\mathsf{Y}\big)^T \mathbf{\Lambda}_1\mathbf{\Lambda}_2\mathsf{X} \\
&= \big(\mathbf{\Lambda}_2\mathsf{Y}\big)^T \big(\boldsymbol{\eta}\mathbf{\Lambda}_1\big)^T \mathbf{\Lambda}_1\mathbf{\Lambda}_2\mathsf{X} \\
&= \mathsf{Y}^T \mathbf{\Lambda}_2^T \big(\mathbf{\Lambda}_1^T \boldsymbol{\eta} \mathbf{\Lambda}_1\big) \mathbf{\Lambda}_2\mathsf{X} \\
&= \mathsf{Y}^T \big(\mathbf{\Lambda}_2^T \boldsymbol{\eta} \mathbf{\Lambda}_2\big) \mathsf{X} \\
&= \mathsf{Y}^T \boldsymbol{\eta} \mathsf{X} = \big\langle \mathsf{X}, \mathsf{Y} \big\rangle_M \quad .
\end{aligned}$$

The theorem is proven, and it is clear why it was sufficient to demonstrate $(\mathbf{\Lambda}_1\mathbf{\Lambda}_2)^T \boldsymbol{\eta} (\mathbf{\Lambda}_1\mathbf{\Lambda}_2) = \boldsymbol{\eta}$.

Theorem 4.1.13 *Every Lorentz transformation has an inverse* $\mathbf{\Lambda}^{-1} \in \mathcal{L}$.

Proof. Given $\mathbf{\Lambda}\boldsymbol{\eta}\mathbf{\Lambda} = g$, we may take the determinant of both sides to obtain

$$\det(\mathbf{\Lambda}\boldsymbol{\eta}\mathbf{\Lambda}) = \det(g) \quad .$$

The determinant of a product is the product of determinants, so

$$\det(\mathbf{\Lambda})\det(\boldsymbol{\eta})\det(\mathbf{\Lambda}) = \det(\boldsymbol{\eta}) \qquad \Longrightarrow \qquad \big(\det(\mathbf{\Lambda})\big)^2 = 1 \quad .$$

Clearly, $\det \mathbf{\Lambda} = \pm 1$, and all matrices with non-zero determinant have an inverse:

$$\forall \mathbf{\Lambda} \in \mathcal{L} \qquad \exists \mathbf{\Lambda}^{-1} \qquad \text{s.t.} \qquad \mathbf{\Lambda}^{-1}\mathbf{\Lambda} = \mathbf{\Lambda}\mathbf{\Lambda}^{-1} = \mathbb{1} \quad ,$$

where $\mathbb{1}$ is the identity matrix in the same dimension as $\mathbf{\Lambda}$. Now we must demonstrate that $\mathbf{\Lambda}^{-1}$ is a Lorentz transformation, meaning $\mathbf{\Lambda}^{-1} \in \mathcal{L}$. This is easily demonstrated by inserting the identity into the self-equality of the metric:

$$\begin{aligned} \boldsymbol{\eta} &= \mathbb{1}^T \boldsymbol{\eta} \mathbb{1} \\ &= \big(\mathbf{\Lambda}\mathbf{\Lambda}^{-1}\big)^T \boldsymbol{\eta} \mathbf{\Lambda}\mathbf{\Lambda}^{-1} \\ &= \big(\mathbf{\Lambda}^{-1}\big)^T \mathbf{\Lambda}^T \boldsymbol{\eta} \mathbf{\Lambda}\mathbf{\Lambda}^{-1} \\ &= \big(\mathbf{\Lambda}^{-1}\big)^T \boldsymbol{\eta} \mathbf{\Lambda}^{-1} \quad . \end{aligned}$$

Since $\boldsymbol{\eta} = \mathbf{\Lambda}^{-1}\boldsymbol{\eta}\mathbf{\Lambda}$, we have established that $\mathbf{\Lambda}^{-1} \in \mathcal{L}$. This proves the theorem.

Definition 4.1.14 Given a set S, $S_x \subset S$ is a **connected component** of S if every element of S_x can be obtained by the continuous parametric variation of any other element of S_x. A connected component is maximal so every element that can be obtained by smooth variation of any element of S_x is also in S_x. Formally, S_x is connected component of S if

$$\forall S_1, S_2 \in S_x \qquad \exists \lambda \in [0,1]\ ,\ S(\lambda) \in S_x \qquad \text{s.t.} \qquad S(0) = S_1\ ,\ S(1) = S_2 \quad .$$

If two components of a set are not connect, then they are **disconnected**, or disjoint.

Remark 4.1.15 When we work in the convention such that the vectors associated with our spacetime manifold are $V^\mu \in \mathbb{R}^4$, it is standard to restrict the components of our $\mathbf{\Lambda}$ matrices to be real-valued. If they weren't, we would run into problems with our endomorphism definition that Lorentz transformations map vectors in a space to other vectors in the space. Accordingly, the set of such real-valued matrices that preserve the Minkowski length is called $\mathcal{L}(\mathbb{R})$, but we might work in another representation such that the components of $\mathbf{\Lambda}$ are complex. Observing that the 2×2 matrix form of a spacetime event uses complex numbers as

$$\begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} \quad \longrightarrow \quad \begin{bmatrix} t+z & x-iy \\ x+iy & t-z \end{bmatrix} \quad ,$$

it is to be expected that the corresponding Lorentz transformations will acquire complex components. The set of such transformations is called $\mathcal{L}(\mathbb{C})$, and it has a different number of connected components than $\mathcal{L}(\mathbb{R})$. Therefore, we will treat the properties of three different representations in the following sections, and then in Section 5 we will organize what the material in the framework of group theory. First, however, we will work through different representations with specific examples rather than jumping straight into the universal properties that are easily cataloged and enumerated with group theory.

§4.2 The Real 4-Vector Representation

Although we have already shown a lot of the properties of Lorentz transformations in the real representation when we developed special relativity in Section 3.1, we will repeat some examples here to set the stage for easy comparisons with other realizations: spinors in Section 4.4 and $V^\mu \in \mathbb{R}^{1,3}$ in Section 4.5.

Definition 4.2.1 $\mathcal{L}(\mathbb{R})$ is the set of **real, homogeneous Lorentz transformations**. The elements of this set are 4×4 matrices such that

$$\mathbf{\Lambda}_{\mu\nu} \in \mathbb{R} \quad , \qquad \text{and} \qquad \mathbf{\Lambda}^T \boldsymbol{\eta} \mathbf{\Lambda} = \boldsymbol{\eta} \quad ,$$

where $\boldsymbol{\eta}$ is the Minkowski metric $\text{diag}(\pm1, \mp1, \mp1, \mp1)$.

Theorem 4.2.2 *Every $\mathbf{\Lambda} \in \mathcal{L}(\mathbb{R})$ is such that $|\mathbf{\Lambda}_{00}| \geq 1$.*

Proof. We will will consider the 00 component of

$$\mathbf{\Lambda}^T \boldsymbol{\eta} \mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_{00} & \mathbf{\Lambda}_{10} & \mathbf{\Lambda}_{20} & \mathbf{\Lambda}_{30} \\ \mathbf{\Lambda}_{01} & \mathbf{\Lambda}_{11} & \mathbf{\Lambda}_{21} & \mathbf{\Lambda}_{31} \\ \mathbf{\Lambda}_{02} & \mathbf{\Lambda}_{12} & \mathbf{\Lambda}_{22} & \mathbf{\Lambda}_{32} \\ \mathbf{\Lambda}_{03} & \mathbf{\Lambda}_{13} & \mathbf{\Lambda}_{23} & \mathbf{\Lambda}_{33} \end{bmatrix} \begin{bmatrix} \pm1 & 0 & 0 & 0 \\ 0 & \mp1 & 0 & 0 \\ 0 & 0 & \mp1 & 0 \\ 0 & 0 & 0 & \mp1 \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_{00} & \mathbf{\Lambda}_{01} & \mathbf{\Lambda}_{02} & \mathbf{\Lambda}_{03} \\ \mathbf{\Lambda}_{10} & \mathbf{\Lambda}_{11} & \mathbf{\Lambda}_{12} & \mathbf{\Lambda}_{13} \\ \mathbf{\Lambda}_{20} & \mathbf{\Lambda}_{21} & \mathbf{\Lambda}_{22} & \mathbf{\Lambda}_{23} \\ \mathbf{\Lambda}_{30} & \mathbf{\Lambda}_{31} & \mathbf{\Lambda}_{32} & \mathbf{\Lambda}_{33} \end{bmatrix} \quad ,$$

where the reversed indices in the matrix to the left indicates the transpose of the one on the right. The 00 component is

$$\left(\mathbf{\Lambda}^T \boldsymbol{\eta} \mathbf{\Lambda}\right)_{00} = \pm\left(\mathbf{\Lambda}_{00}^2 - \mathbf{\Lambda}_{10}^2 - \mathbf{\Lambda}_{20}^2 - \mathbf{\Lambda}_{30}^2\right) \quad .$$

This is equal to $\boldsymbol{\eta}_{00}$, so

$$1 = \mathbf{\Lambda}_{00}^2 - \mathbf{\Lambda}_{10}^2 - \mathbf{\Lambda}_{20}^2 - \mathbf{\Lambda}_{30}^2 \quad .$$

This may be rearranged as

$$|\mathbf{\Lambda}_{00}| = \left(1 + \sum_k \mathbf{\Lambda}_{k0}^2\right)^{1/2} \quad .$$

Since $\mathbf{\Lambda}_{k0} \in \mathbb{R}$, the squares are non-negative, and we obtain

$$|\mathbf{\Lambda}_{00}| \geq 0 \quad .$$

The theorem is proven. ⌀

Theorem 4.2.3 $\mathcal{L}(\mathbb{R})$ *has four disconnected components determined by the signs of* $\mathbf{\Lambda}_{00}$ *and* $\det \mathbf{\Lambda}$.

Proof. By Definition 4.1.14, two components of $\mathcal{L}(\mathbb{R})$ are disconnected if they are not connected. If a transform with $\mathbf{\Lambda}_{00} \geq 1$ is connected to one with with $\mathbf{\Lambda}_{00} \leq 0$, then there exists a continuous parameterization $\mathbf{\Lambda}(s)$ with $s \in [0,1]$ such that

$$\mathbf{\Lambda}_{00}(0) \geq 1 \quad , \quad \text{and} \quad \mathbf{\Lambda}_{00}(1) \leq 0 \quad .$$

By the intermediate value theorem, the continuity of the parameterization in s guarantees some $s_0 \in [0,1]$ such that

$$\mathbf{\Lambda}_{00}(s_0) = 0 \quad .$$

This contradicts the result of Theorem 4.2.2 that

$$\mathbf{\Lambda} \in \mathcal{L}(\mathbb{R}) \quad \Longrightarrow \quad |\mathbf{\Lambda}_{00}| \geq 0 \quad .$$

We have shown in Theorem 4.1.13 that every $\mathbf{\Lambda} \in \mathcal{L}$ has $\det(\mathbf{\Lambda}) = \pm 1$, and it follows that $\mathbf{\Lambda} \in \mathcal{L}(\mathbb{R})$ inherits this property. By the intermediate value theorem, if two matrices with oppositely signed determinants were connected, then there would exist some $\det(\mathbf{\Lambda}(s_0)) = 0$, but that contradicts the stated restriction. Therefore, the theorem is proven by contradiction. $\mathcal{L}(\mathbb{R})$ has four disjoint components. ⌀

Definition 4.2.4 Connected part of $\mathcal{L}(\mathbb{R})$ with $\det(\mathbf{\Lambda}) = 1$ is the set of **proper** Lorentz transformations, and the part with $\det(\mathbf{\Lambda}) = -1$ is the set of **improper** Lorentz transformations. These are denoted $\mathcal{L}_{\pm}$ respectively. The transformations in the connected component with $\mathbf{\Lambda}_{00} \geq 1$ are called **orthochronous**, and those with with $\mathbf{\Lambda}_{00} \leq 1$ are called **non-orthochronous**. The connected components of $\mathcal{L}(\mathbb{R})$ are labeled as follows:

- Proper, orthochronous (*restricted*) transformations: $\mathcal{L}_+^\uparrow \begin{cases} \mathbf{\Lambda}_{00} \geq 1 \\ \det(\mathbf{\Lambda}) = 1 \end{cases}$
- Improper, orthochronous transformations: $\mathcal{L}_-^\uparrow \begin{cases} \mathbf{\Lambda}_{00} \geq 1 \\ \det(\mathbf{\Lambda}) = -1 \end{cases}$
- Proper, non-orthochronous transformations: $\mathcal{L}_+^\downarrow \begin{cases} \mathbf{\Lambda}_{00} \leq 1 \\ \det(\mathbf{\Lambda}) = 1 \end{cases}$

• Improper, non-orthochronous transformations: $\mathcal{L}_{-}^{\downarrow}\begin{cases}\mathbf{\Lambda}_{00}\leq 1\\ \det(\mathbf{\Lambda})=-1\end{cases}$

Example 4.2.5 Restrictions on $\mathbf{\Lambda_{00}}$. Given the boosts and rotations that we have already considered, it is easy to understand the restriction $|\mathbf{\Lambda}_{00}|\geq 1$. Consider

$$\mathbf{\Lambda}_B(\hat{e}_1,\gamma)=\begin{bmatrix}\gamma & -\beta\gamma & 0 & 0\\ -\beta\gamma & \gamma & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\quad,\quad\text{and}\qquad \mathbf{\Lambda}_R(\hat{e}_1,\theta)=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & \mathrm{C} & -\mathrm{S}\\ 0 & 0 & \mathrm{S} & \mathrm{C}\end{bmatrix}\quad.$$

Rotations have $\mathbf{\Lambda}_{00}=1$ and from

$$\gamma=\frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}}\quad,$$

we see that boosts have $\mathbf{\Lambda}_{00}\geq 1$. However, the definition of $\mathcal{L}(\mathbb{R})$ allows negatively signed $\mathbf{\Lambda}_{00}$ which exceed the scope of the transformations that we have examined in Section 3.1. The issue is that we are mostly concerned with *proper, orthochronous Lorentz transformations* in special relativity.

Definition 4.2.6 Proper, orthochronous Lorentz transformations are called **restricted** Lorentz transformations. This convention is chosen mainly for brevity.

Example 4.2.7 Orthochonicity refers to the preservation of the arrow of time. Consider a (proper) orthochronous Lorentz boost acting on $x^\mu\in\mathbb{R}^4$

$$\mathbf{\Lambda}_B(\hat{e}_1,\gamma)\mathsf{X}=\mathsf{X}'=\begin{bmatrix}ct'\\ x'\\ y'\\ z'\end{bmatrix}=\begin{bmatrix}\gamma & -\beta\gamma & 0 & 0\\ -\beta\gamma & \gamma & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}ct\\ x\\ y\\ z\end{bmatrix}=\begin{bmatrix}\gamma\,(ct-\beta x)\\ \gamma\big(x-vt\big)\\ y\\ z\end{bmatrix}\quad.$$

As t increases in the S frame, t' also increases in the S' frame:

$$t_2>t_1\qquad\Longrightarrow\qquad \gamma\big(ct_2-\beta x\big)>\gamma\big(ct_1-\beta x\big)\quad.$$

We may contrast this with a (proper or improper) non-orthochronous transformation

$$\mathbf{\Lambda}'_B(\hat{e}_1,\gamma)\mathsf{X}=\mathsf{X}'=\begin{bmatrix}ct'\\ x'\\ y'\\ z'\end{bmatrix}=\begin{bmatrix}-\gamma & -\beta\gamma & 0 & 0\\ \mp\beta\gamma & \pm\gamma & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}ct\\ x\\ y\\ z\end{bmatrix}=\begin{bmatrix}-\gamma\,(ct-\beta x)\\ \pm\gamma\big(x-vt\big)\\ y\\ z\end{bmatrix}\quad.$$

By changing the sign of $\mathbf{\Lambda}_{00}$, increasing time in the S frame becomes decreasing time in the Lorentz transformed S' frame:

$$t_2 > t_1 \quad\implies\quad -\gamma\big(ct_2 - \beta x\big) < -\gamma\big(ct_1 - \beta x\big) \quad .$$

$\mathbf{\Lambda}'$ obviously has $\mathbf{\Lambda}'_{00} \leq 1$, and it is easily verified that $\det(\mathbf{\Lambda}') = \pm 1$, so it is a proper or improper, orthochronous Lorentz transformation.

Example 4.2.8 Propriety refers to the preservation of the relative orientation of space and time. Following the previous example, the proper or improper (orthochronous) version of a boost in the $\hat{e}_1$ direction acts as

$$\mathbf{\Lambda}''_B(\hat{e}_1, \gamma)\mathsf{X} = \mathsf{X}' = \begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & \beta\gamma & 0 & 0 \\ \mp\beta\gamma & \pm\gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \gamma\,(ct - \beta x) \\ \pm\gamma\big(x - vt\big) \\ y \\ z \end{bmatrix} \quad .$$

Increasing t in the S frame is now associated with increasing or decreasing x distance in the S' frame, depending on the sign of $\det(\mathbf{\Lambda}'')$.

Definition 4.2.9 For $V^\mu \in \mathbb{R}^4$, the **parity**, **time inversion**, and **spacetime inversion** matrices are defined as follows:

$$\begin{aligned} \boldsymbol{P} &= \text{diag}(1, -1, -1, -1) \\ \boldsymbol{T} &= \text{diag}(-1, 1, 1, 1) \\ \boldsymbol{Y} &= \text{diag}(-1, -1, -1, -1) = \boldsymbol{PT} \quad . \end{aligned}$$

These relate the connected components of $\mathcal{L}(\mathbb{R})$ as

$$\mathcal{L}_+^\downarrow = \boldsymbol{Y}\mathcal{L}_+^\uparrow \qquad\qquad \mathcal{L}_-^\uparrow = \boldsymbol{P}\mathcal{L}_+^\uparrow \qquad\qquad \mathcal{L}_-^\downarrow = \boldsymbol{T}\mathcal{L}_+^\uparrow \quad .$$

Theorem 4.2.10 *The four connected components of $\mathcal{L}(\mathbb{R})$ are connected to the identity, parity, time inversion, and spacetime inversion.*

Proof. Per Definition 4.1.14, connected components are maximal so every element of a connected component is connected to every other element. We have already proven that there are dour disconnected components of $\mathcal{L}(\mathbb{R})$ (Theorem 4.2.3), so it will suffice to show that one matrix from each connected part is related to the identity, parity, or time or spacetime inversion. A pure boost $\mathbf{\Lambda}_B(\hat{e}_1, \gamma)$ is a restricted transformation in $\mathcal{L}_+^\uparrow$, and we may take the parameter $\beta = v/c \in [0, 1]$ to write

$$\mathbf{\Lambda}_+^\uparrow(\hat{e}_1, \beta) = \begin{bmatrix} \gamma(\beta) & -\beta\gamma(\beta) & 0 & 0 \\ -\beta\gamma(\beta) & \gamma(\beta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad .$$

Clearly, $\mathbf{\Lambda}_+^\uparrow(\hat{e}_1, 0) = \mathbb{1}_4$. This boost in the x-direction can be smoothly varied by letting $\hat{e}_1 \to \hat{n}$ where $\hat{n}$ is a unit vector in an arbitrary direction, so every restricted Lorentz transformation $\mathbf{\Lambda} \in \mathcal{L}_+^\uparrow$ is connected to the identity matrix. Similar demonstrations prove the connectedness of $\mathcal{L}_-^\uparrow$ to $\boldsymbol{P}$, $\mathcal{L}_+^\downarrow$ to $\boldsymbol{Y}$, and $\mathcal{L}_-^\downarrow$ to $\boldsymbol{T}$. The theorem is proven.

Remark 4.2.11 In special relativity, we are work almost exclusively with the proper Lorentz transformations. In quantum field theory, parity and other inversion operators become important, and likewise they will be important for non-quantum physics in the MCM unit cell. Chirological time is reversed between $\Sigma^\pm$, and the handedness of the coordinates is also expected to be reversed in $\Sigma^\pm$. This already seems to require parity and time inversion, and the context in which $\mathcal{A}$ and Ω are embedded chronological spacetime in $\Sigma^\pm$ suggests a context for spacetime inversion when transforming vectors in one space into vectors in another as is required for

$$\hat{M}^3 : \mathcal{H} \to \Omega \to \mathcal{A} \to \mathcal{H} \quad .$$

We will also need to examine the case in which $\hat{M}^3$ isn't strictly an endomorphism when it is written as

$$\hat{M}^3 : \mathcal{H}_1 \to \Omega \to \mathcal{A} \to \mathcal{H}_2 \quad .$$

Definition 4.2.12 $\mathbf{\Lambda}_R \in \mathcal{L}_+^\uparrow$ is a **proper rotation** if $(\mathbf{\Lambda}_R)_{00} = 1$, meaning it leaves time unchanged. Such transformations take the form

$$\mathbf{\Lambda}_R = \begin{bmatrix} 1 & 0 \\ 0 & \mathcal{R} \end{bmatrix} \quad .$$

where $\mathcal{R}$ is a 3×3 orthogonal matrix such that $\mathcal{R}^T\mathcal{R} = \mathbb{1}$ and $\det(\mathcal{R}) = 1$. (In other words, $\mathcal{R} \in \mathrm{SO}(3)$.) For example,

$$\mathcal{R}(\hat{e}_3, \theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \Longrightarrow \qquad \mathbf{\Lambda}_R(\hat{e}_3, \theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathrm{C} & -\mathrm{S} & 0 \\ 0 & \mathrm{S} & \mathrm{C} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad .$$

Rotations can rotate coordinates, or space. Often times, a convention is such that the sign of θ indicates one or the other, according to Jaffe. However, we should be able to rotate one or the other in either direction, so that convention is very strange.

Definition 4.2.13 $\mathbf{\Lambda}_B \in \mathcal{L}_+^\uparrow$ is a **pure boost** in the $\hat{n}$-direction if it leaves unchanged any vectors in 3-space which are in the plane orthogonal to $\hat{n}$. A pure boost in an arbitrary direction is written

$$\mathbf{\Lambda}_B(\boldsymbol{\beta}) = \begin{bmatrix} \gamma & -\boldsymbol{\beta}^T\gamma \\ -\boldsymbol{\beta}\gamma & \mathbb{1}_3 + \big(\gamma - 1\big)\dfrac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2} \end{bmatrix} \quad ,$$

where $\boldsymbol{\beta} = \boldsymbol{v}/c$, $\boldsymbol{v}$ is the boost velocity, $\boldsymbol{\beta}\boldsymbol{\beta}^T$ is a 3×3 matrix, and $\beta^2 = \boldsymbol{\beta}^T\boldsymbol{\beta}$. When $\boldsymbol{v} = v\boldsymbol{n}$, it is often useful to introduce the rapidity ρ such that

$$\tanh\rho = \frac{v}{c} = \beta \qquad \Longrightarrow \qquad \sinh\rho = \beta\gamma \quad , \quad \cosh\rho = \gamma \quad .$$

For example:

$$\mathbf{\Lambda}_B(\hat{e}_1, \chi) = \begin{pmatrix} \cosh\chi & \sinh\chi & 0 & 0 \\ \sinh\chi & \cosh\chi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad .$$

Pure boosts are always represented by symmetric matrices. By choosing the direction of the boost with $\pm\hat{n}$, we can restrict the rapidity as $\chi \geq 0$.

Remark 4.2.14 The following theorem is cumbersome, and it is proven much more easily in the spinor representation of $\mathcal{L}$. To highlight the fact that different representations are more useful or less useful for certain things, we will prove the theorem in the present representation and again in the following section with the spinor representation.

Theorem 4.2.15 *Any restricted Lorentz transformation $\mathbf{\Lambda} \in \mathcal{L}_+^\uparrow$ can be factored into a proper rotation $\mathbf{\Lambda}_R$ followed by a pure boost $\mathbf{\Lambda}_B$:*

$$\mathbf{\Lambda} \in \mathcal{L}_+^\uparrow \qquad \Longleftrightarrow \qquad \mathbf{\Lambda} = \mathbf{\Lambda}_B\mathbf{\Lambda}_R \quad .$$

Proof. Proof of this theorem requires

$$\mathbf{\Lambda} = \begin{bmatrix} \gamma & -\boldsymbol{\beta}^T\gamma \\ -\boldsymbol{\beta}\gamma & \mathbb{1}_3 + \big(\gamma - 1\big)\dfrac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \mathcal{R} \end{bmatrix} = \begin{bmatrix} \gamma & -\boldsymbol{\beta}^T\gamma\mathcal{R} \\ -\boldsymbol{\beta}\gamma & \mathcal{R} + \big(\gamma - 1\big)\dfrac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2}\mathcal{R} \end{bmatrix} , \tag{4.1}$$

where we have taken the general forms of $\mathbf{\Lambda}_B$ and $\mathbf{\Lambda}_R$ from Definitions 4.2.12 and 4.2.13. Equating the first column of $\mathbf{\Lambda}$ to the first column of $\mathbf{\Lambda}_B\mathbf{\Lambda}_R$ on the right gives

$$\begin{bmatrix} \mathbf{\Lambda}_{00} & \mathbf{\Lambda}_{10} & \mathbf{\Lambda}_{20} & \mathbf{\Lambda}_{30} \end{bmatrix}^T = \begin{bmatrix} \gamma & -\beta_x\gamma & -\beta_y\gamma & -\beta_z\gamma \end{bmatrix}^T \quad ,$$

or, more specifically,

$$\mathbf{\Lambda}_{00} = \gamma \quad , \quad \text{and} \qquad \mathbf{\Lambda}_{i0} = -\boldsymbol{\beta}\gamma \quad . \tag{4.2}$$

It remains to show that $\mathbf{\Lambda}$'s other 12 matrix elements are consistent with $\mathcal{R}$ being a 3D rotation matrix.

Every $\mathbf{\Lambda} \in \mathcal{L}$ has an inverse (Theorem 4.1.13), so we may left multiply the statement of the theorem by $\mathbf{\Lambda}_B^{-1}$ to isolate the rotation part as

$$\mathbf{\Lambda}_B^{-1}\mathbf{\Lambda} = \mathbf{\Lambda}_B^{-1}\mathbf{\Lambda}_B\mathbf{\Lambda}_R = \mathbf{\Lambda}_R \quad . \tag{4.3}$$

Using $\mathbf{\Lambda}_B^{-1}(\boldsymbol{\beta}) = \mathbf{\Lambda}_B(-\boldsymbol{\beta})$, we may expand the identity as

$$\mathbb{1}_4 = \mathbf{\Lambda}_B^{-1}\mathbf{\Lambda}_B = \begin{bmatrix} \gamma & \boldsymbol{\beta}^T\gamma \\ \boldsymbol{\beta}\gamma & \mathbb{1}_3 + \big(\gamma - 1\big)\dfrac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2} \end{bmatrix} \begin{bmatrix} \gamma & -\boldsymbol{\beta}^T\gamma \\ -\boldsymbol{\beta}\gamma & \mathbb{1}_3 + \big(\gamma - 1\big)\dfrac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2} \end{bmatrix} \quad .$$

The inverse is laboriously verified as

$$\begin{aligned}
\big(\mathbf{\Lambda}_B^{-1}\mathbf{\Lambda}_B\big)_{00} &= \gamma^2 - \boldsymbol{\beta}^T\boldsymbol{\beta}\gamma^2 = \gamma^2\big(1 - \beta^2\big) = 1 \\
\big(\mathbf{\Lambda}_B^{-1}\mathbf{\Lambda}_B\big)_{0i} &= -\boldsymbol{\beta}^T\gamma^2 + \boldsymbol{\beta}^T\gamma\left[\mathbb{1}_3 + \big(\gamma - 1\big)\frac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2}\right] \\
&= -\boldsymbol{\beta}^T\gamma^2 + \boldsymbol{\beta}^T\gamma + \big(\gamma - 1\big)\boldsymbol{\beta}^T\gamma \\
&= \boldsymbol{\beta}^T\big(-\gamma^2 + \gamma + \gamma^2 - \gamma\big) = 0 \\
\big(\mathbf{\Lambda}_B^{-1}\mathbf{\Lambda}_B\big)_{j0} &= \boldsymbol{\beta}\gamma^2 - \left[\mathbb{1}_3 + \big(\gamma - 1\big)\frac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2}\right]\boldsymbol{\beta}\gamma \\
&= \boldsymbol{\beta}\gamma^2 - \boldsymbol{\beta}\gamma - \big(\gamma - 1\big)\boldsymbol{\beta}\gamma \\
&= \boldsymbol{\beta}\big(\gamma^2 - \gamma - \gamma^2 + \gamma\big) = 0 \\
\big(\mathbf{\Lambda}_B^{-1}\mathbf{\Lambda}_B\big)_{ij} &= -\boldsymbol{\beta}\boldsymbol{\beta}^T\gamma^2 + \left[\mathbb{1}_3 + \big(\gamma - 1\big)\frac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2}\right]^2 \\
&= -\boldsymbol{\beta}\boldsymbol{\beta}^T\gamma^2 + \left[\mathbb{1}_3 + 2\big(\gamma - 1\big)\frac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2} + \big(\gamma - 1\big)^2\frac{\boldsymbol{\beta}\boldsymbol{\beta}^T\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^4}\right] \\
&= \mathbb{1}_3 + \frac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2}\left[-\beta^2\gamma^2 + 2\gamma - 2 + \gamma^2 - 2\gamma + 1\right] \\
&= \mathbb{1}_3 + \frac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2}\left[\gamma^2\big(1 - \beta^2\big) - 1\right] = \mathbb{1}_3 \quad ,
\end{aligned}$$

where $\mathbb{1}_3 = \delta_{ij}$ balances out the indices in the final equality. Now that we have verified the matrix form of $\mathbf{\Lambda}_B^{-1}$, we continue from Equation (4.3) as

$$\mathbf{\Lambda}_B^{-1}\mathbf{\Lambda} = \mathbf{\Lambda}_R \quad \longrightarrow \quad \begin{bmatrix} \gamma & \boldsymbol{\beta}^T\gamma \\ \boldsymbol{\beta}\gamma & \mathbb{1}_3 + \big(\gamma - 1\big)\dfrac{\boldsymbol{\beta}\boldsymbol{\beta}^T}{\beta^2} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_{00} & \mathbf{\Lambda}_{0j} \\ \mathbf{\Lambda}_{i0} & \mathbf{\Lambda}_{ij} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \mathcal{R} \end{bmatrix} \quad ,$$

and it remains to prove that the ij components of the matrix product are a rotation matrix $\mathcal{R}$. After computing the matrix product $\mathbf{\Lambda}_B^{-1}\mathbf{\Lambda}$, we find

$$\mathcal{R} = \boldsymbol{\beta}\gamma\mathbf{\Lambda}_{0j} + \mathbf{\Lambda}_{ij} + \frac{\big(\gamma-1\big)}{\beta^2}\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{\Lambda}_{ij} \quad .$$

The convenient hodgepodge of vector and index notation should be rendered now in consistent index notation. Equation (4.2) gives $\boldsymbol{\beta} = -\mathbf{\Lambda}_{i0}/\gamma$ from which it follows that

$$\boldsymbol{\beta}^T = -\frac{\big(\mathbf{\Lambda}_{i0}\big)^T}{\gamma} \quad , \quad \text{and} \qquad \beta^2 = \boldsymbol{\beta}^T\boldsymbol{\beta} = \frac{\big(\mathbf{\Lambda}_{i0}\big)^T\mathbf{\Lambda}_{i0}}{\gamma^2} \quad .$$

Note that the transpose of the vector with components $\mathbf{\Lambda}_{i0}$ does not necessarily satisfy $(\mathbf{\Lambda}_{i0})^T = \mathbf{\Lambda}_{0i}$. This is easily seen in Equation (4.1). Now we may simplify $\mathcal{R}$ as

$$\mathcal{R}_{ij} = \mathbf{\Lambda}_{ij} - \mathbf{\Lambda}_{i0}\mathbf{\Lambda}_{0j} + \frac{\big(\gamma-1\big)}{\beta^2\gamma^2}\mathbf{\Lambda}_{i0}\big(\mathbf{\Lambda}_{k0}\big)^T\mathbf{\Lambda}_{kj} \quad ,$$

where we have inserted the k index to indicate that the 3×3 matrix $\boldsymbol{\beta}\boldsymbol{\beta}^T = \mathbf{\Lambda}_{i0}(\mathbf{\Lambda}_{k0})^T/\gamma^2$ is matrix multiplied into the spatial part of $\mathbf{\Lambda}$, which is another 3×3 matrix $\mathbf{\Lambda}_{kj}$ (as inferred from $i,j,k\in\{1,2,3\}$.) Now we use the identity $\beta^2\gamma^2 = (\gamma+1)(\gamma-1)$ along with $\gamma = \mathbf{\Lambda}_{00}$ (Equation (4.2)) to write

$$\begin{aligned}
\mathcal{R}_{ij} &= \mathbf{\Lambda}_{ij} - \mathbf{\Lambda}_{i0}\mathbf{\Lambda}_{0j} + \frac{\mathbf{\Lambda}_{i0}\big(\mathbf{\Lambda}_{k0}\big)^T\mathbf{\Lambda}_{kj}}{\mathbf{\Lambda}_{00}+1} \\
&= \mathbf{\Lambda}_{ij} - \frac{\mathbf{\Lambda}_{i0}\mathbf{\Lambda}_{0j}\big(\mathbf{\Lambda}_{00}+1\big) - \mathbf{\Lambda}_{i0}\big(\mathbf{\Lambda}_{k0}\big)^T\mathbf{\Lambda}_{kj}}{\mathbf{\Lambda}_{00}+1} \\
&= \mathbf{\Lambda}_{ij} - \mathbf{\Lambda}_{i0}\left\{\frac{\mathbf{\Lambda}_{0j} + \big[\mathbf{\Lambda}_{0j}\mathbf{\Lambda}_{00} - \big(\mathbf{\Lambda}_{k0}\big)^T\mathbf{\Lambda}_{kj}\big]}{\mathbf{\Lambda}_{00}+1}\right\} \quad . \qquad (4.4)
\end{aligned}$$

The term in square brackets is expanded as

$$\big[\mathbf{\Lambda}_{0j}\mathbf{\Lambda}_{00} - \big(\mathbf{\Lambda}_{k0}\big)^T\mathbf{\Lambda}_{kj}\big] = \mathbf{\Lambda}_{0j}\mathbf{\Lambda}_{00} - \big(\mathbf{\Lambda}_{10}\mathbf{\Lambda}_{1j} + \mathbf{\Lambda}_{20}\mathbf{\Lambda}_{2j} + \mathbf{\Lambda}_{30}\mathbf{\Lambda}_{3j}\big) \quad .$$

This vanishes as the $\mu\nu = 0k$ component of the definition of the Lorentz transformation:

$$\begin{aligned}
0 = \boldsymbol{\eta}_{0k} &= \big(\mathbf{\Lambda}^T\boldsymbol{\eta}\mathbf{\Lambda}\big)_{0k} \\
&= \left(\begin{bmatrix}\mathbf{\Lambda}_{00} & \mathbf{\Lambda}_{0j} \\ \mathbf{\Lambda}_{i0} & \mathbf{\Lambda}_{ij}\end{bmatrix}^T \begin{bmatrix}-1 & 0 \\ 0 & \mathbb{1}_3\end{bmatrix}\begin{bmatrix}\mathbf{\Lambda}_{00} & \mathbf{\Lambda}_{0k} \\ \mathbf{\Lambda}_{j0} & \mathbf{\Lambda}_{jk}\end{bmatrix}\right)_{0k} \\
&= \left(\begin{bmatrix}\mathbf{\Lambda}_{00} & \mathbf{\Lambda}_{j0} \\ \mathbf{\Lambda}_{0i} & \mathbf{\Lambda}_{ji}\end{bmatrix}\begin{bmatrix}-\mathbf{\Lambda}_{00} & -\mathbf{\Lambda}_{0k} \\ \mathbf{\Lambda}_{j0} & \mathbf{\Lambda}_{jk}\end{bmatrix}\right)_{0k}
\end{aligned}$$

$$= -\mathbf{\Lambda}_{00}\mathbf{\Lambda}_{0k} + \mathbf{\Lambda}_{j0}\mathbf{\Lambda}_{jk} \quad .$$

Thus, Equation (4.4) becomes

$$\mathcal{R}_{ij} = \mathbf{\Lambda}_{ij} - \frac{\mathbf{\Lambda}_{i0}\mathbf{\Lambda}_{0j}}{\mathbf{\Lambda}_{00}+1} \quad , \tag{4.5}$$

and it only remains to show that these are the components of a proper rotation matrix $\mathcal{R} \in \mathrm{SO}(3)$, namely

$$\det(\mathcal{R}) = 1 \quad , \quad \text{and} \qquad \mathcal{R}^T\mathcal{R} = \mathbb{1}_3 \quad .$$

The orthogonality condition $\mathcal{R}^T\mathcal{R} = \mathbb{1}_n$ imposes $n(n+1)/2$ constraints on an $n \times n$ orthogonal matrix, and, presently, $n = 3$, so we are left with $9 - 6 = 3$ free parameters in $\mathbf{\Lambda}_R$ such that $\mathcal{R} \in \mathrm{O}(3)$. The orthogonality condition automatically satisfies $\det(\mathcal{R}) = \pm 1$, so we do not need an additional constraint to ensure that $\mathbf{\Lambda}_R$ is a *proper* rotation $\mathcal{R} \in \mathrm{SO}(3)$ with $\det(\mathcal{R}) = 1$. That is only a matter of choosing an overall sign. It was demonstrated in Corollary 4.1.10 that every $\mathbf{\Lambda} \in \mathcal{L}$ has six free parameters, and we only fixed three as the components of $\boldsymbol{\beta}$ in Equation (4.1). Therefore, a restricted (proper, orthochronous) Lorentz transformation $\mathbf{\Lambda} \in \mathcal{L}_+^\uparrow$ has sufficient freedom in its remaining three parameters to ensure that $\mathcal{R}$ is orthogonal. For any $\mathbf{\Lambda}_B \in \mathcal{L}_+^\uparrow$, there exists a $\mathbf{\Lambda}_R \in \mathcal{L}_+^\uparrow$ such that $\mathbf{\Lambda} = \mathbf{\Lambda}_B\mathbf{\Lambda}_R$. This proves the theorem in principle. An *exceedingly laborious* calculation shows that Equation (4.5) is, in fact, an orthogonal matrix with unit determinant.

Remark 4.2.16 In the remainder of this section, we will demonstrate the Lorentz invariance or non-invariance of a few physical quantities under transformations by $\mathbf{\Lambda} \in \mathcal{L}_+^\uparrow(\mathbb{R})$.

Theorem 4.2.17 *A separation 4-vector $x^\mu \in \mathbb{R}^4$ is a Lorentz vector with invariant $x_M^2 = (\Delta s)^2$.*

Proof. We will demonstrate the invariance of x^μ under a specific case of $\mathbf{\Lambda} = \mathbf{\Lambda}_B\mathbf{\Lambda}_R$, and then we will assume that it is invariant under all $\mathbf{\Lambda} \in \mathcal{L}$. When x^μ is a strictly $x^\mu \in \mathbb{R}^4$, a boost and rotation may be expressed as

$$\mathbf{\Lambda}_B(\beta,\hat{e}_3)\mathbf{\Lambda}_R(\theta,\hat{e}_2) = \begin{bmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\theta & 0 & \cos\theta \end{bmatrix} \quad .$$

so the matrix representation of the Lorentz transformed position is

$$\mathsf{X}' = \mathbf{\Lambda}\mathsf{X} = \begin{bmatrix} \gamma & \mathrm{S}\beta\gamma & 0 & -\mathrm{C}\beta\gamma \\ 0 & \mathrm{C} & 0 & \mathrm{S} \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & -\mathrm{S}\gamma & 0 & \mathrm{C}\gamma \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \gamma\big(ct + \mathrm{S}\beta x - \mathrm{C}\beta z\big) \\ \mathrm{C}x + \mathrm{S}z \\ y \\ -\gamma\big(\beta ct + \mathrm{S}x - \mathrm{C}z\big) \end{bmatrix} \quad .$$

We observe that the Minkowski square is Lorentz invariant:

$$\begin{aligned}
\langle \mathsf{X}', \mathsf{X}' \rangle_M &= \big(\mathsf{X}'\big)^T \boldsymbol{\eta} \mathsf{X}' \\
&= \begin{bmatrix} \gamma\big(ct+\mathrm{S}\beta x-\mathrm{C}\beta z\big) \\ \mathrm{C}x+\mathrm{S}z \\ y \\ -\gamma\big(\beta ct+\mathrm{S}x-\mathrm{C}z\big) \end{bmatrix}^T \begin{bmatrix} \pm1 & 0 & 0 & 0 \\ 0 & \mp1 & 0 & 0 \\ 0 & 0 & \mp1 & 0 \\ 0 & 0 & 0 & \mp1 \end{bmatrix} \begin{bmatrix} \gamma\big(ct+\mathrm{S}\beta x-\mathrm{C}\beta z\big) \\ \mathrm{C}x+\mathrm{S}z \\ y \\ -\gamma\big(\beta ct+\mathrm{S}x-\mathrm{C}z\big) \end{bmatrix} \\
&= \begin{bmatrix} \gamma\big(ct+\mathrm{S}\beta x-\mathrm{C}\beta z\big) \\ \mathrm{C}x+\mathrm{S}z \\ y \\ -\gamma\big(\beta ct+\mathrm{S}x-\mathrm{C}z\big) \end{bmatrix} \begin{bmatrix} \pm\gamma\big(ct+\mathrm{S}\beta x-\mathrm{C}\beta z\big) \\ \mp\big(\mathrm{C}x+\mathrm{S}z\big) \\ \mp y \\ \pm\gamma\big(\beta ct+\mathrm{S}x-\mathrm{C}z\big) \end{bmatrix} \\
&= \pm\Big[\gamma^2\big(ct+\mathrm{S}\beta x-\mathrm{C}\beta z\big)^2 - \big(\mathrm{C}x+\mathrm{S}z\big)^2 - y^2 + \dots \\
&\qquad\qquad \dots - \gamma^2\big(\beta ct+\mathrm{S}x-\mathrm{C}z\big)^2\Big] \\
&= \pm\Big[\big(ct\big)^2\big(\gamma^2-\gamma^2\beta^2\big) + x^2\big(\gamma^2\beta^2\mathrm{S}^2-\gamma^2\mathrm{S}^2-\mathrm{C}^2\big) - y^2 + \dots \\
&\qquad \dots + z^2\big(-\mathrm{S}^2+\gamma^2\beta^2\mathrm{C}^2-\gamma^2\mathrm{C}^2\big) + 2\mathrm{SC}xz\big(-\gamma^2\beta^2-1+\gamma^2\big)\Big] \\
&= \pm\Big\{\big(ct\big)^2\gamma^2\big(1-\beta^2\big) - x^2\big[\mathrm{C}^2+\mathrm{S}^2\gamma^2\big(1-\beta^2\big)\big] - y^2 + \dots \\
&\qquad \dots - z^2\big[\mathrm{S}^2+\mathrm{C}^2\gamma^2\big(1-\beta^2\big)\big] + 2\mathrm{SC}xz\big[\gamma^2\big(1-\beta^2\big)-1\big]\Big\} \\
&= \pm\Big[\big(ct\big)^2 - x^2 - y^2 - z^2\Big] \\
&= \big(\Delta s\big)^2 = \langle \mathsf{X}, \mathsf{X} \rangle_M \quad .
\end{aligned}$$

The separation 4-vector $x^\mu \in \mathbb{R}^4$ properly invariant with Lorentz scalar $(\Delta s)^2$ under $\boldsymbol{\Lambda}_B\boldsymbol{\Lambda}_R \in \mathcal{L}(\mathbb{R})$.

Theorem 4.2.18 (Reproof of Theorem 3.1.20.) *The 4-velocity is a Lorentz vector with invariant* $U^\mu U_\mu = \pm c^2$.

Proof. The 4-velocity of is the derivative of position with respect to proper time:

$$\mathsf{U} = \frac{d}{d\tau}\begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} c \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

where the proper time τ is the observer's time when he is stationary in the S frame. It is obvious that $\langle \mathsf{U}, \mathsf{U} \rangle_M = \pm c^2$. In the boosted, rotated frame S' associated with $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_B\boldsymbol{\Lambda}_R$ as in the proof of the previous theorem, the

4-velocity becomes

$$\mathsf{U}' = \boldsymbol{\Lambda}\mathsf{U} = \begin{bmatrix} \gamma & \mathrm{S}\beta\gamma & 0 & -\mathrm{C}\beta\gamma \\ 0 & \mathrm{C} & 0 & \mathrm{S} \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & -\mathrm{S}\gamma & 0 & \mathrm{C}\gamma \end{bmatrix} \begin{bmatrix} c \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \gamma c \\ 0 \\ 0 \\ -\beta\gamma c \end{bmatrix} \quad .$$

The Lorentz scalar is

$$\left\langle \mathsf{U}', \mathsf{U}' \right\rangle_M = \left(\mathsf{U}'\right)^T \boldsymbol{\eta} \mathsf{U}' = \begin{bmatrix} \gamma c \\ 0 \\ 0 \\ -\beta\gamma c \end{bmatrix}^T \begin{bmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \mp 1 & 0 & 0 \\ 0 & 0 & \mp 1 & 0 \\ 0 & 0 & 0 & \mp 1 \end{bmatrix} \begin{bmatrix} \gamma c \\ 0 \\ 0 \\ -\beta\gamma c \end{bmatrix} = \pm c^2 \quad .$$

Example 4.2.19 The electromagnetic 4-potential A^μ is not a Lorentz vector in general.

Not every vector that transforms as a vector under spatial rotations is a Lorentz vector. Therefore, it is good to tie things back to position when possible. All derivatives of position transform as Lorentz vectors. First we will show that the form of A^μ common in the introduction to electrodynamics does not transform as a Lorentz vector, and then we will show that it possible to express A^μ in *the Lorenz gauge*—after Ludvig Lorenz rather than Hendrik Lorentz—so that it does transform as a Lorentz vector, and becomes one identically. Given Gaussian units such that $A^\mu = (\phi, \boldsymbol{A})$ where ϕ and $\boldsymbol{A}$ are the scalar and vector potentials satisfying

$$\Box\phi = -\frac{1}{\varepsilon_0}\rho \quad , \quad \text{and} \qquad \Box\boldsymbol{A} = \mu_0\boldsymbol{J} \quad ,$$

for a charge distribution ρ and a current distribution $\boldsymbol{J}$.

$$\big\langle \mathsf{A}, \mathsf{A} \big\rangle_M = \begin{bmatrix} \phi \\ A_x \\ A_y \\ A_z \end{bmatrix}^T \begin{bmatrix} \pm\phi \\ \mp A_x \\ \mp A_y \\ \mp A_z \end{bmatrix} = \pm\left[\phi^2 - A_x^2 - A_y^2 - A_z^2\right] \quad .$$

"LORENTZ GAUGE"

Remark 4.2.20 Maybe also talk about why the transpose was weird in the long proof of Theorem 4.2.15

Remark 4.2.21 FROM WIKI `https://en.wikipedia.org/wiki/Lorentz_transformation`

The transformations are not defined if v is outside these limits. At the speed of light ($v = c$) γ is infinite, and faster than light ($v > c$) γ is an imaginary number, each of which make the transformations unphysical. The space and time coordinates are measurable quantities and numerically must be real numbers.

Remark 4.2.22 Maybe use this for changing the level of aleph:

$$\gamma(c) = \widehat{\infty}$$

Maybe boost direction favors one spatial direction for an eigenspinor basis.

§4.3 The Hermitian Matrix Representation

Before demonstrating the Lorentz transformations of spinors, which are usually written as two-component complex vectors (what we will call rank-1 spinors after developing a context), it will be instructive to show that points in spacetime can be represented as 2×2 matrices in addition to their usual representation as 4-vectors, and that the restricted (proper, orthochronous) Lorentz transformations we have studied as 4×4 matrices acting on 4-vectors have a corresponding representation as 2×2 matrices acting on other 2×2 matrices. Although it is true that an arbitrary complex 4-vector $\mathbb{C}^4$ can be put in one-to-one correspondence with a complex 2×2 matrix (each contain eight numbers: four real parts and imaginary parts), we will continue with the usual treatment of real 4-vectors $V^\mu \in \mathbb{R}^4$ that can be put into one-to-one correspondence with 2×2 hermitian matrices. Later, when we introduce the $\mathbb{R}^(1,3)$ vectors that exceed $\mathbb{R}^4$ while still falling short of the eight parameters in $\mathbb{C}^4$, we will say more about the case for general 2×2 matrices that are not necessarily hermitian. The material in this section mostly follows Jaffe [22] and Williams [18].

==============================

WHY DID I INCLUDE ALL THIS STUFF ABOUT VECTOR SPACES???

==============================

Definition 4.3.1 The **space of 2×2 complex matrices** is called $\mathbb{M}_{2\times2}(\mathbb{C})$. The defining property is

$$\mathbb{M}_{2\times2}(\mathbb{C}) \ni \mathsf{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Longleftrightarrow \quad \{a,b,c,d\} \in \mathbb{C} \ .$$

Theorem 4.3.2 $\mathbb{M}_{2\times2}(\mathbb{C})$ *is vector space but not a scalar product space.*

Proof. Given $\mathsf{M}_1, \mathsf{M}_2 \in \mathbb{M}_{2\times2}(\mathbb{C})$, we have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix} \ ,$$

so he sum of two vectors in $\mathbb{M}_{2\times2}(\mathbb{C})$ is another vector in $\mathbb{M}_{2\times2}(\mathbb{C})$. The 2×2 matrix with four zero entries is the zero vector; such a matrix is the additive identity in the space:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \ .$$

For every $\mathsf{M} \in \mathbb{M}_{2\times2}(\mathbb{C})$, there exists an additive inverse $\mathsf{M}^{-1} \in \mathbb{M}_{2\times2}(\mathbb{C})$ with which M sums to the zero vector:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \ .$$

(Normally M^{-1} will refer to the inverse under matrix multiplication.) Given a scalar $z \in \mathbb{C}$, scalar multiplication of $\mathsf{M} \in \mathbb{M}_{2\times2}(\mathbb{C})$ returns another vector $\mathsf{M}' \in \mathbb{M}_{2\times2}(\mathbb{C})$:

$$z \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} za & zb \\ zc & zd \end{bmatrix} \ .$$

However, the inherent matrix multiplication operation between $\mathsf{M}_1, \mathsf{M}_2 \in \mathbb{M}_{2\times2}(\mathbb{C})$ does not yield a scalar. Rather, this product yields another vector in $\mathbb{M}_{2\times2}(\mathbb{C})$, so this space is not inherently equipped as a scalar product space:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} \big(ae+bg\big) & \big(af+bh\big) \\ \big(ce+dg\big) & \big(cf+dh\big) \end{bmatrix} \quad .$$

Definition 4.3.3 To equip $\mathbb{M}_{2\times2}(\mathbb{C})$ as a scalar product space, we define the **matrix scalar product**

$$\big\langle\big\langle \mathsf{M}_1, \mathsf{M}_2 \big\rangle\big\rangle = \frac{1}{2}\mathrm{tr}(M_1^\dagger M_2) \quad ,$$

where tr indicates the trace, which is the sum of the diagonal entries, and the dagger indicates the hermitian conjugate, which is the complex conjugated transpose.

Definition 4.3.4 A matrix M is **hermitian** if it is equal to its hermitian conjugate, i.e.: $\mathsf{M}^\dagger = \mathsf{M}$, and it is **anti-hermitian** if hermitian conjugation is a sign inversion, i.e.: $\mathsf{M}^\dagger = -\mathsf{M}$. These conditions require that M is a square, $n \times b$ matrix. For any $\boldsymbol{V}, \boldsymbol{W} \in \mathbb{C}^4$, Hermitian matrices satisfy the vector inner product as

$$\big\langle \boldsymbol{V}, \mathsf{M}\boldsymbol{W} \big\rangle = \big\langle \mathsf{M}\boldsymbol{V}, \boldsymbol{W} \big\rangle \quad , \quad \text{with} \qquad \big\langle \boldsymbol{V}, \mathsf{M}\boldsymbol{W} \big\rangle \in \mathbb{R} \quad .$$

Definition 4.3.5 M is a **unitary** matrix if its inverse is equal to its hermitian conjugate, i.e.: $\mathsf{M}^{-1} = \mathsf{M}^\dagger$. M is **special unitary** if it is unitary with unit determinant, i.e.: $\det(\mathsf{M}) = 1$.

Definition 4.3.6 The **space of 2 × 2 hermitian matrices** equipped with the matrix inner product $\langle\langle \cdot, \cdot \rangle\rangle$ is called $\mathbb{H}_2$. The defining property of $\mathsf{M} \in \mathbb{H}_2$ is

$$\mathbb{H} \ni \mathsf{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \Longleftrightarrow \qquad \big\{a, b, c, d\big\} \in \mathbb{C} \quad .$$

Definition 4.3.7 The **Pauli matrices** are

$$\hat{\sigma}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \hat{\sigma}_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad \hat{\sigma}_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad .$$

We define the identity matrix as

$$\hat{\sigma}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad ,$$

so that we may speak of a 2×2 matrix 4-vector $\hat{\sigma}_\mu$. These matrices are hermitian, unitary, and $\hat{\sigma}_i$ is traceless. Pauli matrices commute and anti-commute as

$$[\hat{\sigma}_i, \hat{\sigma}_j] \equiv \hat{\sigma}_i\hat{\sigma}_j - \hat{\sigma}_j\hat{\sigma}_i = 2i\varepsilon_{ijk}\hat{\sigma}_k \ , \quad \text{and} \quad \{\hat{\sigma}_i, \hat{\sigma}_j\} \equiv \hat{\sigma}_i\hat{\sigma}_j + \hat{\sigma}_j\hat{\sigma}_i = 2\delta_{ij}\mathbb{1}_2 \ .$$

The following, equivalent identity is often useful as well:

$$\hat{\sigma}_i\hat{\sigma}_j = \delta_{ij}\mathbb{1}_2 + i\varepsilon_{ijk}\hat{\sigma}_k \quad .$$

Remark 4.3.8 The matrices are really just a representation of the Pauli matrices. Anything that obeys the same commutation relations is automatically another representation. The quaternions obey this, and we are going to incorporate quaternions into the structure of space time. Once we put a quaternion in the metric, we can *complexigate* it by replacing it with a matrix. That may or may not be too complicate with the 4×4 metric, but we should look at it for the 2×2 spinor metric.

Theorem 4.3.9 *The set* $\{\hat{\sigma}_\mu\}$ *is a spanning basis for* $\mathbb{H}_2$*.*

Proof. A spanning basis for a vector space is a set of linearly independent vectors whose scalar multiples can be summed to equal any vector in the space. Linear independence is demonstrated with the scalar product:

$$\big\langle\big\langle \hat{\sigma}_0, \hat{\sigma}_1 \big\rangle\big\rangle = \frac{1}{2}\text{tr}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 0$$

$$\big\langle\big\langle \hat{\sigma}_0, \hat{\sigma}_2 \big\rangle\big\rangle = \frac{1}{2}\text{tr}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = 0$$

$$\big\langle\big\langle \hat{\sigma}_0, \hat{\sigma}_3 \big\rangle\big\rangle = \frac{1}{2}\text{tr}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = 0$$

$$\big\langle\big\langle \hat{\sigma}_1, \hat{\sigma}_2 \big\rangle\big\rangle = \frac{1}{2}\text{tr}\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = 0$$

$$\big\langle\big\langle \hat{\sigma}_1, \hat{\sigma}_3 \big\rangle\big\rangle = \frac{1}{2}\text{tr}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = 0$$

$$\big\langle\big\langle \hat{\sigma}_2, \hat{\sigma}_3 \big\rangle\big\rangle = \frac{1}{2}\text{tr}\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = 0 \quad .$$

The other condition will be satisfied if

$$\mathsf{M}_1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad ,$$

is hermitian and there exist some $z_\mu \in \mathbb{C}$ such that

$$\mathsf{M}_2 = z_0\hat{\sigma}_0 + z_1\hat{\sigma}_1 + z_2\hat{\sigma}_2 + z_3\hat{\sigma}_3 = \begin{bmatrix} z_0 + z_3 & z_1 - iz_2 \\ z_1 + iz_2 & z_0 - z_3 \end{bmatrix} = \mathsf{M}_1 \quad .$$

Equating the real and imaginary parts of the components of M_1 and M_2 gives eight equations in eight unknowns: the real and imaginary parts of the four z_μ. A solution to this system of equations obviously exists for any $a, b, c, d \in \mathbb{C}$, so $\hat{\sigma}_\mu$ is a spanning basis for $\mathbb{H}_2$.

Remark 4.3.10 A simpler basis for $\mathbb{H}_2$ might be

$$\hat{e}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \hat{e}_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \hat{e}_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad \hat{e}_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad ,$$

but $\hat{\sigma}_\mu$ has properties that make it easier to do certain things. Mainly, hermitian basis vectors for the vector space of Hermitian matrices are *convenient*. Also note that in the latter part of the proof of Theorem 4.3.9, we have shown only that every $\mathsf{M} \in \mathbb{H}_2$ can be expressed as some $z_\mu\hat{\sigma}_\mu$. We have not shown that every $z_\mu\hat{\sigma}_\mu$ is hermitian, and we could not possibly show it because this is not the case. If we select $x_\mu \in \mathbb{R}$, however, then $x_\mu\hat{\sigma}_\mu$ is always hermitian, as we will prove now.

Theorem 4.3.11 *Every real-valued 4-vector $x^\mu \in \mathbb{R}^4$ is such that $x^\mu\hat{\sigma}_\mu \in \mathbb{H}_2$.*

Proof. The hermitian conjugate of $\mathsf{M} = x^\mu\hat{\sigma}_\mu$ is

$$\begin{aligned} \mathsf{M}^\dagger &= \begin{bmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{bmatrix}^\dagger \\ &= \begin{bmatrix} \left(x^0 + x^3\right)^* & \left(x^1 + ix^2\right)^* \\ \left(x^1 - ix^2\right)^* & \left(x^0 - x^3\right)^* \end{bmatrix} \\ &= \begin{bmatrix} \left(x^0\right)^* + \left(x^3\right)^* & \left(x^1\right)^* - i\left(x^2\right)^* \\ \left(x^1\right)^* + i\left(x^2\right)^* & \left(x^0\right)^* - \left(x^3\right)^* \end{bmatrix} \quad . \end{aligned}$$

Since x^μ is real, we have $(x^\mu)^* = x^\mu$, and necessarily $\mathsf{M}^\dagger = \mathsf{M}$. This proves that $x^\mu\hat{\sigma}_\mu$ is always a hermitian 2×2 matrix.

Definition 4.3.12 The **hermitian matrix representation of a 4-vector** x^μ is denoted $\widehat{x}$.

Theorem 4.3.13 *Given the transformation $\hat{\sigma} : \mathbb{R}^4 \to \mathbb{H}_2$ such that $\hat{\sigma}(x^\mu) = x^\mu\hat{\sigma}_\mu$, there exists an inverse transformation $\hat{\sigma}^{-1} : \mathbb{H}_2 \to \mathbb{R}^4$ such that*

$$\left(\hat{\sigma}^{-1}(\widehat{x})\right)^\mu = \left\langle\left\langle\hat{\sigma}_\mu, \widehat{x}\right\rangle\right\rangle = x^\mu \quad .$$

Proof. Given

$$\mathsf{M} = \begin{bmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{bmatrix} \quad ,$$

we have

$$\langle\langle \sigma_0, \widehat{x} \rangle\rangle = \frac{1}{2}\mathrm{tr} \begin{bmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{bmatrix} = x^0$$

$$\langle\langle \sigma_1, \widehat{x} \rangle\rangle = \frac{1}{2}\mathrm{tr} \begin{bmatrix} x^1 + ix^2 & x^0 - x^3 \\ x^0 + x^3 & x^1 - ix^2 \end{bmatrix} = x^1$$

$$\langle\langle \sigma_2, \widehat{x} \rangle\rangle = \frac{1}{2}\mathrm{tr} \begin{bmatrix} -ix^1 + x^2 & -i\big(x^0 - x^3\big) \\ i\big(x^0 + x^3\big) & ix^1 + x^2 \end{bmatrix} = x^2$$

$$\langle\langle \sigma_3, \widehat{x} \rangle\rangle = \frac{1}{2}\mathrm{tr} \begin{bmatrix} x^0 + x^3 & x^1 - ix^2 \\ -\big(x^1 + ix^2\big) & -\big(x^0 - x^3\big) \end{bmatrix} = x^3 \quad .$$

Theorem 4.3.14 *The determinant of $\widehat{x}$ is the Minkowski square of x^μ: $x_M^2 = \det(\widehat{x})$.*

CHANGE THIS SO THAT WE HAVE DEFINED A MATRIX COVECTOR ALREADY, NOT A MATRIX CONTRAVECTOR!!!

TALK ABOUT "COMPLETENESS"

IF I HARD CODE THE SIGN OF THE MATRIX COVECTOR, WHAT HAPPENS WHEN I ADD THE SIGN FREEDOM IN THE METRIC????

Proof. Given $x^\mu = (ct, x, y, z)$, we have

$$\widehat{x} = x^\mu \hat{\sigma}_\mu = x^\mu \hat{\sigma}^\nu \eta_{\mu\nu} = \mp \begin{bmatrix} z - ct & x - iy \\ x + iy & -\big(z + ct\big) \end{bmatrix} \quad ,$$

from which it follows that

$$\mp \det(\widehat{x}) = -\big(z - ct\big)\big(z + ct\big) - \big(x - iy\big)\big(x + iy\big) = \big(ct\big)^2 - x^2 - y^2 - z^2 = \mp x_M^2 \quad .$$

Remark 4.3.15 Now that we have established an invertible correspondence between real-valued 4-vectors $x^\mu \in \mathbb{R}^4$ and 2×2 matrices $\widehat{x} \in \mathbb{H}_2$, we will develop the corresponding Lorentz transformations for the $\widehat{x}$ representation.

Lorentz transformations are matrices that preserve the Minkowski square (Definition 4.1.3). They are endomorphisms on a vector space—maps from vectors in the space to other vectors in the space—and we have established that $\mathbb{H}_2$ is a vector space. For a column vector, or a 1×4 matrix, the rules of matrix multiplication dictate that endomorphism will be multiplication from the left with a 4×4 matrix, or from the right with with a 1×1 matrix:

$$\begin{bmatrix} 4\times 4 & \\ & \end{bmatrix}\begin{bmatrix} 4\times 1 \\ \\ \end{bmatrix} = \begin{bmatrix} 4\times 1 \\ \\ \end{bmatrix} \quad , \quad \text{or} \qquad \begin{bmatrix} 4\times 1 \\ \\ \end{bmatrix}\begin{bmatrix} 1\times 1 \end{bmatrix} = \begin{bmatrix} 4\times 1 \\ \\ \end{bmatrix} \quad .$$

The latter case is trivial multiplication by a scalar, so it cannot suffice for general Lorentz transformations that mix space and time. Presently, however, endomorphisms on 2×2 matrices allow matrix multiplication from the left or right with 2×2 matrices, and we have no way to prefer one direction over the other. Thus, we must write Lorentz transformations for $\widehat{x}$ as

$$\widehat{x}_{\mu'\nu'} = \Lambda^{\mu}_{\mu'}\Lambda^{\nu}_{\nu'}\widehat{x}_{\mu\nu} \quad \longleftrightarrow \quad \widehat{x}' = \mathbf{\Lambda}\widehat{x}\mathbf{\Lambda} \quad .$$

WHICH ONE IS THE CONJUGATE?????????
COMPARE TO METRIC
We also know that a square matrix is the representation of a tensor with two lower indices, so the formalism $\mathbf{\Lambda} = \Lambda^{\mu'}_{\mu}$ suggests that we will have to use to matrices to Lorentz transform $\widehat{x}$.

Definition 4.3.16 The **hermitian matrix representation of a Lorentz transformation** is denoted $\mathsf{\Lambda}(\mathbf{\Lambda})$, and it has the property

$$\mathbf{\Lambda} \in \mathcal{L} \quad \Longleftrightarrow \quad \det\big(\mathsf{\Lambda}(\mathbf{\Lambda})\widehat{x}\big) = \det(\widehat{x}) \quad .$$

Example 4.3.17 Lorentz transformations for $\widehat{\mathbf{x}} \in \mathbb{H}_2$. It is well established by now that a pure boost in the x-direction is written

$$\mathsf{X}' = \mathbf{\Lambda}_{4\times4}\mathsf{X} = \begin{bmatrix} \gamma\big(ct - \beta x\big) \\ \pm\gamma\big(x - vt\big) \\ y \\ z \end{bmatrix} \quad .$$

Now $\hat{\sigma} : \mathbb{R}^4 \to \mathbb{H}_2$ acts (in matrix index notation) as

$$\hat{\sigma}(\mathsf{X}') = \mathsf{X}'_{\mu}\hat{\sigma}_{\mu} = \begin{bmatrix} \gamma\big(ct - \beta x\big) + z & \pm\gamma\big(x - vt\big) - iy \\ \pm\gamma\big(x - vt\big) + iy & \gamma\big(ct - \beta x\big) - z \end{bmatrix} = \widehat{x}' \quad ,$$

and we must find what $\mathbf{\Lambda}_{2\times2}$ satisfy

$$\mathbf{\Lambda}_{2\times2}\begin{bmatrix} ct + x & x - iy \\ x + iy & ct - z \end{bmatrix}\mathbf{\Lambda}_{2\times2} = \hat{\sigma}(\mathsf{X}') \quad .$$

We will avoid making distinctions among the different kinds of $\mathbf{\Lambda}$, and it will always be clear from the context what is meant. Proving without heuristics that Lorentz transformations for $\widehat{x} \in \mathbb{H}_2$ take the form $\widehat{x}' = \mathbf{\Lambda}\widehat{x}\mathbf{\Lambda}^{\dagger}$ requires a modest amount of of further machinery, and we must prove that this the correct form of the Lorentz transformation before we continue to deduce the properties these $\mathbf{\Lambda}_{2\times2}$ matrices, which are the elements of the hermitian matrix representation of $\mathcal{L}$.

WE WANT TO PRESERVE THE DETERMINANT!!!!

Definition 4.3.18 A set of matrices is said to be an **irreducible set** if any matrix that commutes with every matrix in the set is a multiple of the identity $\mathbb{1}_n$.

Definition 4.3.19 The group of all 2×2, complex matrices with unit determinant is called the **special linear group** and it is denoted $\mathrm{SL}(2,\mathbb{C})$.

Lemma 4.3.20 *The Pauli matrices are an irreducible set.*

Lemma 4.3.21 *The transformation $\hat{\sigma}:\mathbb{R}^4\to\mathbb{H}_2$ and its inverse $\hat{\sigma}^{-1}:\mathbb{H}_2\to\mathbb{R}^4$ are linear transformations.*

Proof. Following Definition 4.1.1, we must show $T(\boldsymbol{V}+\boldsymbol{W})=T(\boldsymbol{V})+T(\boldsymbol{W})$ and $T(c\boldsymbol{V})=cT(\boldsymbol{V})$ for $\hat{\sigma}$ and its inverse $\hat{\sigma}^{-1}$. Given

$$x^\mu=(ct,x,y,z)\quad,\quad\text{and}\qquad y^\mu=(ct',x',y',z')\quad,$$

such that

$$\hat{\sigma}(x^\mu)=\begin{bmatrix}ct+x & x-iy\\ x+iy & ct-z\end{bmatrix}\quad,\quad\text{and}\qquad \hat{\sigma}(y^\mu)=\begin{bmatrix}ct'+x' & x'-iy'\\ x'+iy' & ct'-z'\end{bmatrix}\quad,$$

$\hat{\sigma}$ preserves additivity as

$$\begin{aligned}\hat{\sigma}(x^\mu+y^\mu)&=\begin{bmatrix}c\big(t+t'\big)+\big(x+x'\big) & \big(x+x'\big)-i\big(y+y'\big)\\ \big(x+x'\big)+i\big(y+y'\big) & c\big(t+t'\big)-\big(x+x'\big)\end{bmatrix}\\ &=\hat{\sigma}(x^\mu)+\hat{\sigma}(y^\mu)\quad.\end{aligned}$$

$\hat{\sigma}$ preserves scalar multiplication as

$$\hat{\sigma}(\xi x^\mu)=\widehat{x}=\begin{bmatrix}\xi\big(ct+x\big) & \xi\big(x-iy\big)\\ \xi\big(x+iy\big) & \xi\big(ct-z\big)\end{bmatrix}=\xi\hat{\sigma}(x^\mu)\quad.$$

$\hat{\sigma}$ satisfies the definition of linearity, and it remains to demonstrate the same for $\hat{\sigma}^{-1}$. Observe that

$$\begin{aligned}\big[\hat{\sigma}^{-1}(\widehat{x}+\widehat{y})\big]^\mu&=\frac{1}{2}\mathrm{tr}\big[\hat{\sigma}_\mu\big(\widehat{x}+\widehat{y}\big)\big])\\ &=\frac{1}{2}\mathrm{tr}\big[\hat{\sigma}_\mu\widehat{x}+\hat{\sigma}_\mu\widehat{y}\big]\end{aligned}$$

$$= \frac{1}{2}\mathrm{tr}\big(\hat{\sigma}_\mu \widehat{x}\big) + \frac{1}{2}\mathrm{tr}\big(\hat{\sigma}_\mu \widehat{x}\big)$$
$$= \hat{\sigma}^{-1}(\widehat{x}) + \hat{\sigma}^{-1}(\widehat{y}) \quad ,$$

and that, likewise,

$$\big[\hat{\sigma}^{-1}(c\widehat{x})\big]^\mu = \frac{1}{2}\mathrm{tr}(\hat{\sigma}_\mu c\widehat{x}) = \frac{c}{2}\mathrm{tr}(\hat{\sigma}_\mu \widehat{x}) \quad .$$

$\hat{\sigma}$ and its inverse are linear transformations.

Remark 4.3.22 We have proven that the transformations between are $\mathbb{R}^4$ and $\mathbb{H}_2$ are linear, but we have only written the exact form of the operation for $\hat{\sigma}^{-1}$. Namely, we do not want matrix algebra operation takes $\widehat{x} \to \widehat{\mathbf{\Lambda} x}$ in one-to-one correspondence with $x^\mu \to \mathbf{\Lambda} x^\mu$. So, now we will prove that $\hat{\sigma}$ has a certain form, and we will proceed to identify that form as the hermitian matrix representation of the restricted Lorentz transformation $\mathbf{\Lambda} \in \mathcal{L}_+^\uparrow$.

Main Theorem 4.3.23 *Every determinant-preserving, orthochronous linear transformation $T : \mathbb{H}_2 \to \mathbb{H}_2$ may be expressed as*

$$T(\widehat{x}) = \mathsf{M}^\dagger \widehat{x} \mathsf{M} \quad ,$$

where $\mathsf{M} \in \mathrm{SL}(2, \mathbb{C})$.

Proof. Following Jaffe [22], we first consider $\mathsf{A}, \mathsf{B} \in \mathrm{SL}(2, \mathbb{C})$ such that

$$T : \mathbb{H}_2 \to \mathbb{H}_2 \quad , \qquad \text{and} \qquad T(\widehat{x}) = \mathsf{A}\widehat{x}\mathsf{B} \quad . \tag{4.6}$$

As T is an endomorphism mapping vectors in $\mathbb{H}_2$ to other vectors in $\mathbb{H}_2$, the constraint that $T(\widehat{x})$ must be hermitian is expressed as

$$\mathsf{A}\widehat{x}\mathsf{B} = \big(\mathsf{A}\widehat{x}\mathsf{B}\big)^\dagger \qquad \iff \qquad \mathsf{A}\widehat{x}\mathsf{B} = \mathsf{B}^\dagger \widehat{x} \mathsf{A} \quad ,$$

where $\widehat{x} = \widehat{x}^\dagger$ also follows from the hermiticity of $\widehat{x}$. We multiply the latter expression from the left by $(\mathsf{B}^\dagger)^{-1}$ and from the right by B^{-1} to obtain

$$\big(\mathsf{B}^\dagger\big)^{-1} \mathsf{A} \widehat{x} = \widehat{x} \mathsf{A}^\dagger \mathsf{B}^{-1} \quad .$$

Without loss of generality (since $\widehat{x}$ is an arbitrary hermitian matrix), we may choose $\widehat{x} = \mathbb{1}_2$ so that

$$\big(\mathsf{B}^\dagger\big)^{-1} \mathsf{A} = \mathsf{A}^\dagger \mathsf{B}^{-1} \quad .$$

Since the inverse of the hermitian conjugate is the conjugate of the inverse, the product $(\mathsf{B}^\dagger)^{-1}\mathsf{A} = \Sigma$ is a hermitian matrix; the matrices A and B are determined by the transformation T, not the matrix $\widehat{x}$, so Σ must be hermitian for any $\widehat{x}$. We will proceed in the hermitian case, and then we will rule out the case when $\mathbf{\Sigma}$ is not Hermitian at the end. Hermitian matrices act equally from the left and right (Definition 4.3.4), so this matrix commutes with $\widehat{x}$:

$$\big[\mathsf{\Sigma},\widehat{x}\big] \equiv \mathsf{\Sigma}\widehat{x} - \widehat{x}\mathsf{\Sigma} = 0 \quad .$$

Again relying on the fact that $\widehat{x}$ is an arbitrary hermitian matrix, we can choose it to be one of the Pauli matrices, in which case

$$\big[\mathsf{\Sigma},\sigma_i\big] = 0 \quad .$$

Since we have proven in Lemma 4.3.20 that $\{\sigma_i\}$ is an *irreducible set*, the matrix $\mathsf{\Sigma}$ must be a multiple of $\mathbb{1}_2$ (Definition 4.3.18), and since $\mathsf{\Sigma}$ is hermitian, the scalar multiple must be real-valued:

$$\xi\mathbb{1} = \big(\xi\mathbb{1}\big)^\dagger = \xi^*\mathbb{1} \qquad \Longleftrightarrow \qquad \xi\in\mathbb{R} \quad .$$

Furthermore, since (i) we have selected $\mathsf{A},\mathsf{B}\in\mathrm{SL}(2,\mathbb{C})$ meaning A and B have unit determinant, since (ii) the determinant of the hermitian conjugate is the complex conjugate of the determinant, since (iii) the determinant of an inverse is the inverse of the determinant, and since (iv) the determinant of a product is the product of determinants, we know

$$\det(\mathsf{\Sigma}) = \det\Big[\big(\mathsf{B}^\dagger\big)^{-1}\Big]\det(\mathsf{A}) = \frac{\det(\mathsf{A})}{\det(\mathsf{B}^\dagger)} = \frac{1}{1^*} = 1 \qquad \Longrightarrow \qquad \xi = \pm 1 \quad .$$

If $\mathsf{\Sigma}$ is a multiple of the identity with unit determinant, then the multiple is ± 1. Therefore,

$$\big(\mathsf{B}^\dagger\big)^{-1}\mathsf{A} = \pm\mathbb{1} \qquad \Longrightarrow \qquad \mathsf{A} = \pm\mathsf{B}^\dagger.$$

Therefore, referring to (4.6), every linear transformation $T:\mathbb{H}_2\to\mathbb{H}_2$ in the form $T(\widehat{x}) = \mathsf{A}\widehat{x}\mathsf{B}$ for some $\mathsf{A},\mathsf{B}\in\mathrm{SL}(2,\mathbb{C})$ must be expressed as

$$T(\widehat{x}) = \pm\mathsf{A}^\dagger\widehat{x}\mathsf{A} \quad .$$

Since $\det(\mathsf{A}) = 1$, this is automatically determinant-preserving.

Now we will show that the freedom to choose a sign reflects the choice between orthochronous and non-orthochronous $\mathbf{\Lambda}\in\mathcal{L}$. We have established

$$\hat{\sigma}(x^\mu) = \qquad \widehat{x^\mu} \qquad = x^\mu\hat{\sigma}_\mu$$
$$\Big[\hat{\sigma}^{-1}\left(\widehat{x^\mu}\right)\Big]^\rho = \big\langle\big\langle\hat{\sigma}_\rho,\widehat{x^\mu}\big\rangle\big\rangle = x^\rho \quad ,$$

so, given a Lorentz transformation $x^{\mu'} = \Lambda^{\mu'}{}_\mu x^\mu$, we will let $\mu'\to\nu$ to write

$$\hat{\sigma}(\Lambda^\nu{}_\mu x^\mu) = \qquad \widehat{\Lambda^\nu{}_\mu x^\mu} \qquad = \Lambda^\nu{}_\mu x^\mu\hat{\sigma}_\nu$$
$$\left[\hat{\sigma}^{-1}\left(\widehat{\Lambda^\nu{}_\mu x^\mu}\right)\right]^\rho = \big\langle\big\langle\hat{\sigma}_\rho,\widehat{\Lambda^\nu{}_\mu x^\mu}\big\rangle\big\rangle = \Lambda^\rho{}_\mu x^\mu \quad ,$$

and so far in this proof we have determined that

$$\widehat{\Lambda^\nu{}_\mu x^\mu} \equiv \Lambda^\nu{}_\mu x^\mu \hat{\sigma}_\nu = \pm \mathsf{A}^\dagger \widehat{x^\mu} \mathsf{A} = \pm \mathsf{A}^\dagger x^\mu \hat{\sigma}_\mu \mathsf{A} \quad .$$

Inserting this expression into the expression for the inverse yields

$$\Lambda^\rho{}_\mu x^\mu = \left[\hat{\sigma}^{-1} \left(\widehat{\Lambda^\nu{}_\mu x^\mu} \right) \right]^\rho = \pm \big\langle \big\langle \hat{\sigma}_\rho, \mathsf{A}^\dagger x^\mu \hat{\sigma}_\mu \mathsf{A} \big\rangle \big\rangle = \pm \frac{1}{2} \mathrm{tr} \left[\hat{\sigma}_\rho \mathsf{A}^\dagger x^\mu \hat{\sigma}_\mu \mathsf{A} \right]$$

Orthochronicity refers to the sign of $\Lambda^0{}_0$, so we choose $\rho = 0$ and use $\hat{\sigma}_0 = \mathbb{1}$ to write

$$\Lambda^0{}_\mu x^\mu = \pm \frac{1}{2} \mathrm{tr}[\mathsf{A}^\dagger x^\mu \hat{\sigma}_\mu \mathsf{A}]$$

Now we recall x^μ are the components of a vector—they are scalars and not the vector itself—so they may be factored out of the trace and divided away to give

$$\Lambda^0{}_\mu = \pm \frac{1}{2} \mathrm{tr}[\mathsf{A}^\dagger \hat{\sigma}_\mu \mathsf{A}] \quad \Longrightarrow \quad \Lambda^0{}_0 = \pm \frac{1}{2} \mathrm{tr}[\mathsf{A}^\dagger \mathsf{A}] \quad .$$

Now we have demonstrated that given some $\mathsf{M} \in \mathrm{SL}(2, \mathbb{C})$

$$T(\widehat{x}) = \pm \mathsf{M} \widehat{x} \mathsf{M}^\dagger \quad ,$$

is a determinant-preserving, *orthochronous* linear transformation on $\mathbb{H}_2$.

Remark 4.3.24 "Since the linear, determinant-preserving transformations $\widehat{x} \to \widehat{x}'$ are in one-to-one correspondence with Lorentz transformations, we know that they have six parameters. (JAFFE p11 TOP)

Corollary 4.3.25 *NO!!!!!!! THERE ARE TWO DIFFERENT X HATS FOR EACH LORENTZ XFORM. THIS COROLLARY PROVES THAT THERE IS ONE AND ONLY ONE* $\mathsf{M} \in \mathrm{SL}(2, \mathbb{C})$ *FOR EACH X HAT!!!!!!*

There exists one and only one $\mathsf{M} \in \mathrm{SL}(2, \mathbb{C})$ *corresponding to each restricted Lorentz transformation* $\mathbf{\Lambda} \in \mathcal{L}_+^\uparrow$.

Proof. Continuing to follow Jaffe [22], we will prove this corollary by contradiction. Suppose $\mathsf{A}, \mathsf{B} \in \mathrm{SL}(2, \mathbb{C})$ and

$$\mathsf{A}^\dagger \widehat{x} \mathsf{A} = \mathsf{B}^\dagger \widehat{x} \mathsf{B} \quad .$$

Multiplying from the left by $(\mathsf{A}^\dagger)^{-1}$ and from the right by A^{-1} gives

$$\widehat{x} = (\mathsf{A}^\dagger)^{-1} \mathsf{B}^\dagger \widehat{x} \mathsf{B} \mathsf{A}^{-1} = \big(\mathsf{B} \mathsf{A}^{-1} \big)^\dagger \widehat{x} \mathsf{B} \mathsf{A}^{-1} \quad . \tag{4.7}$$

Letting $\widehat{x} = \mathbb{1}$ shows that $(\mathsf{B}\mathsf{A}^{-1})$ is unitary:

$$\mathbb{1} = \big(\mathsf{B} \mathsf{A}^{-1} \big)^\dagger \mathsf{B} \mathsf{A}^{-1} \quad ,$$

and this $(\mathsf{B}\mathsf{A}^{-1})$ commutes with an arbitrary $\hat{x}$. Commutation is verified by writing

$$\big(\mathsf{B} \mathsf{A}^{-1} \big) \widehat{x} = \widehat{x} \big(\mathsf{B} \mathsf{A}^{-1} \big) \quad ,$$

and then multiplying from the left with $(\mathsf{BA}^{-1})^{-1}$, which is equal to $(\mathsf{BA}^{-1})^{\dagger}$ by unitarity:

$$\widehat{x} = \big(\mathsf{BA}^{-1}\big)^{\dagger}\widehat{x}\big(\mathsf{BA}^{-1}\big) = \big(\mathsf{A}^{-1}\big)^{\dagger}\mathsf{B}^{\dagger}\widehat{x}\big(\mathsf{BA}^{-1}\big) \quad .$$

Noting that the conjugate inverse is the inverse conjugate, this expression appears in (4.7), so we have verified commutativity. Now we choose $\widehat{x} = \hat{\sigma}_i$ and conclude that since BA^{-1} commutes with every element of an irreducible set, it must be a multiple of the identity: $\mathsf{BA}^{-1} = \xi\mathbb{1}_2$. Since A and B have unit determinant, and $\det(\xi\mathbb{1}_2) = \xi^2 = 1$, we have found again that $\xi = \pm 1$. Only $\xi = +1$ is consistent with orthochronicity, so the corollary is proven.

Main Theorem 4.3.26 *THERE IS A TWO-TO-ONE CORRESPONDENCE!!!!!!!*

Article 4.3.27 PROVE AGAIN THAT EVERY $\mathbf{\Lambda}$ is $\mathbf{\Lambda}_B\mathbf{\Lambda}_B$ [22]

Remark 4.3.28 $\widehat{x}$ is a rank-2 spinor. now we move to rank-1 spinors in the next section.

§4.4 The Spinor Representation

PROBABLY QUOTE FIRST FOUR PARAGRAPHS FROM STEANE [1] PLUS 7,8?? PLUS THE WHOLE DEFINITION with a, b on p2???

"To every tensor of rank k there corresponds a spinor of rank $2k$." We saw this in the previous section when we obtained the hermitian rank-2 spinor $(\widehat{x} = \widehat{x})_{\mu\nu}$ from from the rank-1 4-vector x^μ. Some tensors are associated with a a spinor of the same rank: a null 4-vector can be associated with a spinor of the same rank. Therefore, when we try to make ordinary angular momentum states from the MCM unit cell, we should look for the different kinds of null vectors that we can construct in the unit cell. When the spinor actually corresponds to a null 4-vector with an angle and sign, we will want to associate the choice of sign with $\Sigma^\pm$, and, to the extent that we want to build time arrow spinors for eigenvalue algebras of time arrow operators, we can probably use the arrow of time in $\mathcal{H}$—the direction of x^0—for setting reference angles.

The minus -1 for spinors have no effect on individual spinors, but they matter when we compare different ones. This is like what we want to do by constructing the complex plane with a quaternion rather than the imaginary number. In the absence of any other quaternion, it will be indistinguishable from the imaginary number, and in the presence of another one we get the identities $\mathbf{q_i}\mathbf{q_j} = \varepsilon_{ijk}\mathbf{q_k}$ and $\mathbf{q_i}\mathbf{q_j}\mathbf{q_k} = -1$, with the former containing the all-relevant minus sign for odd permutations of ijk.

=======================

As in the previous section where we developed the correspondence $x^\mu \in \mathbb{R}^4$ and $\widehat{x} \in \mathbb{H}_2$, it will be valuable for pedagogy if we likewise develop in this section the relationship of rank-1 spinors to objects in $\mathbb{R}^4$. When we introduce the new complex vector representation in the following section, we will then be able to smoothly state the similar results for adapting the new representation of $\mathcal{L}$ to hermitian matrices and rank-1 spinors. When we examine the Lorentz transformations of bispinors in Section XXX, we will examine the real and complex representations side by side.

==========================

Theorem 4.4.1 *When a spinor is a vector with a flagpole represented as a two-component complex vector (CHECK VECTOR??), two spinors pointing oppositely in the direction are orthogonal, and two spinors pointing in the same direction are shifted by a complex phase if their flags point in opposite directions.*

Article 4.4.2 This choice of sign that leaves the Lorentz norm unchanged associated with right- and left-handed spinors, so this should be adapted to real or complex components in $\mathbb{R}^{1,3}$.

Theorem 4.4.3 *Unitary* 2×2 *matrices preserve the length of the flag pole, which requires preserving* $|r|^2 = |a|^2 + |b|^2$ *in the spinor.*

Article 4.4.4 In the previous section, we found that preservation of the Lorentz norm for 4-vectors was adapted to 2×2 matrices as the preservation of the determinant. Now we need to establish a similar constraint.

Theorem 4.4.5 *The components of the flagpole vector of a given rank-1 spinor are*

$$x = \boldsymbol{s}^\dagger \hat{\sigma}_x \boldsymbol{s} \qquad y = \boldsymbol{s}^\dagger \hat{\sigma}_y \boldsymbol{s} \qquad z = \boldsymbol{s}^\dagger \hat{\sigma}_z \boldsymbol{s} \quad .$$

$$\boldsymbol{r} = \boldsymbol{s}^\dagger \hat{\boldsymbol{\sigma}} \boldsymbol{s} = \big\langle s \big| \boldsymbol{\sigma} \big| s \big\rangle \quad .$$

Definition 4.4.6 The exponential of a matrix is defined by the infinite series expansion of the exponential function.

Example 4.4.7 Rotation of rank-1 spinors AND ANGLE DOUBLING.
Steane says this can be done with trig or with matrices, but he only shows matrices.

Example 4.4.8 Boost of rank-1 spinors.

Theorem 4.4.9 *Show that every Lorentz transform can be represented as* $\Lambda_B \Lambda_R$*, and that since this is a linear transformation, it must have been the case in the* $x^\mu \in \mathbb{R}^4$ *representation as well.*

Example 4.4.10 Spinors are weird because rotation by 2π causes the spinor to acquire a sign -1: all spin rotation matrices are $-\mathbb{1}_2$ when $\theta = 2\pi$. The flag and flagpole were rotated correctly, but there is some overall minus sign with no ready interpretation in the spinor case. We need to show this in this example.

STEANE [1] says that issue lies with SO(3) not being a simply connected group.

Article 4.4.11 Rank-2 spinors. We have a correspondence between $\boldsymbol{x} \in \mathbb{R}^3$ and $\widehat{\boldsymbol{x}}$ in $\mathbb{H}_2$, just like in the previous section. Rather than writing $\hat{\sigma} : \mathbb{R}^4 \to \mathbb{H}_2$ such that $\hat{\sigma}(x^\mu) = x^\mu \hat{\sigma}_\mu$, now we write $\hat{\sigma} : \mathbb{R}^3 \to \mathbb{H}_2$ such that $\hat{\sigma}(\boldsymbol{x}) = \boldsymbol{x} \cdot \hat{\boldsymbol{\sigma}}$. The determinant is $|\boldsymbol{x}|^2$ and we have already established that determinant preserving linear transformations on $\mathbb{H}_1$ must take the form

$$\widehat{\boldsymbol{x}'} = \mathsf{M}^\dagger \widehat{\boldsymbol{x}} \mathsf{M} \quad .$$

WHY IS M UNITARY? I know it has unit determinant.

Theorem 4.4.12 *STEANE [1] top of p7: prove cases of transformations.*

Remark 4.4.13 If a rank-2 spinor in $\mathbb{H}_2$ is the outer product of two rank-1 spinors, and then

$$\boldsymbol{s}\boldsymbol{s}^\dagger = \mathsf{M}^\dagger \boldsymbol{s}\boldsymbol{s}^\dagger \mathsf{M} \qquad \Longrightarrow \qquad \boldsymbol{s}' = \mathsf{M}^\dagger \boldsymbol{s} \quad ,$$

and the transformation of $\boldsymbol{s}^\dagger$ follows from the conjugation of the implied expression. Since M and $\mathsf{M}^\dagger$ are both in $\mathrm{SL}(2, \mathbb{C})$, we can choose $\Lambda = \mathsf{M}^\dagger$ to group the hermitian conjugates as

$$\boldsymbol{s}' = \Lambda \boldsymbol{s} \quad , \qquad \text{and} \qquad \left(\boldsymbol{s}'\right)^\dagger = \boldsymbol{s}^\dagger \Lambda^\dagger \quad .$$

KKKKKKKKKKKKKKKKKKKKK

Definition 4.4.14 Given a null 4-vector V^μ, u is a **right-handed rank-1 spinor** if

$$V^\mu = \langle u | \hat{\sigma}^\mu | u \rangle \quad ,$$

and it is a **left-handed rank-1 spinor** if

$$g_{\mu\nu} V^\nu = V_\mu = V = \langle u | \hat{\sigma}^\mu | u \rangle \quad .$$

Right-handed spinors are called contraspinors because they corresponded to null contravectors with upper indices, and left-handed one are called cospinors because they correspond to null covectors with lower indices.

NEED TO CHECK CONVENTION FOR INDEX PLACEMENT ON $\hat{\sigma}$!!!!!!!

Theorem 4.4.15 *Every 2×2 matrix with unit determinant Lorentz transforms as a spinor.*

Theorem 4.4.16 *If $\boldsymbol{s}' = \Lambda(v)\boldsymbol{s}$ is the Lorentz transform of a right-handed rank-1 spinor, then the same change of reference is written for a left-handed spinor as $\tilde{\boldsymbol{s}}' = (\Lambda^\dagger)^{-1}\tilde{\boldsymbol{s}} = \Lambda(-v)\tilde{\boldsymbol{s}}$.*

Theorem 4.4.17 *A Dirac bi-spinor $\Psi = (\phi_R, \phi_L)$ is composed of a pair of spinors, one of each handedness. From the two associated null 4-vectors, one can extract two orthogonal non-null 4-vectors*

$$V^\mu = \Psi^\dagger \gamma^0 \gamma^\mu \Psi \quad , \quad \textit{and} \qquad W^\mu = \Psi^\dagger \gamma^0 \gamma^\mu \gamma^5 \Psi \quad .$$

ZEE SAID SOMETHING ABOUT HOW γ^5 IS REQUIRED SOMEWHERE DUE TO THE METRIC, SO WE MAY NEED TO MAKE SOME COMMENTS ABOUT THIS!

Proof. These vectors are the 4-velocity and 4-spin. Therefore, after we select the null 4-vectors for the Pauli spinors, we will want to examine what kinds of 4-vectors we can make in general for the Dirac bi-spinor. Obviously, the 4-velocity is timelike. What is the restriction on the 4-spin? Is this the Pauli–Lubanski vector? (WHO IS LUBANSKI?)

Theorem 4.4.18 *Lorentz transform of a Dirac bi-spinor is*

$$\begin{bmatrix} -m & E + \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ E - \boldsymbol{\sigma}\cdot\boldsymbol{p} & -m \end{bmatrix} \begin{bmatrix} \phi_R(\boldsymbol{p}) \\ \chi_L(\boldsymbol{p}) \end{bmatrix} = 0$$

This is the Dirac equation. Under parity, the parts of a Dirac spinor swap and $\boldsymbol{\sigma} \to bm\sigma$, so the Dirac equation is parity invariant.

I THINK I READ THAT THE DIRAC BI-SPINOR IS THE ONLY POSSIBLE BI-SPINOR REPRESENTATION OF THE LORENTZ GROUP.

Theorem 4.4.19 *Angle doubling*

kkkkkkkkkkkkkkkkkk
kkkkkkkkkkkkkkkkkk
kkkkkkkkkkkkkkkkkk

The Weyl representation or spinor map is a pair of surjective homomorphisms from SL(2,C) to SO+(1, 3). They form a matched pair under parity transformations, corresponding to left and right chiral spinors.

URL:

Remark 4.4.20 If $\boldsymbol{\eta\Lambda\eta} = \boldsymbol{\Lambda}^{-1}$ in the spinor representation, then add this result to the FUNDAMENTALS section.

§4.5 The Complex Vector Representation and its Extension to Spinors

Williams [18]: end of p4 onto p5: talks about the complex inner product

========================

WRONG: The fundamental representation of the Lorentz group is a real-valued 4-vector $x^\mu \in \mathbb{R}^4$. The realization of the Lorentz group in this representation consists of boosts and rotations, and their products. Weyl spinors also form a representation of the Lorentz group. By the two-to-one SOMETHING-ism between SO(3) and SU(2), we have two different spinor representations called $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$. This comes from our ability to write points in space-time as

$$x^\mu = (x^0, x^1, x^2, x^3) \quad\longleftrightarrow\quad \tilde{x} = \begin{bmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{bmatrix}$$

The matrix is a rank-2 spinor and we pull out rank-1 spinors via the 2-to-1 correspondence (somehow).

So, while such things are very well known [1,3,5,22], presently, we will use another, less common representation of the Lorentz group $x^\mu \in \mathcal{H}$. To prove certain results and make appropriate generalizations to the MCM, we will need to find the realizations of the Lorentz group for these complex vectors and the spinors constructed from them. Luckily, Roman has worked out the case for imaginary time and real space in Appendix 3-2 to [3].

> "Consider the four-dimensional Minkowski–Lorentz space of events,
>
> $$x_1 = x, \quad x_2 = y, \quad x_3 = z, \quad x_4 = ict.$$
>
> The points of this space form the manifold of special relativity. Let
>
> $$x_\mu \to x'_\mu = \alpha_{\mu\nu} x_\nu$$
>
> be a homogeneous linear mapping of the space onto itself. (The usual dummy index summation convention is adopted; summation over

repeated Greek indices from 1 to 4 is understood.) The coefficients of the transformation are subject to the reality conditions

$$\begin{array}{rll} \alpha_{ik}, & i,k=1,2,3 & \textit{real}, \\ \alpha_{i4} \text{ and } \alpha_{4i}, & i=1,2,3 & \textit{imaginary}, \\ \alpha_{44}, & & \textit{real}. \end{array}$$

Let the coefficients be further restricted by the *orthogonality conditions*

$$\alpha_{\mu\nu}\alpha_{\mu\rho} = \delta_{\nu\rho}$$
$$\alpha_{\nu\mu}\alpha_{\rho\mu} = \delta_{\nu\rho}$$

It then follows that [*these*] transformations leave the length of a Minkowski vector invariant:

$$x_\mu x_\mu \to x'_\mu x'_\mu = x_\mu x_\mu,$$

and also

$$x'_\mu x'_\mu \to x_\mu x_\mu = x'_\mu x'_\mu.$$

It can be easily seen that the set of all transformations with this property form a group. This group is called the *homogeneous Lorentz group L*. Its unit element is given by

$$\alpha_{\mu\nu} = \delta_{\mu\nu}$$

and the inverse of [*Equation (4.5)*] reads

$$x'_\mu \to x_\mu = \alpha_{\nu\mu}x'_\nu.$$

It is easy to see that [...] there are six independent quantities $\alpha_{\mu\nu}$ (three real and three imaginary). We will consider them as *parameters* of a group element."

===========================

The structure of Lorentz transformations is different in the cases of real- or complex-valued 4 vectors. The conventions are laid out by Roman as follows [3].

Roman has used the convention for strictly imaginary time, but we want to use either time or space. The conventions generalize as

======================

By making timelike part of a Minkowski vector imaginary, the boost parameters which are usually described as rotation by a complex angle acquire a requisite factor of i. Other than that, the realization of the Lorentz group is the same. Below, we will lay out the full case for making space and time alternatingly imaginary in some convention, we will use the 0-index to describe time rather than the 4-index used by Roman, and we will use tensor indices rather than the all-lower matrix indices used by Roman.

For the MCM specifically, we have many times raised a question about how timelike motion in one of $\Sigma^\pm$ might be linked to spacelike motion in $\Sigma^\mp$ at $\mathcal{H}$ or $\varnothing$ when the metric signature in $\Sigma^\pm$ is $\{+---\pm\}$ or $\{-+++\pm\}$. The purpose in using a complex vector representation of the Lorentz group will be to remove the non-Euclidean character of the metric (where we will call a metric Euclidean if its eigenvalues are all $+1$ or all -1). To create matching conditions for smooth evolutions across bounding branes, we will employ the alternating phase condition suggested in [10]. This will require that we alternate between the scheme for imaginary time detailed by Roman and another one in which time is real and space is imaginary. There exist many notations for inserting a conditional imaginary number into an expression, for example

$$e^{\imath\theta} = \cos\theta + i\sin\theta \equiv \text{cis}\theta$$

with an appropriate restriction on θ. However, since we will want to consider cases for 1 and i without considering -1 and $-i$, we will introduce a new symbol $\mathbb{i}^{\updownarrows}$ which allows us to write $x^\mu \in \mathcal{H}$ as

$$x^0 = \mathbb{i}^{\updownarrows} t \quad , \quad x^1 = \mathbb{i}^{\downuparrows} x \quad , \quad x^2 = \mathbb{i}^{\downuparrows} y \quad , \quad x^3 = \mathbb{i}^{\downuparrows} z \quad .$$

Hopefully this cluttered notation is not too cluttered: $\mathbb{i}^\uparrow = 1$ and $\mathbb{i}^\downarrow = 1$ with $\uparrow\downarrow$ functioning as the $\pm$ symbol does. After we show the realization of the Lorentz group and prove that the manifold of special relativity works as usual in the $\uparrow$ and $\downarrow$ permutations, we will make extensions to simultaneous, different combinations in different submanifolds of the MCM unit cell, and then we introduce non-trivial new behaviors by considering the cases for

$$\mathbb{i}^{\updownarrows} \quad \to \quad \mathbb{q}^{\uparrow\downarrow} \quad ,$$

where $\mathbb{q}^{\uparrow\downarrow}$ is used to assign real and quaternion phase rather than real and imaginary. By doing so, we will intend to construct the algebra of quantum mechanical spin operators from the phase convention on spacetime structure in the MCM unit cell. Namely, where he have defined physics in $\mathcal{H}$ as a sum of contributions from physics in $\Sigma^\pm$, we might examine the free particle Hamiltonian $H = xp$ in each of $\Sigma^\pm$ to pick up one quaternion phase or another so that the sum

$$H_\mathcal{H} = \frac{1}{2}\big(H_\mathcal{A} + H_\Omega\big) \quad ,$$

begins to look like the quaternion commutator

$$[\mathbf{i}, \mathbf{j}] = \varepsilon_{ijk}\mathbf{k} \quad .$$

Another thing we will look at is

$$H_{\mathcal{H}} = \Phi H_{\mathcal{A}} - \varphi H_{\Omega} \quad ,$$

so that there is a remainder.

Also, since we have determined that the MCM fifth dimension $\chi^4_\pm$ is not a local variable, meaning it information and correlations are not restricted by the speed of light, we will need to consider things like boosts with boost parameters greater than one. For this reason, we will not follow the usual convention to set $c = 1$.

KKKKKKKKKKKKKKKKKKKKKKKKKKKK
KKKKKKKKKKKKKKKKKKKKKKKKKKK
KKKKKKKKKKKKKKKKKKKKKKKKKK
maybe add $\widetilde{x}^\mu \in \mathcal{H}$ to title
Group theory, c.f.: [3], Appendix 2.
Review JAFFE, maybe add stuff from STEANE: [5, 22]

Example 4.5.1 In our way, we use the Euclidean metric $\delta_{\mu\nu}$:

$$\begin{aligned}
\big(x^\mu\big)^2 &= x^\mu x^\nu \delta_{\mu\nu} && (4.8)\\
&= x^\mu x_\mu = (ct, i\vec{x}) \cdot (ct, i\vec{x}) && (4.9)\\
&= \big(ct\big)^2 - \vec{x} \cdot \vec{x} && (4.10)\\
&= \big(ct\big)^2 - \sum x_i^2 && (4.11)
\end{aligned}$$

In Jaffe's way with points in $\mathbb{R}^4$ but a non-Euclidean metric, we have:

$$\begin{aligned}
\big(x^\mu\big)^2 &= x^\mu x^\nu \eta_{\mu\nu} && (4.12)\\
&= x^\mu x_\mu = (ct, \vec{x}) \cdot (ct, -\vec{x}) && (4.13)\\
&= \big(ct\big)^2 - \vec{x} \cdot \vec{x} && (4.14)\\
&= \big(ct\big)^2 - \sum x_i^2 && (4.15)
\end{aligned}$$

Overall, Jaffe uses matrix rather than tensor notations.

Remark 4.5.2 EXPLAIN WHY WE ARE GOING TO USE THESE OTHER SYMBOLS $\mathbb{i}^{\updownarrows}$ and $\beta^{\updownarrows}$

Definition 4.5.3 The symbol $\mathbb{i}^{\updownarrows}$ is such that

$$\mathbb{i}^{\uparrow} = i \quad , \quad \text{and} \qquad \mathbb{i}^{\downarrow} = 1 \quad .$$

The superscript $\uparrow\downarrow$ functions like the $\pm$ symbol, so

$$\mathbb{i}^{\updownarrows}\mathbb{i}^{\downuparrows} = i \quad .$$

The complex conjugation behavior is

$$\mathbb{i}^{\uparrow *} = -i \quad , \quad \mathbb{i}^{\downarrow *} = 1 \qquad \Longrightarrow \qquad |\mathbb{i}^{\uparrow\downarrow}|^2 = \big(\mathbb{i}^{\uparrow\downarrow}\big)^* \mathbb{i}^{\uparrow\downarrow} = 1 \quad .$$

The direct (non-conjugated) square of this symbol is

$$\big(\mathbb{i}^{\uparrow}\big)^2 = -1 \quad , \quad \text{and} \qquad \big(\mathbb{i}^{\downarrow}\big)^2 = 1 \quad .$$

This symbol will allow us to consider the cases for imaginary time and space parts of 4-vectors in unified expressions (unlike the previous article). Since we will want to consider sign and phase permutations separately, the more obvious notation $\mathbb{i}^{\pm}$ will not suffice. While it might be more convenient to use 10 and 01 than $\uparrow\downarrow$ and $\downarrow\uparrow$, we will want to link these cases to “spin up” and “spin down” later, and the notation is introduced with foresight. The reader is so advised.

Definition 4.5.4 To reverse the order of the arrows on the $\mathbb{i}^{\uparrow\downarrow}$ symbol, we will have to introduce one more symbol $\beta^{\uparrow\downarrow}$ such that

$$\mathbb{i}^{\uparrow\downarrow}\beta^{\downarrow\uparrow} = \mathbb{i}^{\downarrow\uparrow} \qquad \Longrightarrow \qquad \begin{cases} \beta^{\uparrow} = i \\ \beta^{\downarrow} = -i \quad . \end{cases}$$

Again, it will not suffice to use $\mp i$ because the $\pm$ symbol will be varied separately from the $\uparrow\downarrow$ symbol, and we have chosen the convention to be such that “beta down” is negative while “beta up” is positive. It follows that

$$\mathbb{i}^{\uparrow\downarrow}\beta^{\uparrow\downarrow} = -\mathbb{i}^{\downarrow\uparrow} \qquad \Longrightarrow \qquad \pm\beta^{\uparrow\downarrow} = \mp\beta^{\downarrow\uparrow} \quad .$$

MAKE LIST OF OTHER PROPERTIES
ALSO MAKE TABLE

Definition 4.5.5 The complex vector representation of points $x^\mu \in \mathcal{H}$ in Minkowski space is

$$\begin{aligned} x^0 &= \mathbb{i}^{\uparrow\downarrow} ct \\ x^1 &= \mathbb{i}^{\downarrow\uparrow} x \\ x^2 &= \mathbb{i}^{\downarrow\uparrow} y \\ x^3 &= \mathbb{i}^{\downarrow\uparrow} z \quad . \end{aligned}$$

where $(t, x, y, z) \in \mathbb{R}^4$. Since this does not allow four complex numbers in the general case, this is sometimes called $\mathbb{R}^{1,3}$ or $\mathbb{R}^{3,1}$.

Remark 4.5.6 It must be demonstrated that the complex 4-vectors transform appropriately under Lorentz transformations $x^\mu \to x'^\mu$:

$$x'^\mu = \Lambda^\mu_\nu x^\nu \quad .$$

We should also show that the Euclidean metric is invariant under Lorentz transformations:

$$\gamma'_{\mu\nu} = \Lambda^{\rho}_{\mu}\gamma_{\rho\sigma}\Lambda^{\sigma}_{\nu} = \gamma_{\mu\nu}$$

Theorem 4.5.7 *Given the Euclidean 4-metric $\gamma_{\mu\nu}$ and a complex vector $x^\mu \in \mathcal{M}$, the line element in Minkowski space is the usual one.*

Proof. Given

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu \quad ,$$

we have

$$\begin{aligned} ds^2 &= \gamma_{\mu\nu}dx^\mu dx^\nu = \pm\delta_{\mu\nu} \\ &= \pm\left[\left(\mathbb{i}^{\updownarrows}cdt\right)^2 + \left(\mathbb{i}^{\downuparrows}dx\right)^2 + \left(\mathbb{i}^{\downuparrows}dy\right)^2 + \left(\mathbb{i}^{\downuparrows}dz\right)^2\right] \quad . \end{aligned}$$

The cases of $\updownarrows$ are

$$\begin{aligned} ds^2_{\uparrow} &= \pm\left[\left(\mathbb{i}^{\uparrow}cdt\right)^2 + \left(\mathbb{i}^{\downarrow}dx\right)^2 + \left(\mathbb{i}^{\downarrow}dy\right)^2 + \left(\mathbb{i}^{\downarrow}dz\right)^2\right] \\ &= \pm\left[i^2c^2dt^2 + dx^2 + dy^2 + dz^2\right] \\ &= \pm\left[-c^2dt^2 + dx^2 + dy^2 + dz^2\right] \quad , \end{aligned}$$

and

$$\begin{aligned} ds^2_{\downarrow} &= \pm\left[\left(\mathbb{i}^{\downarrow}cdt\right)^2 + \left(\mathbb{i}^{\uparrow}dx\right)^2 + \left(\mathbb{i}^{\uparrow}dy\right)^2 + \left(\mathbb{i}^{\uparrow}dz\right)^2\right] \\ &= \pm\left[c^2dt^2 + i^2dx^2 + i^2dy^2 + i^2dz^2\right] \\ &= \pm\left[c^2dt^2 - dx^2 - dy^2 - dz^2\right] \quad . \end{aligned}$$

These expressions are identical they agree with the line element obtained from the Lorentzian metric $\eta_{\mu\nu} = \text{diag}(\pm1, \mp1, \mp1, \mp1)$ when points in Minkowski space are specified with $x^\mu \in \mathbb{R}^4$:

$$ds^2 = -c^2dt^2 + \sum_i \left(dx^i\right)^2 \quad , \qquad \text{or} \qquad ds^2 = c^2dt^2 - \sum_i \left(dx^i\right)^2 \quad .$$

Theorem 4.5.8 *The Minkowski square of $\widetilde{x}^\mu$ is the scalar product*

Proof.

$$\begin{aligned} \widetilde{x}^2_M &= \gamma_{\mu\nu}\widetilde{x}^\mu\widetilde{x}^\nu \\ &= \pm\left[\left(\mathbb{i}^{\updownarrows}\right)^2t^2 + \left(\mathbb{i}^{\downuparrows}\right)^2x^2 + \left(\mathbb{i}^{\downuparrows}\right)^2y^2 + \left(\mathbb{i}^{\downuparrows}\right)^2z^2\right] \\ &= \left(\mathbb{i}^{\downuparrows}\right)^2t^2 + \left(\mathbb{i}^{\updownarrows}\right)^2\left[x^2 + y^2 + z^2\right] \end{aligned}$$

Remark 4.5.9 We will need to find the set of Lorentz transformations that leaves this invariant. Then we will call this quantity "a Lorentz scalar." The positive, negative, and vanishing conditions that determine whether $\widetilde{x}^\mu$ is time-like, spacelike, or null, will depend on the chosen conventions. The important thing is that $(\mathbb{i}^{\upharpoonleft\downharpoonright})^2$ and $(\mathbb{i}^{\downharpoonleft\upharpoonright})^2$ are oppositely signed, so the required behavior is available in some form.

NOTE THAT THIS IS NOT THE INNER PRODUCT!!!

Remark 4.5.10 Since the complex 4-vector representation makes either the time or space part of a real-valued 4-vector imaginary, we should expect that $\widetilde{\Lambda} = \Lambda$ when the transformation is a pure rotation. For a boost, however, which is considered to be a rotation by an imaginary angle, or a rotation of time space and time axes, then we should find $\widetilde{\Lambda} \neq \Lambda$.

Definition 4.5.11 The elements of the realization of the Lorentz group in the complex-valued 4-vector representation are $\widetilde{\Lambda}^{\mu'}_{\mu}$:

$$\begin{aligned}\widetilde{\Lambda}^0_0 &= \Lambda^0_0 \\ \widetilde{\Lambda}^0_i &= \beta^{\upharpoonleft\downharpoonright}\Lambda^0_i \\ \widetilde{\Lambda}^i_0 &= \beta^{\downharpoonleft\upharpoonright}\Lambda^i_0 \\ \widetilde{\Lambda}^i_j &= \Lambda^i_j \quad .\end{aligned}$$

Note that these agree with the conventions given in 4.5. The freedom to choose real or imaginary space or time in $\widetilde{x}$ is reflected in the choice of sign for the imaginary parts of $\widetilde{\Lambda}$. Recall that the Lorentz transformation is not a tensor. It is a transformation and we don't need to worry about raising and lowering operations changing the signs of the entries. These are matrices with an upper index indicating a row, and a lower index indicating a column.

Remark 4.5.12 Every element of the Lorentz group has an inverse

$$x^{\mu'} = \Lambda^{\mu'}_{\mu} x^\mu \qquad \Longrightarrow \qquad x^\nu = \left(\Lambda^{-1}\right)^{\nu}_{\mu'} \Lambda^{\mu'}_{\mu} x^\mu = x^\mu \quad ,$$

and every inverse Lorentz transformation belongs to the Lorentz group as well. Now we will demonstrate that realization of the Lorentz group in the complex 4-vector representation with a few brief examples.

THE METRIC IS INVARIANT

BOOSTS

ROTATIONS

SHOW THAT THE METRIC IS ITS OWN INVERSE IN THE NEW CONVENTION

Λ IS UNITARY, CHECK FOR $\widetilde{\Lambda}$

Example 4.5.13 *Spatial rotations of real and complex 4-vectors.*

In the $x^\mu \in \mathbb{R}^4$ representation, a rotation by angle θ in the xy-plane is written

$$\Lambda^{\mu'}{}_{\mu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad .$$

It follows that

$$x^{\mu'} = \Lambda^{\mu'}{}_{\mu} x^\mu = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathrm{C} & \mathrm{S} & 0 \\ 0 & -\mathrm{S} & \mathrm{C} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t \\ x\mathrm{C} + y\mathrm{S} \\ -x\mathrm{S} + y\mathrm{C} \\ z \end{bmatrix} \quad .$$

Likewise for $\widetilde{x}^\mu$:

$$\widetilde{x}^{\mu'} = \widetilde{\Lambda}^{\mu'}{}_{\mu} \widetilde{x}^\mu = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathrm{C} & \mathrm{S} & 0 \\ 0 & -\mathrm{S} & \mathrm{C} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbb{i}^{\upharpoonleft\downharpoonright} t \\ \mathbb{i}^{\downharpoonleft\upharpoonright} x \\ \mathbb{i}^{\downharpoonleft\upharpoonright} y \\ \mathbb{i}^{\downharpoonleft\upharpoonright} z \end{bmatrix} = \begin{bmatrix} \mathbb{i}^{\upharpoonleft\downharpoonright} t \\ \mathbb{i}^{\downharpoonleft\upharpoonright} \big(x\mathrm{C} + y\mathrm{S} \big) \\ \mathbb{i}^{\downharpoonleft\upharpoonright} \big(-x\mathrm{S} + y\mathrm{C} \big) \\ \mathbb{i}^{\downharpoonleft\upharpoonright} z \end{bmatrix} \quad .$$

This example has confirmed the convention in Definition 4.5.11: NEED TO FIX

Example 4.5.14 *Boosts of real and complex 4-vectors.*

In the $x^\mu \in \mathbb{R}^4$ representation, a boost with rapidity ϕ in the z-direction is written

$$\Lambda^{\mu'}{}_{\mu} = \begin{bmatrix} \cosh\phi & 0 & 0 & -\sinh\phi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\phi & 0 & 0 & \cosh\phi \end{bmatrix} \quad .$$

It follows that

$$x^{\mu'} = \Lambda^{\mu'}{}_{\mu} x^\mu = \begin{bmatrix} \mathrm{Ch} & 0 & 0 & -\mathrm{Sh} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\mathrm{Sh} & 0 & 0 & \mathrm{Ch} \end{bmatrix} \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t\mathrm{Ch} - z\mathrm{Sh} \\ x \\ y \\ -t\mathrm{Sh} + z\mathrm{Ch} \end{bmatrix} \quad .$$

The usual special relativity formulae are obtained by defining the boost parameter, c.f.: Carroll [2] or Jaffe [22], to be such that

$$\phi = \tanh^{-1} v \quad ,$$

where v is the velocity of the boosted frame. Then

$$\cosh(\tanh^{-1} v) = \frac{1}{\sqrt{1-v^2}} \quad , \quad \text{and} \qquad \sinh(\tanh^{-1} v) = \frac{v}{\sqrt{1-v^2}} \quad ,$$

to obtain

$$t' = \frac{t}{\sqrt{1-v^2}} - \frac{zv}{\sqrt{1-v^2}} = \gamma\big(t - zv\big)$$
$$z' = -\frac{tv}{\sqrt{1-v^2}} + \frac{z}{\sqrt{1-v^2}} = \gamma\big(z - vt\big) \quad ,$$

where

$$\gamma = \frac{1}{\sqrt{1-v^2}} \quad .$$

Obviously, the correct form of the transformation of $\widetilde{x}^\mu$ under a Lorentz boost will be

$$\widetilde{x}^\mu \to \widetilde{x}^{\mu'} = \begin{bmatrix} \mathbb{i}^{\Updownarrow}\big(t\text{Ch} - z\text{Sh}\big) \\ \mathbb{i}^{\Downarrow\Uparrow}x \\ \mathbb{i}^{\Downarrow\Uparrow}y \\ \mathbb{i}^{\Downarrow\Uparrow}\big(-t\text{Sh} + z\text{Ch}\big) \end{bmatrix} \quad .$$

It is trivial to determine that the requisite form of $\widetilde{\Lambda}$ is

$$\widetilde{x}^{\mu'} = \widetilde{\Lambda}^{\mu'}{}_{\mu}\widetilde{x}^\mu = \begin{bmatrix} \text{Ch} & 0 & 0 & -\beta^{\Updownarrow}\text{Sh} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta^{\Downarrow\Uparrow}\text{Sh} & 0 & 0 & \text{Ch} \end{bmatrix} \begin{bmatrix} \mathbb{i}^{\Updownarrow}t \\ \mathbb{i}^{\Downarrow\Uparrow}x \\ \mathbb{i}^{\Downarrow\Uparrow}y \\ \mathbb{i}^{\Downarrow\Uparrow}z \end{bmatrix} = \begin{bmatrix} \mathbb{i}^{\Updownarrow}\big(t\text{Ch} - z\text{Sh}\big) \\ \mathbb{i}^{\Downarrow\Uparrow}x \\ \mathbb{i}^{\Downarrow\Uparrow}y \\ \mathbb{i}^{\Downarrow\Uparrow}\big(-t\text{Sh} + z\text{Ch}\big) \end{bmatrix} \quad .$$

This example has confirmed the convention in Definition 4.5.11: NEED TO FIX

Example 4.5.15 Check the inverse for complex realization. Given a boost in the x-direction

$$\Lambda^{\mu'}_{\mu} = \begin{bmatrix} \text{Ch} & \text{Sh} & 0 & 0 \\ \text{Sh} & \text{Ch} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad ,$$

the inverse is easily verified to be

$$\big(\Lambda^{-1}\big)^{\nu}_{\mu'} = \begin{bmatrix} \text{Ch} & -\text{Sh} & 0 & 0 \\ -\text{Sh} & \text{Ch} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad ,$$

Applying the conventions of Definition 4.5.11 to obtain this transformation and its inverse for $x^\mu \in \mathcal{H}$, we obtain

$$
\begin{aligned}
\left(\Lambda^{-1}\right)^{\nu}_{\mu'}\Lambda^{\mu'}_{\mu} &= \begin{bmatrix} \mathbb{i}^{\downuparrows}\text{Ch} & -\mathbb{i}^{\updownarrows}\text{Sh} & 0 & 0 \\ -\mathbb{i}^{\updownarrows}\text{Sh} & \mathbb{i}^{\downuparrows}\text{Ch} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbb{i}^{\downuparrows}\text{Ch} & \mathbb{i}^{\updownarrows}\text{Sh} & 0 & 0 \\ \mathbb{i}^{\updownarrows}\text{Sh} & \mathbb{i}^{\downuparrows}\text{Ch} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \left(\mathbb{i}^{\downuparrows}\right)^2\text{Ch}^2 - \left(\mathbb{i}^{\updownarrows}\right)^2\text{Sh}^2 & i\text{ChSh} - i\text{ShCh} & 0 & 0 \\ -i\text{ShCh} + i\text{ChSh} & -\left(\mathbb{i}^{\updownarrows}\right)^2\text{Sh}^2 + \left(\mathbb{i}^{\downuparrows}\right)^2\text{Ch} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \left(\mathbb{i}^{\downuparrows}\right)^2\text{Ch}^2 - \left(\mathbb{i}^{\updownarrows}\right)^2\text{Sh}^2 & 0 & 0 & 0 \\ 0 & -\left(\mathbb{i}^{\updownarrows}\right)^2\text{Sh}^2 + \left(\mathbb{i}^{\downuparrows}\right)^2\text{Ch} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned}
$$

Remark 4.5.16 In the real-valued $x^\mu \in \mathbb{R}^{\ngeq}$ representation of the Lorentz group, the realization of Λ is such that all of its entries are real, and the metric satisfies

$$\Lambda^T g \Lambda = g \quad .$$

Since the realization for $\widetilde{x}^\mu \in \mathcal{H}$ contains complex numbers, we should expect that the metric is invariant between a transform and its conjugate transpose.

Theorem 4.5.17 *The Euclidean metric $\gamma_{\mu\nu} = diag\{\pm1, \pm1, \pm1, \pm1\}$ satisfies*

$$\widetilde{\Lambda}^\dagger \gamma \widetilde{\Lambda} = \gamma \quad ,$$

where $\dagger$ denotes the conjugate transpose.

Proof. Given the known behavior of the Lorentzian metric $\eta_{\mu\nu} = \text{diag}\{\mp1, \pm1, \pm1, \pm1\}$

$$\eta_{\mu\nu} \to \eta_{\mu'\nu'} = \Lambda^{\mu}_{\mu'}\eta_{\mu\nu}\Lambda^{\nu}_{\nu'}$$

we may write in matrix index notation

$$\eta = \Lambda^T \eta \Lambda = \Lambda^T_{\rho\mu}\eta_{\mu\nu}\Lambda_{\nu\sigma} = \Lambda_{\mu\rho}\eta_{\mu\nu}\Lambda_{\nu\sigma} \quad .$$

The metric contains off-diagonal zeros so the constraints on the entries in Λ must come from the case in which $\mu = \nu$:

$$\eta_{\rho\sigma} = \mp\Lambda_{0\rho}\Lambda_{0\sigma} \pm \Lambda_{1\rho}\Lambda_{1\sigma} \pm \Lambda_{2\rho}\Lambda_{2\sigma} \pm \Lambda_{3\rho}\Lambda_{3\sigma}$$

Choosing the 00 component of $\eta_{\rho\sigma}$, we find

$$\mp1 = \mp\Lambda_{00}\Lambda_{00} \pm \Lambda_{10}\Lambda_{10} \pm \Lambda_{20}\Lambda_{20} \pm \Lambda_{30}\Lambda_{30} \quad .$$

This does not agree with the orthogonality condition given by Roman in Equation (4.5):

$$\widetilde{\Lambda}_{\mu\nu}\widetilde{\Lambda}_{\mu\rho} = \widetilde{\Lambda}_{\nu\mu}\widetilde{\Lambda}_{\rho\mu} = \delta_{\nu\rho} \quad .$$

So, we can see that the $\widetilde{\Lambda}$ realization of the Lorentz group has constraints not present for Λ. Using

$$\widetilde{\Lambda}^0_0 = \Lambda^0_0 \quad , \quad \widetilde{\Lambda}^0_i = \beta^{\Updownarrow}\Lambda^0_i \quad , \quad \widetilde{\Lambda}^i_0 = \beta^{\Downarrow\Uparrow}\Lambda^i_0 \quad , \quad \text{and} \qquad \widetilde{\Lambda}^i_j = \Lambda^i_j \quad ,$$

(as in XXXXXX) we may directly compute the theorem as

$$\gamma = \widetilde{\Lambda}^\dagger\gamma\widetilde{\Lambda}$$

$$\gamma_{\mu\nu} = \begin{bmatrix} \Lambda_{00} & \left(\beta^{\Downarrow\Uparrow}\right)^*\Lambda_{10} & \left(\beta^{\Downarrow\Uparrow}\right)^*\Lambda_{20} & \left(\beta^{\Downarrow\Uparrow}\right)^*\Lambda_{30} \\ \left(\beta^{\Updownarrow}\right)^*\Lambda_{01} & \Lambda_{11} & \Lambda_{21} & \Lambda_{31} \\ \left(\beta^{\Updownarrow}\right)^*\Lambda_{02} & \Lambda_{12} & \Lambda_{22} & \Lambda_{32} \\ \left(\beta^{\Updownarrow}\right)^*\Lambda_{03} & \Lambda_{13} & \Lambda_{23} & \Lambda_{33} \end{bmatrix} \times \ldots$$

$$\ldots \times \begin{bmatrix} \pm\Lambda_{00} & \pm\beta^{\Updownarrow}\Lambda_{01} & \pm\beta^{\Updownarrow}\Lambda_{02} & \pm\beta^{\Updownarrow}\Lambda_{03} \\ \pm\beta^{\Downarrow\Uparrow}\Lambda_{10} & \pm\Lambda_{11} & \pm\Lambda_{12} & \pm\Lambda_{13} \\ \pm\beta^{\Downarrow\Uparrow}\Lambda_{20} & \pm\Lambda_{21} & \pm\Lambda_{22} & \pm\Lambda_{23} \\ \pm\beta^{\Downarrow\Uparrow}\Lambda_{30} & \pm\Lambda_{31} & \pm\Lambda_{32} & \pm\Lambda_{33} \end{bmatrix} \quad .$$

Letting suffice again to consider only the 00 component, we obtain

$$\pm 1 = \pm\left[\Lambda_{00}\Lambda_{00} + \left(\beta^{\Downarrow\Uparrow}\right)^*\left(\beta^{\Downarrow\Uparrow}\right)\Lambda_{i0}\Lambda_{i0}\right] \quad .$$

Using $(\beta^{\Downarrow\Uparrow})^*(\beta^{\Downarrow\Uparrow}) = 1$, we obtain

$$\pm 1 = \pm\left[\Lambda_{00}\Lambda_{00} + \Lambda_{i0}\Lambda_{i0}\right] \quad .$$

Using the $\widetilde{\Lambda}_{\mu\nu}\widetilde{\Lambda}_{\mu\rho} = \delta_{\nu\rho}$ condition stated by Roman, we have

$$\pm 1 = \pm\widetilde{\Lambda}_{\mu 0}\widetilde{\Lambda}_{\mu 0} = \pm\delta_{00} \quad ,$$

which is correct. Since Lorentz transformation are *defined* to satisfy $\widetilde{\Lambda}^T g\widetilde{\Lambda} = g$, we have independently arrived at the condition cited by Roman, and theorem is proven. ∎

Theorem 4.5.18 *The inner product of $x^\mu \in \mathcal{H}$ with itself is not a Lorentz scalar under rotations.*

Proof. The inner product for $x^\mu \in \mathcal{H}$ is defined as

$$\left(x^\mu\right)^2 \equiv |x^\mu|^2 = \left(x^\mu\right)^* x_\mu = \left(x^\mu\right)^*\eta_{\mu\nu}x^\nu$$

If the 4-vector is complex and the metric is Euclidean, and if we use the same rotation matrix as we have used for real 4-vectors, then

$$x^{\mu'} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathrm{C} & \mathrm{S} & 0 \\ 0 & -\mathrm{S} & \mathrm{C} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbb{i}^{\uparrow\downarrow} t \\ \mathbb{i}^{\downarrow\uparrow} x \\ \mathbb{i}^{\downarrow\uparrow} y \\ \mathbb{i}^{\downarrow\uparrow} z \end{bmatrix} = \begin{bmatrix} \mathbb{i}^{\uparrow\downarrow} t \\ \mathbb{i}^{\downarrow\uparrow}\big(x\mathrm{C} + y\mathrm{S}\big) \\ \mathbb{i}^{\downarrow\uparrow}\big(y\mathrm{C} - x\mathrm{S}\big) \\ \mathbb{i}^{\downarrow\uparrow} z \end{bmatrix} \quad . \tag{4.16}$$

The inner product of the rotated complex 4-vector with itself is:

$$\begin{aligned}
\big(x^{\mu'}\big)^2 &= \big(x^{\mu'}\big)^* \gamma_{\mu\nu} x^{\nu'} \\
&= \begin{bmatrix} \mathbb{i}^{\uparrow\downarrow} t \\ \mathbb{i}^{\downarrow\uparrow}\big(x\mathrm{C} + y\mathrm{S}\big) \\ \mathbb{i}^{\downarrow\uparrow}\big(y\mathrm{C} - x\mathrm{S}\big) \\ \mathbb{i}^{\downarrow\uparrow} z \end{bmatrix}^\dagger \begin{bmatrix} \pm1 & 0 & 0 & 0 \\ 0 & \pm1 & 0 & 0 \\ 0 & 0 & \pm1 & 0 \\ 0 & 0 & 0 & \pm1 \end{bmatrix} \begin{bmatrix} \mathbb{i}^{\uparrow\downarrow} t \\ \mathbb{i}^{\downarrow\uparrow}\big(x\mathrm{C} + y\mathrm{S}\big) \\ \mathbb{i}^{\downarrow\uparrow}\big(y\mathrm{C} - x\mathrm{S}\big) \\ \mathbb{i}^{\downarrow\uparrow} z \end{bmatrix} \\
&= \begin{bmatrix} \big(\mathbb{i}^{\uparrow\downarrow}\big)^* t \\ \big(\mathbb{i}^{\downarrow\uparrow}\big)^*\big(x\mathrm{C} + y\mathrm{S}\big) \\ \big(\mathbb{i}^{\downarrow\uparrow}\big)^*\big(y\mathrm{C} - x\mathrm{S}\big) \\ \big(\mathbb{i}^{\downarrow\uparrow}\big)^* z \end{bmatrix}^T \begin{bmatrix} \pm\mathbb{i}^{\uparrow\downarrow} t \\ \pm\mathbb{i}^{\downarrow\uparrow}\big(x\mathrm{C} + y\mathrm{S}\big) \\ \pm\mathbb{i}^{\downarrow\uparrow}\big(y\mathrm{C} - x\mathrm{S}\big) \\ \pm\mathbb{i}^{\downarrow\uparrow} z \end{bmatrix} \\
&= \pm|\mathbb{i}^{\uparrow\downarrow} t|^2 \pm |\mathbb{i}^{\downarrow\uparrow}|^2 \Big[\big(x\mathrm{C} + y\mathrm{S}\big)^2 + \big(y\mathrm{C} - x\mathrm{S}\big)^2 + z^2\Big] \\
&= \pm|\mathbb{i}^{\uparrow\downarrow} t|^2 \pm |\mathbb{i}^{\downarrow\uparrow}|^2 \big[\big(x^2\mathrm{C}^2 + 2xy\mathrm{CS} + y^2\mathrm{S}^2\big) + \big(y^2\mathrm{C}^2 - 2xy\mathrm{CS} + x^2\mathrm{S}^2\big) + z^2\big] \\
&= \pm|\mathbb{i}^{\uparrow\downarrow} t|^2 \pm |\mathbb{i}^{\downarrow\uparrow}|^2 \big[x^2\big(\mathrm{C}^2 + \mathrm{S}^2\big) + y^2\big(\mathrm{S}^2 + \mathrm{C}^2\big) + z^2\big] \\
&= \pm|\mathbb{i}^{\uparrow\downarrow} t|^2 \pm |\mathbb{i}^{\downarrow\uparrow}|^2 \big(x^2 + y^2 + z^2\big)
\end{aligned}$$

Since $|\mathbb{i}^{\uparrow\downarrow}|^2 = 1$, the inner product has not been preserved. The spacelike and timelike parts have the same sign, and there is no possibility for a null 4-vector with vanishing Minkowski length (SQUARED?).

Theorem 4.5.19 *The contraction of $x^\mu, x^\nu \in \mathcal{H}$ with the metric is a Lorentz scalar under rotations.*

Proof. Using $x^{\mu'}$ as in Equation (4.16), the contraction with the metric is

$$\begin{aligned}
\big(x^{\mu'}\big)^2 &= x^{\mu'} \gamma_{\mu\nu} x^{\nu'} \\
&= \begin{bmatrix} \mathbb{i}^{\uparrow\downarrow} t \\ \mathbb{i}^{\downarrow\uparrow}\big(x\mathrm{C} + y\mathrm{S}\big) \\ \mathbb{i}^{\downarrow\uparrow}\big(y\mathrm{C} - x\mathrm{S}\big) \\ \mathbb{i}^{\downarrow\uparrow} z \end{bmatrix}^T \begin{bmatrix} \pm1 & 0 & 0 & 0 \\ 0 & \pm1 & 0 & 0 \\ 0 & 0 & \pm1 & 0 \\ 0 & 0 & 0 & \pm1 \end{bmatrix} \begin{bmatrix} \mathbb{i}^{\uparrow\downarrow} t \\ \mathbb{i}^{\downarrow\uparrow}\big(x\mathrm{C} + y\mathrm{S}\big) \\ \mathbb{i}^{\downarrow\uparrow}\big(y\mathrm{C} - x\mathrm{S}\big) \\ \mathbb{i}^{\downarrow\uparrow} z \end{bmatrix}
\end{aligned}$$

$$
\begin{aligned}
&= \begin{bmatrix} \mathbb{i}^{\updownarrows}t \\ \mathbb{i}^{\downuparrows}\big(x\mathrm{C}+y\mathrm{S}\big) \\ \mathbb{i}^{\downuparrows}\big(y\mathrm{C}-x\mathrm{S}\big) \\ \mathbb{i}^{\downuparrows}z \end{bmatrix}^T \begin{bmatrix} \pm\mathbb{i}^{\updownarrows}t \\ \pm\mathbb{i}^{\downuparrows}\big(x\mathrm{C}+y\mathrm{S}\big) \\ \pm\mathbb{i}^{\downuparrows}\big(y\mathrm{C}-x\mathrm{S}\big) \\ \pm\mathbb{i}^{\downuparrows}z \end{bmatrix} \\
&= \pm\big(\mathbb{i}^{\updownarrows}t\big)^2 \pm \big(\mathbb{i}^{\downuparrows}\big)^2 \Big[\big(x\mathrm{C}+y\mathrm{S}\big)^2 + \big(y\mathrm{C}-x\mathrm{S}\big)^2 + z^2\Big] \\
&= \pm\big(\mathbb{i}^{\updownarrows}t\big)^2 \pm \big(\mathbb{i}^{\downuparrows}\big)^2 \Big[\big(x^2\mathrm{C}^2 + 2xy\mathrm{CS} + y^2\mathrm{S}^2\big) + \big(y^2\mathrm{C}^2 - 2xy\mathrm{CS} + x^2\mathrm{S}^2\big) + z^2\Big] \\
&= \pm\big(\mathbb{i}^{\updownarrows}t\big)^2 \pm \big(\mathbb{i}^{\downuparrows}\big)^2 \Big[x^2\big(\mathrm{C}^2+\mathrm{S}^2\big) + y^2\big(\mathrm{S}^2+\mathrm{C}^2\big) + z^2\Big] \\
&= \pm\big(\mathbb{i}^{\updownarrows}t\big)^2 \pm \big(\mathbb{i}^{\downuparrows}\big)^2\big(x^2+y^2+z^2\big)
\end{aligned}
$$

Since $(\mathbb{i}^{\updownarrows})^2$ and $(\mathbb{i}^{\downuparrows})^2$ are oppositely signed, the contraction with the metric is a Lorentz scalar. The scalar output of the contraction operation agrees with the result obtain in Theorem 4.2.17.

Example 4.5.20 If we use the Lorentz boost defined for $x^\mu \in \mathbb{R}^4$, the result is not a Lorentz scalar. Namely, the representation of the Lorentz group is different for complex-valued 4-vectors.

Given

$$
x^{\mu'} = \begin{bmatrix} \mathrm{Ch} & 0 & 0 & -\mathrm{Sh} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\mathrm{Sh} & 0 & 0 & \mathrm{Ch} \end{bmatrix} \begin{bmatrix} \mathbb{i}^{\updownarrows}t \\ \mathbb{i}^{\downuparrows}x \\ \mathbb{i}^{\downuparrows}y \\ \mathbb{i}^{\downuparrows}z \end{bmatrix} = \begin{bmatrix} \mathbb{i}^{\updownarrows}t\mathrm{Ch} - \mathbb{i}^{\downuparrows}z\mathrm{Sh} \\ \mathbb{i}^{\downuparrows}x \\ \mathbb{i}^{\downuparrows}y \\ -\mathbb{i}^{\updownarrows}t\mathrm{Sh} + \mathbb{i}^{\downuparrows}z\mathrm{Ch} \end{bmatrix} \quad ,
$$

the contraction with the metric is

$$
\begin{aligned}
\big(x^{\mu'}\big)^2 &= x^{\mu'}\gamma_{\mu\nu}x^{\nu'} \\
&= \begin{bmatrix} \mathbb{i}^{\updownarrows}t\mathrm{Ch} - \mathbb{i}^{\downuparrows}z\mathrm{Sh} \\ \mathbb{i}^{\downuparrows}x \\ \mathbb{i}^{\downuparrows}y \\ -\mathbb{i}^{\updownarrows}t\mathrm{Sh} + \mathbb{i}^{\downuparrows}z\mathrm{Ch} \end{bmatrix}^T \begin{bmatrix} \pm1 & 0 & 0 & 0 \\ 0 & \pm1 & 0 & 0 \\ 0 & 0 & \pm1 & 0 \\ 0 & 0 & 0 & \pm1 \end{bmatrix} \begin{bmatrix} \mathbb{i}^{\updownarrows}t\mathrm{Ch} - \mathbb{i}^{\downuparrows}z\mathrm{Sh} \\ \mathbb{i}^{\downuparrows}x \\ \mathbb{i}^{\downuparrows}y \\ -\mathbb{i}^{\updownarrows}t\mathrm{Sh} + \mathbb{i}^{\downuparrows}z\mathrm{Ch} \end{bmatrix} \\
&= \begin{bmatrix} \mathbb{i}^{\updownarrows}t\mathrm{Ch} - \mathbb{i}^{\downuparrows}z\mathrm{Sh} \\ \mathbb{i}^{\downuparrows}x \\ \mathbb{i}^{\downuparrows}y \\ -\mathbb{i}^{\updownarrows}t\mathrm{Sh} + \mathbb{i}^{\downuparrows}z\mathrm{Ch} \end{bmatrix}^T \begin{bmatrix} \pm\mathbb{i}^{\updownarrows}t\mathrm{Ch} \mp \mathbb{i}^{\downuparrows}z\mathrm{Sh} \\ \pm\mathbb{i}^{\downuparrows}x \\ \pm\mathbb{i}^{\downuparrows}y \\ \mp\mathbb{i}^{\updownarrows}t\mathrm{Sh} \pm \mathbb{i}^{\downuparrows}z\mathrm{Ch} \end{bmatrix} \\
&= \big(\mathbb{i}^{\updownarrows}t\mathrm{Ch} - \mathbb{i}^{\downuparrows}z\mathrm{Sh}\big)\big(\pm\mathbb{i}^{\updownarrows}t\mathrm{Ch} \mp \mathbb{i}^{\downuparrows}z\mathrm{Sh}\big) \pm \big(\mathbb{i}^{\downuparrows}\big)^2x^2 \pm \big(\mathbb{i}^{\downuparrows}\big)^2y^2 + \dots \\
&\qquad\qquad \dots + \big(-\mathbb{i}^{\updownarrows}t\mathrm{Sh} + \mathbb{i}^{\downuparrows}z\mathrm{Ch}\big)\big(\mp\mathbb{i}^{\updownarrows}t\mathrm{Sh} \pm \mathbb{i}^{\downuparrows}z\mathrm{Ch}\big)
\end{aligned}
$$

$$
\begin{aligned}
&= \pm\big(\mathbb{i}^{\uparrow\downarrow}\big)^2 t^2 \mathrm{Ch}^2 \mp 2itz\mathrm{ChSh} \pm \big(\mathbb{i}^{\downarrow\uparrow}\big)^2 z^2 \mathrm{Sh}^2 \pm \big(\mathbb{i}^{\downarrow\uparrow}\big)^2 \big(x^2 + y^2\big) + \dots \\
&\qquad\qquad \dots + \Big[\pm\big(\mathbb{i}^{\uparrow\downarrow}\big)^2 t^2 \mathrm{Sh}^2 \mp 2itz\mathrm{ChSh} \pm \big(\mathbb{i}^{\downarrow\uparrow}\big)^2 z^2 \mathrm{Ch}^2\Big] \\
&= \pm\big(\mathbb{i}^{\uparrow\downarrow}\big)^2 t^2 \mathrm{Ch}^2 \pm \big(\mathbb{i}^{\downarrow\uparrow}\big)^2 z^2 \mathrm{Sh}^2 \pm \big(\mathbb{i}^{\downarrow\uparrow}\big)^2 \big(x^2 + y^2\big) + \dots \\
&\qquad\qquad \dots + \Big[\pm\big(\mathbb{i}^{\uparrow\downarrow}\big)^2 t^2 \mathrm{Sh}^2 \pm \big(\mathbb{i}^{\downarrow\uparrow}\big)^2 z^2 \mathrm{Ch}^2\Big]
\end{aligned}
$$

Since $\mathbb{i}^{\uparrow\downarrow}\mathbb{i}^{\downarrow\uparrow} = i$, we obtain an i in the cross terms, as should be expected when mixing the time and space parts of $x^\mu \in \mathcal{H}$. The cross terms have the same sign and cannot cancel, so we have proven that we need a different representation of the Lorentz group for $x^\mu \in \mathcal{H}$ than we use when $x^\mu \in \mathbb{R}^4$.

Theorem 4.5.21 *A complex 4-vector $x^\mu \in \mathcal{H}$ is a Lorentz scalar under boosts.*

§4.6 The Bi-Spinor Representation

General properties of the Lorentz covariance of Dirac bi-spinors may be found in, for example, Weinberg [23] and Roman [3]. Weinberg uses the usual representation corresponding to $V^\mu \in \mathbb{R}^4$ and Roman uses the one corresponding to $\mathbb{R}^{1,3}$.

URL:

ROMAN AND WEINBERG HAVE SECTIONS ON LORENTZ XFORMS FOR DIRAC

ZEE TOO MAYBE???

§5 The Lorentz Group

REFS p693 ROMAN, Lubanski?

====================

"The fundamental group of a manifold is countable" [15]

QUOTE Carroll O(3,1): p13, eq (1.30)

Remark 5.0.1 "which is connected to the fact that the group of motions in hyperbolic space, the Mobius group or projective special linear group, and the Laguerre group are isomorphic to the Lorentz group."

URL:

Remark 5.0.2 "The Lorentz group is a subgroup of the Poincaré group—the group of all isometries of Minkowski spacetime. Lorentz transformations are, precisely, isometries that leave the origin fixed. Thus, the Lorentz group is the isotropy subgroup with respect to the origin of the isometry group of Minkowski spacetime. For this reason, the Lorentz group is sometimes called the homogeneous Lorentz group while the Poincaré group is sometimes called the inhomogeneous Lorentz group."

URL:

" (The vector space equipped with this quadratic form is sometimes written $\mathbb{R}^{1,3}$"

URL:

"The restricted Lorentz group consists of those Lorentz transformations that preserve both the orientation of space and the direction of time. Its fundamental group has order 2, and its universal cover, the indefinite spin group Spin(1,3), is isomorphic to both the special linear group SL(2,C) and to the symplectic group Sp(2,C). These isomorphisms allow the Lorentz group to act on a large number of mathematical structures important to physics, most notably spinors. Thus, in relativistic quantum mechanics and in quantum field theory, it is very common to call SL(2,C) the Lorentz group, with the understanding that $SO^+(1,3)$ is a specific representation (the vector representation) of it."

Remark 5.0.3 I need to make some distinctions about the representations/generators of the group when the objects it acts on are different kinds of vectors.

======================

the endomorphisms of an abelian group form a ring (the endomorphism ring)

URL:

=======================

§5.1 Groups

A multiplicative group G is a set of elements with three properties:

1. Associative multiplication:

$$g_1, g_2 \in \mathbb{G} \quad \Longrightarrow \quad g_1 g_2 \in G \quad , \quad \text{where} \quad \big(g_1 g_2\big) g_3 = g_1 \big(g_2 g_3\big)$$

2. G has an identity $\mathbb{1} \in G$:

$$g \in G \quad \Longrightarrow \quad g\mathbb{1} = \mathbb{1}g = g$$

3. Every $g \in G$ has an inverse $g^{-1} \in G$:

$$gg^{-1} = g^{-1}g = \mathbb{1}$$

Definition 5.1.1 An **orthogonal group** is

Definition 5.1.2 A **unitary group** is

Definition 5.1.3 A **spin group** is

Definition 5.1.4 A **symplectic group** is

Definition 5.1.5 A **linear group** is

Definition 5.1.6 A **Lie group** is

Definition 5.1.7 The **dimension** of a group, or group dimension, is the number of real parameters needed to specify on of its elements.

Definition 5.1.8 Groups G_1, G_2 are **isomorphic groups** if STEANE [**?**]4, col2

Definition 5.1.9 Groups G_1, G_2 are **homomorphic groups** if STEANE [**?**]4, col2

Theorem 5.1.10 *SU(2) and SO(3) are homomorphic but not isomorphic.*

KKKKKKKKKKKKKKKKKKK

§5.2 The Lorentz Group

Article 5.2.1 Steane [1] on p4 relates the rotation group to SU(2) and says its very important for physics.

3D rotation group is different than, but isomorphic to, SO(3). We might write the rotations as $\mathcal{R}(\hat{\theta}, \hat{n})$, in which case they form a group without every mentioning matrices. However, the group of all such $\mathcal{R}(\hat{\theta}, \hat{n})$ is in one-to-one correspondence with the group of 3×3 orthogonal matrices, and they are said to be isomorphic.

"The group manifold of O(3, 1) can be thought of as the 6-dimensional surface in 16-dimensional matrix space (the space of 4 × 4, real matrices) on which Eq. (8) is satisfied. "

URL:

==

Theorem 5.2.2 *$\mathcal{L}$ is a multiplicative group with properties:*

$$\Lambda \in \mathcal{L} \implies \det \lambda = \pm 1 \quad , \quad \textit{and} \qquad \forall \Lambda \in \mathcal{L} \quad \exists \Lambda^T \in \mathcal{L}$$

Proof. Given $\Lambda^T g \Lambda = g$, let $\Lambda = \Lambda_1 \Lambda_2$ with $\Lambda_1 \Lambda_2 \in \mathcal{L}$. Using $(AB)^T = B^T A^T$, we find:

$$\big(\Lambda_1 \Lambda_2\big)^T g \big(\Lambda_1 \Lambda_2\big) = \Lambda_2^T \big(\Lambda_1^T g \Lambda_1\big) \Lambda_2 = \Lambda_2^T g \Lambda_2 = g$$

Therefore, the set is closed under multiplication because $\Lambda \in \mathcal{L}$. Using $\det AB = \det A \det B$ and $\det A = \det A^T$, we find:

$$\det g = -1 \qquad \implies \qquad \det \Lambda^T \det g \det \Lambda = -1 \qquad \implies \qquad \det \Lambda = \pm 1$$

The inverse exists whenever the determinant is non-zero, so there exists a Λ^{-1}. We use the inverse to write:

$$\begin{aligned} g &= \mathbb{1} g \mathbb{1} \\ &= \big(\Lambda^T\big)^{-1} \big(\Lambda^T g \Lambda\big) \Lambda^{-1} \\ &= \big(\Lambda^T\big)^{-1} g \Lambda^{-1} \\ &= \big(\Lambda^{-1}\big)^T g \Lambda^{-1} \end{aligned}$$

where we have used $(A^{-1})^T = (A^T)^{-1}$. Since the final line is equal to g, we have shown $\Lambda^{-1} \in \mathcal{L}$

$\mathcal{L}_x \subset \mathcal{L}$ is a connected component if one can find a continuous trajectory of matrices:

$$\forall \Lambda_1, \Lambda_2 \in \mathcal{L}_x \quad \exists 0 \leq s \leq 1, \Lambda(s) \in \mathcal{L}_x \qquad \text{s.t.} \qquad \Lambda(0) = \Lambda_1, \quad \Lambda(1) = \Lambda_2$$

============================
- $\mathcal{L}_-^\uparrow = P\mathcal{L}_+^\uparrow \qquad \mathcal{L}_-^\downarrow = T\mathcal{L}_+^\uparrow \qquad \mathcal{L}_+^\downarrow = PT\mathcal{L}_+^\uparrow$

- $\mathcal{L}^\uparrow = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\uparrow$ is a group, and so is $\mathcal{L}^\downarrow = \mathcal{L}_+^\downarrow \cup \mathcal{L}_-^\downarrow$

- Another group is $\mathcal{L}_0 = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\downarrow$ given by $(\Lambda_{00} \det \Lambda) \geq 1$

- $\mathcal{L}_-^\uparrow$ and $\mathcal{L}_+^\downarrow$ are not closed under multiplication, so they are not groups!

=========================

§5.3 Lie Groups and Lie Algebras

Article 5.3.1 "For example a Lie group is defined as a certain differentiable manifold, but what does this mean"
URL:

Article 5.3.2 Actions of Lie groups and Lie algebras on manifolds
URL:

URL:
==================
VERY INTERESTING: URL:
mentions mercury perihelion
=========================

Because it is a Lie group, the Lorentz group is a group and also has a topological description as a smooth manifold. As a manifold, it has four connected components. Intuitively, this means that it consists of four topologically separated pieces.

The four connected components can be categorized by two transformation properties its elements have:

Some elements are reversed under time-inverting Lorentz transformations, for example, a future-pointing timelike vector would be inverted to a past-pointing vector Some elements have orientation reversed by improper Lorentz transformations, for example, certain vierbein (tetrads)
URL:
==========================

§5.4 Clifford Algebras

URL:

§5.5 The Poincaré Group

This group is called the inhomogeneous Lorentz group. It includes all of the usual Lorentz translations, but it also includes translations that obviously won't preserve the Lorentz norm. NICE EXPLANATION BY PENROSE with "vec" function.

THEY SAY SPIN EMERGES NATURALLY IN THE STUDY OF THIS GROUP!!!

§6 MCM Particles

§6.1 The Topology of MCM Spacetime

LETS FORMALIZE THE TOPOLOGY!!!!!!!!!

§6.2 Submanifolds in the MCM Unit Cell

COPY SUBMANIFOLDS FROM CARROLL

What we have called the manifold of special relativity so far is called "one quantum of spacetime" in the MCM, and the spectra of such quanta are said to be the standard model particles [10, 24]. Via the introduction of a new application for holographic duality, the 4D manifold of physical observables in spacetime, called $\mathcal{H}$, is embedded between two 5D spaces $\Sigma^{\pm}$, as in Figure 8. $\mathcal{A}$ and Ω are also 4-spaces with elements that should form a representation of the Lorentz group. Most excitingly, the MCM new fifth dimension $\chi^4_{\pm}$ (horizontal on the page) is left-handed or right-handed with respect to the coordinates $\mathcal{H}$. We will want to form 4-vectors in $\Sigma^{\pm}$ by taking the 3D space part together with either the chronological or chirological time. The 2×2 matrix representation of 4-vectors will be very natural for this because given a 5-vector

$$\chi^A = (\chi^0, \chi^1, \chi^2, \chi^3, \chi^4) \quad ,$$

there exists two 2×2 matrices

$$\widehat{\chi}_0 = \begin{bmatrix} \chi^0+\chi^3 & \chi^1 - i\chi^2 \\ \chi^1 + i\chi^2 & \chi^0 - \chi^3 \end{bmatrix} \quad , \quad \text{and} \qquad \widehat{\chi}_4 = \begin{bmatrix} \chi^4+\chi^3 & \chi^1 - i\chi^2 \\ \chi^1 + i\chi^2 & \chi^4 - \chi^3 \end{bmatrix} \quad .$$

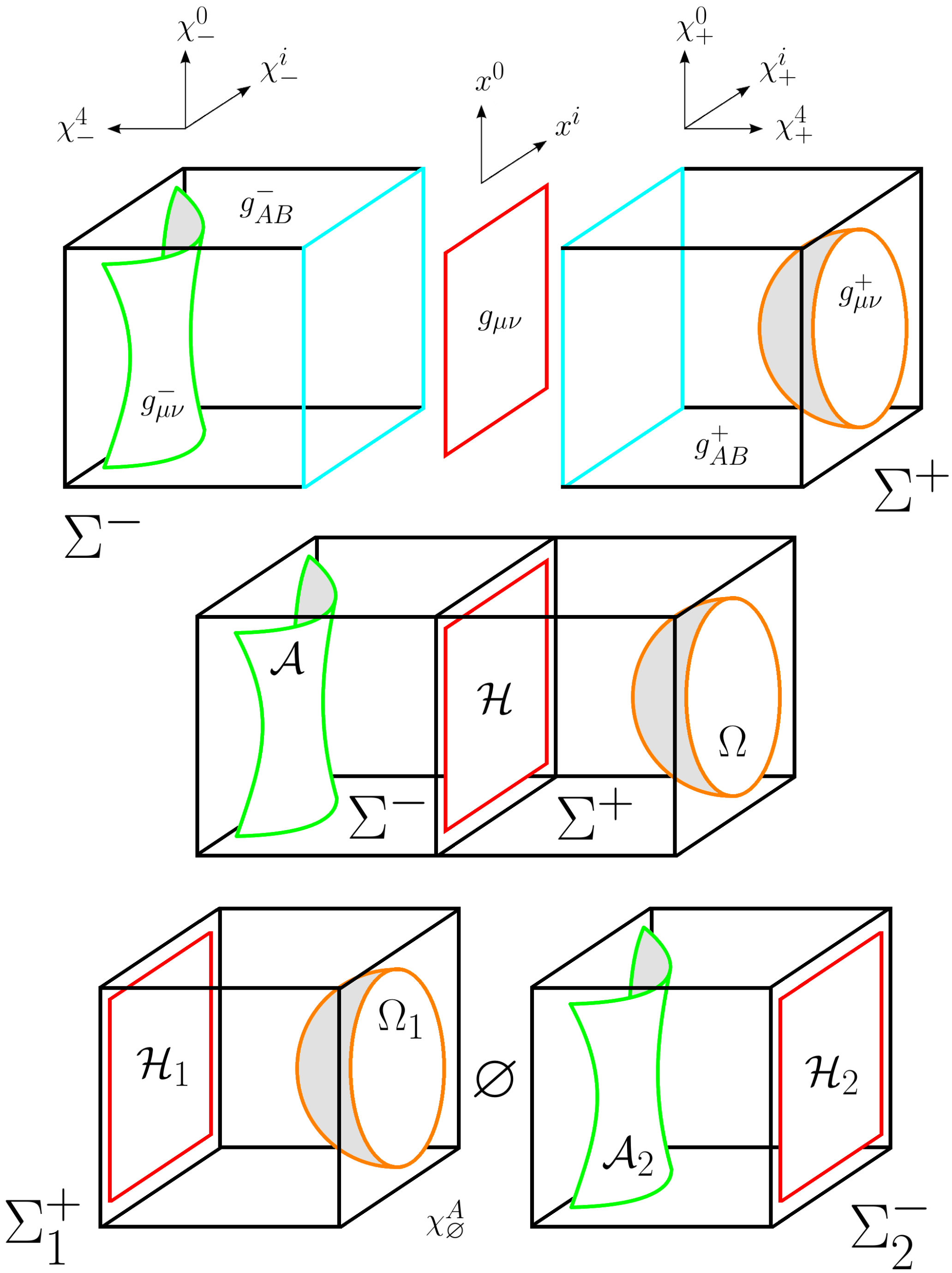
Both of these are rank-2 spinors (CHECK)!!! By developing the different properties in of such matrices in the $\widehat{\chi}_{\pm}$ variants, we will want to obtain the left- and right-handed Weyl spinors ψ^{α}_R and $\bar{\chi}^{\dot{\alpha}}_L$ used to construct Dirac bispinors for the observable quantum physics of relativistic charged particles in $\mathcal{H}$.

Carroll gives the example of something that is not a manifold as a line terminating on a plane, but $\chi^4_{\pm}$ are positive- and negative-definite, so we may treat them as manifolds.

In some way, we will want the property of spinors that rotation by 2π results in sign inversion to reflect the sign of the fifth position in the metric signature.

==============================

Replace the imaginary number with a quaternion. There's no way to tell the difference when they aren't in contact. However, they come in contact



χ^0_-
χ^i_-
χ^4_-
x^0
x^i
χ^0_+
χ^i_+
χ^4_+
g^-_{AB}
$g_{\mu\nu}$
$g^+_{\mu\nu}$
$g^-_{\mu\nu}$
g^+_{AB}
Σ^+
Σ^-
$\mathcal{A}$
$\mathcal{H}$
Ω
Σ^-
Σ^+
Ω_1
$\mathcal{H}_1$
$\varnothing$
$\mathcal{H}_2$
$\mathcal{A}_2$
Σ^+_1
$\chi^A_\varnothing$
Σ^-_2

Figure 8: XXX

when we turn the crank on going back and forth between 5-vectors and 2×2 matrices.

==============================

We have set up the MCM unit cell to be such that the fundamental SM particles are taken as quanta of spacetime: a 3-space joined to a chronological or chirological time part. We have raised some questions about the metric discontinuity between $\Sigma^\pm$ where the $A^\mu = 0$ metric $g^\pm_{AB}$ has signature $\{+---\pm\}$ or $\{-+++\pm\}$ when $\chi^A \in \mathbb{R}^5$. If evolution along $\chi^4_\pm$ is timelike in $\Sigma^\pm$, then it is spacelike in $\Sigma^\mp$, and there is no way to construct a smooth evolution. So, by moving to a manifold in which the vectors are nos strictly real-valued, we migrate to another convention where a solution is not ruled out a priori.

======================

If $x^\mu \in M$ is a position vector, it belongs to $\mathcal{H}$, but the 4-momentum $p^\mu \in M$ belongs to the tangent space to $\mathcal{H}$. Since we are dealing with flat space $\mathcal{H}$ is its own tangent space at every point, but this is not the case in general. For $\mathcal{A}$ and Ω, $x^\mu_\pm, p^\mu_\pm \in M$ are such that

$$\begin{aligned} x^\mu_+ &\in \Omega \\ p^\mu_+ &\in T\Omega \\ x^\mu_- &\in \mathcal{A} \\ p^\mu_- &\in T\mathcal{A} \quad , \end{aligned}$$

where the T indicates the tangent bundle, which is the union of the tangent spaces at every point in $\mathcal{A}$ or Ω.

Maybe say something about the cotangent bundle.

==========================

Define a Frenet frame with $x^0, \chi^4_\pm, x^i$????

$$\begin{aligned} \frac{d\mathbf{T}}{ds} &= \kappa\mathbf{N}, \\ \frac{d\mathbf{N}}{\mathrm{d}s} &= -\kappa\mathbf{T} + \tau\mathbf{B}, \\ \frac{d\mathbf{B}}{ds} &= -\tau\mathbf{N}, \end{aligned}$$

$$\begin{bmatrix} \mathbf{T}' \\ \mathbf{N}' \\ \mathbf{B}' \end{bmatrix} = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{bmatrix}$$

====================

Remark 6.2.1 It has not been established that the MCM's chronological and chirological times have the same units, so we may need to introduce new transposing parameters for $\chi^4_\pm$. Since it is expected that the relative scale of physics changes across sequential MCM worldsheets, the mechanism by which c exchanges small amounts of time for large amounts of space may have further utilities toward the changing the *level of aleph* [10].

Remark 6.2.2 Lorentz factor becomes imaginary for $v > c$. Boosts with imaginary rapidity.

Remark 6.2.3 TALK ABOUT p^0 and ADM THEOREM.

Remark 6.2.4 "The necessity for introducing PSI BAR in addition to PSI DAGGER in relativistic physics is traced back to the (+, -, -, -) signature of the Minkowski metric." FROM ZEE p97

Remark 6.2.5 On p20 in Wald, he talks about a natural isomorphism between a vector space $\mathcal{V}$ and it's double dual $\mathcal{V}^* *$. We have invented $\mathbb{C}^*$ to be such that this identification is broken, and we should explore the details after everything is set up. I don't remember if this was only for conjugation or or if it was for the dual space too. I think this will depend on how we set up the scalar product since the dual of a complex vectors is automatically conjugated.

Since tensors only have upper and lower indices referring to a vector space and its dual, with no possibility for a third space, we can probably "set something in motion" by projecting down into the space of just two spaces in some rolling fashion.

Wald talks about the energy at spacelike infinity in Ch11. The "total gravitational mass" should be useful for dark energy. On p293, WALD cites a few "more geometrical" versions of the ADM theorem that we should examine.

Remark 6.2.6 Lee says alternating tensors are those that change sign whenever two arguments are interchanged. This should have application toward the ADM theorem.

§6.3 MISC

I saw something about how rotations don't commute and it is a fudge when they build up Newtonian force diagrams of rotating vectors from the commuting infinitesimal rotations. This was maybe a source of the problem with precession.

====================

When we're constructing null-4, there will probably be times where we can use two times and two spatial dimensions in the $\mathbb{H}_2$ matrix. Since this is already set up to make spin angular momentum about a spatial axis, we can probably set up some other kind of time arrow spinor formalism by using two times and two spaces, and maybe that will be good for selecting something.

Also, we should construct two orthogonal combinations of the sum of all three spatial directions, and see if we can use those two with two times too.

§A Intrinsic Construction of the Tangent Space

§A.1 Tangent vectors as derivations

§A.2 Maybe Schuller way?

§A.3 Carroll way?

URL:

§A.4 Tangent vectors from equivalence classes of curves

==============================

While there are a number of ways to construct the tangent space to a manifold, we will follow Isham [16].

WALD and ISHAM use DERIVATIONS

LEE uses it p376 in 2ed, toward the beginning ot 1ed.

Carroll kind of uses the equivalence classes p65

Tangent space is spanned by the directional derivatives. Lee has a brief version, and Wald has a long version. Also Carroll, has something I need to revisit.

==============================

Definition A.1 The **tangency relation** $\overset{T_p}{=}$ denotes the tangency of two curves at p. If γ_1 and γ_2 are tangent at p, then

$$\gamma_1 \overset{T_p}{=} \gamma_2 \quad .$$

Definition A.2 We say a relation is an **equivalence relation** if and only if (i) S is a set, (ii) every $x \in S$ is related to x, meaning the relation is reflexive, (iii) for every $x, y \in S$, the relation of x to y implies the relation of y to x, meaning the relation is symmetric, and (iv) for every $x, y, z \in S$, the relation of x to y and the relation of y to z together imply the relation of x to z, meaning the relation is transitive. The **equivalence class** of $x \in S$, namely the set of all objects which are related to x by an equivalence relation, is denoted $[x]$.

Theorem A.3 *The tangency relation is an equivalence relation.*

Proof. Consider the set of smooth curves $S = \{\gamma_k\}$. Every $\gamma_j \in \{\gamma_k\}$ is obviously tangent to itself:

$$\gamma_j \overset{T_p}{=} \gamma_j \qquad \longleftrightarrow \qquad \frac{d}{d\lambda}\big(\varphi \circ \gamma_j\big)\bigg|_p = \frac{d}{d\lambda}\big(\varphi \circ \gamma_j\big)\bigg|_p \quad .$$

For every pair of curves $\gamma_i, \gamma_j \in \{\gamma_k\}$ tangent at p, the tangency relation is obviously symmetric:

$$\frac{d}{d\lambda}\big(\varphi \circ \gamma_i\big)\bigg|_p = \frac{d}{d\lambda}\big(\varphi \circ \gamma_j\big)\bigg|_p \qquad \Longleftrightarrow \qquad \frac{d}{d\lambda}\big(\varphi \circ \gamma_j\big)\bigg|_p = \frac{d}{d\lambda}\big(\varphi \circ \gamma_i\big)\bigg|_p \quad .$$

It is also obvious that this relation is transitive (it inherits transitivity from the = relation in the equality of derivatives), so the tangency relation is an equivalence relation.

Definition A.4 The equivalence class of curves related by $\stackrel{T_p}{=}$ is called the **tangency class of curves** at p and it is denoted $[\gamma_k(\lambda_0)]^T$. The superscript indicates that the elements of the equivalence class are related by the tangency relation rather than the usual equivalence relation denoted with the = symbol.

Remark A.5 At this point, the usual program is to declare that the tangent vectors at p are the equivalence classes of curves tangent at p. However, one might ask why we have introduced the non-standard equivalence relation $\stackrel{T_p}{=}$ at all. If the tangent vectors are going to end up as derivatives later, i.e.: the tangent vectors are the velocities of the curves passing through p, then why not use the standard equivalence relation "=" to define equivalence classes of derivatives of curves? Every object in an equivalence class is equivalent to the whole class, so if we say that a tangent vector is an equivalence class of curves, it is implied that the tangent vector is a curve. Obviously, this false implication hinges on the non-standard $\stackrel{T_p}{=}$ relation, but there is no such implication if we declare the tangent vectors as equivalence classes of curves' derivatives. In many ways, this issue mirrors the main question investigated in this paper: Why do we do build relativity with vectors in $\mathbb{R}^4$ and a non-Euclidean metric when can use the natural vectors in $\mathbb{R}^{1,3}$ with the Euclidean metric? Likewise, we not complicate our tangent spaces with $[\gamma_k(\lambda_0)]^T$. Instead, we will use the usual equivalence relation denoted = to define equivalence classes of curves' derivatives, what we will call $[d_\lambda\gamma(\lambda_0)]$ and then we will declare our tangent vectors as these classes and proceed to show that they are properly vectorial.

====================

Isham writes, "Not clear how this is ξ^μ" [16].

Definition A.6 The **equivalence class of curve derivatives** at p is denoted $[d_\lambda\gamma_k(\lambda_0)]$. The objects in this class are derivatives related by

$$\frac{d}{d\lambda}\big(\varphi\circ\gamma_1\big)\bigg|_{\gamma_1(\lambda_0)} = \frac{d}{d\lambda}\big(\varphi\circ\gamma_2\big)\bigg|_{\gamma_2(\lambda_0)} \quad ,$$

which is the definition of tangency between γ_1, γ_2 at p, as per Definition 3.3.3. Although a map φ shows up in this definition, we know from Theorem 3.3.4 that the tangency property is independent of the map φ.

Definition A.7 A **tangent vector** to a manifold at p is an *equivalence class of curve derivatives*. The **tangent space** to a manifold M at p is the vector space spanned by the tangent vectors at p. It is denoted T_pM.

Remark A.8 We have defined an infinite number of tangent vectors. For a given derivative at p, there are an infinite number of γ_k that might have that derivative there, and there are an infinite number of possible derivatives, and this agrees with the fact that there are usually an infinite number of vectors in a vector space. What we want to do is narrow this down to the linearly independent vectors that we can use as the spanning basis $\hat{e}_\mu$ for tangent space. With these basis vectors, we will write our tangent vectors as

$$V = V^\mu \hat{e}_\mu \quad ,$$

but it is not yet clear how we might relate such objects to the $[d_\lambda \gamma_k(\lambda_0)]$ classes. To do so, we will introduce *restricted equivalence classes of curve derivatives*. The idea will be to separate the derivatives of the coordinate functions x^μ in a given coordinate map φ by adding an index to the composition of the map with the curve:

$$\big(\varphi \circ \gamma\big)^\mu = x^\mu \circ \gamma \quad .$$

Written out explicitly, $(\varphi \circ \gamma)^\mu$ refers to the components of

$$\varphi \circ \gamma = \begin{bmatrix} x^1(U_\gamma) \\ \vdots \\ x^n(U_\gamma) \end{bmatrix} = \begin{bmatrix} x^1(\gamma(\lambda)) \\ \vdots \\ x^n(\gamma(\lambda)) \end{bmatrix} \quad ,$$

where U_γ is the range of γ in $U \subset M$. Since a manifold of dimension n is locally like $\mathbb{R}^n$, we know that there will be a constant number of coordinate functions associated with every coordinate map in the atlas of M. Since n is a constant, we may express the *equivalence classes of curve derivatives* as a class of related matrices:

$$\frac{d}{d\lambda}\big(\varphi \circ \gamma\big) = \begin{bmatrix} \dfrac{d}{d\lambda} x^1(\gamma(\lambda)) \\ \vdots \\ \dfrac{d}{d\lambda} x^n(\gamma(\lambda)) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial x^1}{\partial \gamma}\dfrac{d\gamma}{d\lambda} \\ \vdots \\ \dfrac{\partial x^n}{\partial \gamma}\dfrac{d\gamma}{d\lambda} \end{bmatrix} \quad .$$

The form in the matrix on the right follows from the definition of the total derivative: given a function $\xi(\chi_1, \ldots, \chi_n)$ where $\chi_k = \chi_k(t)$, the total derivative with respect to t is

$$\frac{d\xi}{dt} = \frac{\partial \xi}{\partial \chi_1}\frac{d\chi_1}{dt} + \cdots + \frac{\partial \xi}{\partial \chi_n}\frac{d\chi_n}{dt} \quad .$$

This relates to $x^\mu(\gamma(\lambda))$ as the simplest case of a single $\chi(t)$ function. It will suffice at this point to let the derivative with respect to γ be defined by the total derivative. If two derivatives are in $[d_\lambda \gamma(\lambda_0)]$, then

$$\left(\frac{\partial x^\mu}{\partial \gamma}\bigg|_{\gamma_1(\lambda_0)} \frac{d\gamma}{d\lambda}\bigg|_{\lambda_0} \right) = \left(\frac{\partial x^\mu}{\partial \gamma}\bigg|_{\gamma_1(\lambda_0)} \frac{d\gamma}{d\lambda}\bigg|_{\lambda_0} \right) \quad .$$

The derivatives of $x^\mu : I \subset \mathbb{R} \to \mathbb{R}$ are clearly real numbers, so we can select classes where the derivatives are equal to certain numbers. These will be the restricted equivalence classes, and we define them as basis vectors. Then we will show that they are properly vectorial, that they span a vector space, and that the linear independence of the basis did not depend on the choice of φ or its components x^μ.

Definition A.9 A **restricted tangency class** $[\gamma'_k(\lambda_0)]_\sigma$ contains all derivatives equal at p whose components vanish when $\mu \neq \sigma$. These derivatives satisfy

$$\frac{d}{d\lambda}\big(\varphi \circ \gamma_k\big)^\mu\bigg|_{p=\gamma_k(\lambda_0)} = \left(\frac{\partial x^\mu}{\partial \gamma_k}\bigg|_{\gamma_k(\lambda_0)} \frac{d\gamma_k}{d\lambda}\bigg|_{\lambda_0} \right) = \begin{cases} 0 & \text{if} \quad \mu \neq \sigma \\ c & \text{if} \quad \mu = \sigma \end{cases} \qquad ,$$

where c is any non-zero, real-valued constant.

Example A.10 Restricted tangency classes in $n = 4$. If, for example, $\varphi : M \to \mathbb{R}^4$, then we will have four types of restricted tangency classes associated with $\sigma = \{1, 2, 3, 4\}$:

$$[\gamma'_k(\lambda_0)]_1 = \begin{bmatrix} c \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad [\gamma'_k(\lambda_0)]_2 = \begin{bmatrix} 0 \\ c \\ 0 \\ 0 \end{bmatrix} \quad [\gamma'_k(\lambda_0)]_3 = \begin{bmatrix} 0 \\ 0 \\ c \\ 0 \end{bmatrix} \quad [\gamma'_k(\lambda_0)]_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ c \end{bmatrix} \quad .$$

There is one such class for every non-zero $c \in \mathbb{R}$.

Theorem A.11 *An equivalence class of curve derivatives $[\gamma'_k(\lambda_0)]$ is a vector.*

Proof. THIS NEEDS TO MAKE A DISTINCTION IN THE TRANSFORMATION OF THE COMPONENTS AND THE BASIS VECTORS!!!!!!!!!!!
========================

An object that transforms as a vector is a vector, so it will suffice to demonstrate compliance with the vector transformation law

$$V^\mu = \frac{\partial x^\mu}{\partial x^{\mu'}} V^{\mu'} \quad .$$

Given two different coordinate maps

$$\varphi = (x^1, \ldots, x^n) \quad , \quad \text{and} \qquad \psi = (y^1, \ldots, y^n) \quad ,$$

we have

$$\frac{d}{d\lambda}\big(x^\mu \circ \gamma\big) = \frac{\partial x^\mu}{\partial \gamma}\frac{d\gamma}{d\lambda} \quad , \quad \text{and} \qquad \frac{d}{d\lambda}\big(y^\nu \circ \gamma\big) = \frac{\partial y^\nu}{\partial \gamma}\frac{d\gamma}{d\lambda} \quad .$$

Multiplying the expression on the left by $\partial y^\nu / \partial x^\mu$ clearly yields the expression on the right. The derivatives transform as vectors, so they are vectors.

====================================

Theorem A.12 *The restricted tangency classes form a basis for a vector space.*

Proof. First we will show that they satisfy the vector space axioms. Then we will show that they are linearly independent. If we multiply a restricted tangency class by a scalar,

$$\beta[\gamma'_k(\lambda_0)]_\sigma = \begin{cases} 0 & \text{if} \quad \mu \neq \sigma \\ \beta c & \text{if} \quad \mu = \sigma \end{cases}$$

this merely changes the constant in the non-zero entry. The constant was unrestricted, so the product of a vector and a scalar is still a vector. Considering

$$\big(\beta_1\beta_2\big)[\gamma'_k(\lambda_0)]_\sigma = \beta_1\big(\beta_2[\gamma'_k(\lambda_0)]_\sigma\big)$$
$$\big(\beta_1+\beta_2\big)[\gamma'_k(\lambda_0)]_\sigma = \beta_1[\gamma'_k(\lambda_0)]_\sigma + \beta_2[\gamma'_k(\lambda_0)]_\sigma$$

shows that scalar multiplication is associative and distributive over addition. The sum of two vectors is another vector:

$$[\gamma'_k(\lambda_0)]_\sigma + [\gamma'_k(\lambda_0)]_\rho = \begin{cases} 0 & \text{if} \quad \mu \neq \sigma \quad \text{and} \quad \mu \neq \rho \\ c & \text{if} \quad \mu = \sigma \\ c & \text{if} \quad \mu = \rho \end{cases}$$

This sum clearly belongs to an unrestricted tangency class $[\gamma'_k(\lambda_0)]$. The sum of two basis vectors has not yielded another basis vector, but instead the sum is just some other vector in the vector space. This is as expected since general vectors are constructed from sums of basis vectors and their scalar multiples. Vector addition is obviously commutative, meaning $a+b=b+a$, and it is equally obviously associative, meaning $(a+b)+c=a+(b+c)$. The inverse of any vector is itself scalar multiplied by -1, and the zero vector exists in the space since we have already demonstrated closure under addition, and a vector plus its inverse is the zero vector. Finally, the elements of a basis for a vector space must be linearly independent, and it is obvious that

$$\{[\gamma'_k(\lambda_0)]_1, \dots, [\gamma'_k(\lambda_0)]_n\} = \left\{ \begin{bmatrix} c \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \dots, \begin{bmatrix} \vdots \\ 0 \\ 0 \\ c \end{bmatrix} \right\} \quad ,$$

is a set of linearly independent vectors; if the scalar multiples of two elements sum to zero, then the scalars were zero:

$$\beta_i[\gamma'_k(\lambda_0)]_i + \beta_j[\gamma'_k(\lambda_0)]_j = 0 \quad \Longrightarrow \quad \beta_i = \beta_j = 0 \ .$$

Theorem A.13 *The linear independence of a set of restricted tangency classes does not depend on the choice of x^μ coordinate functions.*

Proof. XXXXXXXXXXXXXXXXXXX

Remark A.14 A Lorentzian manifold is *smooth* (Definition 3.2.26), so it includes a *smooth structure* (Definition 3.2.25). This requires that the *transition functions* (Definition 3.2.19) from one chart to another are C^∞-*compatible* (Definition 3.2.23). Thus, referring to Figure XXX, we know that $\psi \circ \varphi^{-1}$ and $\varphi \circ \psi^{-1}$ (mutual inverses) are diffeomorphisms, and we will call them simply f^{a} and $(f^{\mathsf{a}})^{-1}$ here. As $\mathbb{R}^n \to \mathbb{R}^n$ maps, each coordinate has its own function f^μ or $(f^\mu)^{-1}$ (Example 3.2.13), and the components are scalar diffeomorphisms. As such, it is already established that we may take derivatives of $f^\mu : \mathbb{R} \to \mathbb{R}$ with respect to its variables: the coordinates of $\mathbb{R}^n$. The vector of the derivatives is called the gradient and written

$$\boldsymbol{\nabla} f = (\partial_1 f, \ldots, \partial_n f) \quad , \quad \text{or} \qquad \boldsymbol{\nabla} f = \sum \frac{\partial f}{\partial x^\mu} \hat{e}_\mu \quad .$$

Each component is the rate of change of f along the direction of some basis vector. However, we might also want to know the rate of change of f along an arbitrary direction. Since the gradient is the vector field of the rates of change, determining the scalar rate of change in an arbitrary direction requires a scalar product, which is obviously the dot product with a vector pointing in the direction in question. Just like it was already guaranteed that the derivatives of the transition functions exist, we know such a vector exists because $\mathbb{R}^n$ is a vector space. So, then, by the smooth structure of our manifold, we are automatically equipped with a directional derivative. For obvious reasons, we will choose to define the directional derivative of a function along a vector with the unit vector pointing in that direction:

$$\boldsymbol{\nabla} f \cdot \hat{n} \equiv \frac{\partial f}{\partial \hat{n}}$$

At this point, we say that the various directional derivatives are the basis vectors in the tangent space, but this requires an element of hand-waving since the directional derivatives are scalars. Therefore, we will carefully construct the tangent space here. Among many possible ways to construct the tangent space, we will use *curves* since we have already introduced *coordinate maps.*

============

The main ways to construct the tangent spaces to a manifold without first embedding the manifold in a higher dimensional space are to use equivalence classes of curves in the manifold, or to define a set of *derivations* and then say one or the other are identified with the spanning basis of the tangent space we are looking for. The consensus in the literature seems to be that the derivation approach is easier, but it is unsatisfying in many ways because the derivations themselves are contrived, at least in the present context of manifolds. In that regard, Isham describes the derivation approach as being algebraic while the equivalence classes of curves are more geometric. So, we will carefully develop the equivalence class approach and then show that we have achieved the usual formalism in which the tangent space to a manifold at a point is spanned by the directional derivatives at that point.

==============

The derivative is just a scalar. To make it into a vector, we need to give it a direction.

Example A.15 The tangent vectors are directional derivative operators. In the above, it is not immediately clear what is meant by the derivative with respect to γ. To avoid this notation, Carroll uses:

$$x^\mu \circ \gamma = \big(x^\mu \circ \psi^{-1}\big) \circ \big(\psi \circ \gamma\big) \quad .$$

He also uses f for the coordinate function while we have used x^μ. So, we have applied the chain rule as

$$\frac{d}{d\lambda}\big(x^\mu \circ \gamma\big) = \frac{\partial x^1}{\partial \gamma}\frac{d\gamma}{d\lambda} \quad ,$$

Carroll avoids taking the derivative with respect to γ as

$$\frac{d}{d\lambda}\left[\big(x^\mu \circ \psi^{-1}\big) \circ \big(\psi \circ \gamma\big)\right] = \frac{\partial\big(x^\mu \circ \psi^{-1}\big)}{\partial\big(\psi \circ \gamma\big)}\frac{d\big(\psi \circ \gamma\big)}{d\lambda} \quad ,$$

As before, the compositions $\psi \circ \gamma$ are just the coordinates of the image of γ in $\mathbb{R}^n$, so

$$\frac{d}{d\lambda}\left[\big(x^\mu \circ \psi^{-1}\big) \circ \big(\psi \circ \gamma\big)\right] = \frac{\partial\big(x^\mu \circ \psi^{-1}\big)}{\partial\big(\psi \circ \gamma\big)^\nu}\frac{d\big(\psi \circ \gamma\big)^\nu}{d\lambda} = \frac{\partial\big(x^\mu \circ \psi^{-1}\big)}{\partial y^\nu}\frac{dy^\nu}{d\lambda} \quad .$$

Note that an upper index on the bottom of a derivative is treated as lower index, and the sum reflects the definition of the total derivative:

$$\frac{d\xi}{dt} = \frac{\partial \xi}{\partial \chi_1}\frac{d\chi_1}{dt} + \cdots + \frac{\partial \xi}{\partial \chi_n}\frac{d\chi_n}{dt} \quad ,$$

Noting that $\psi^{-1} : \mathbb{R} \to M$ and $x^\mu : M \to \mathbb{R}$ (meaning that x^μ is component of φ rather than the whole vector), we may introduce notation such that

$$x^\mu \circ \psi^{-1} : \mathbb{R}^n \to \mathbb{R} \qquad \Longrightarrow \qquad x^\mu \circ \psi^{-1} = f(y^1, \ldots, y^n) \quad .$$

Here, this notation means that f is a scalar function of the coordinates y^ν of the $\mathbb{R}^n$ domain (not to be confused with the value of f at some particular y^ν.) Thus, by inserting the inverses at the intermediate step to make a composition of compositions, Carroll is able to write the derivatives in our tangency classes as

$$\frac{d}{d\lambda}\big(x^\mu \circ \gamma\big) = \frac{\partial f(y^1, \ldots, y^n)}{\partial y^\nu}\frac{dy^\nu}{d\lambda} \qquad \Longrightarrow \qquad \frac{d}{d\lambda} = \frac{dy^\nu}{d\lambda}\partial_\nu \quad .$$

Comparing to the directional derivative

$$\boldsymbol{\nabla} f \cdot \boldsymbol{v} = \frac{\partial f(\boldsymbol{y})}{\partial \boldsymbol{y}} \cdot \boldsymbol{v}$$

So, then we can compare the usual definition of a tangent vector to the expression we have obtained for $d/d\lambda$. Given a vector $\boldsymbol{r}(t) = (x^1(t), \ldots, x^n(t))$, the tangent vector is

$$\frac{d\boldsymbol{r}}{dt} = \sum_i \frac{\partial \boldsymbol{r}}{\partial x^i}\frac{dx^i}{dt}$$

========================

we find that the components of $\boldsymbol{v}$ must be the tangent vector to the curve in $\mathbb{R}^n$ parameterized in λ, namely $dy^\nu/d\lambda$. Since we have written the directional derivative of an *arbitrary* function, it is acceptable that have identified f with the coordinate functions x^μ in an arbitrary XXXXXXXXXXX

======================

We have declared the atlas to be maximal so all smooth functions were included, as the maps φ, and the coordinate functions are therefore associated with any arbitrary smooth function. This is the connection to the directional derivative or an arbitrary smooth function.

Definition A.16 The partial derivatives ∂_μ are called the **coordinate basis** for T_pM.

Example A.17 The partial derivatives in a coordinate basis for T_pM are the restricted tangency classes. In $n = 4$, the restricted tangency classes were

$$[\gamma_k(\lambda_0)]_1^T = \begin{bmatrix} c \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad [\gamma_k(\lambda_0)]_2^T = \begin{bmatrix} 0 \\ c \\ 0 \\ 0 \end{bmatrix} \quad [\gamma_k(\lambda_0)]_3^T = \begin{bmatrix} 0 \\ 0 \\ c \\ 0 \end{bmatrix} \quad [\gamma_k(\lambda_0)]_4^T = \begin{bmatrix} 0 \\ 0 \\ 0 \\ c \end{bmatrix} \quad .$$

To obtain these constant derivatives, we needed to pick some certain coordinate map φ to complete the map from I to $\mathbb{R}^n$. The composition $\varphi \circ \gamma$ takes $\lambda \in I$ and returns a parameterized curve in $\mathbb{R}^n$. The parameterized components of the curve are necessarily $x^\mu \circ \gamma$, and we explained in the previous example that these are any arbitrary C^∞ functions because the atlas on M includes all possible C^∞ maps (Theorem 3.2.24). After dividing this into $(x^\mu \circ \psi^{-1}) \circ (\psi \circ \gamma)$, we introduced $x^\mu \circ \psi^{-1} = f(y^1, \ldots, y^n)$. The partials are called the coordinate basis for T_pM because we need to choose the arbitrary y^μ coordinates associated with ψ before we can take the derivative of an arbitrary function. So, the tangency classes are not associated with any particular basis, but the partials must be. The equivalence of the two is demonstrated as

$$\partial_1 f = \begin{bmatrix} c \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \partial_2 f = \begin{bmatrix} 0 \\ c \\ 0 \\ 0 \end{bmatrix} \qquad \partial_3 f = \begin{bmatrix} 0 \\ 0 \\ c \\ 0 \end{bmatrix} \qquad \partial_4 f = \begin{bmatrix} 0 \\ 0 \\ 0 \\ c \end{bmatrix} \quad .$$

The gradient of a scalar is a vector

$$\nabla f = \left(\frac{\partial f}{\partial y^1}, \ldots, \frac{\partial f}{\partial y^n} \right)$$

The basis vectors of a contravariant vector have lower indices. The gradient is a covariant vector field.

??????????????????

§B Major Definitions, Alphabetized

References

[1] Andrew M. Steane. An Introduction to Spinors. *arXiv:1312.3824*, (2013).

[2] Sean M. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*. Pearson, (2004).

[3] Paul Roman. *Advanced Quantum Theory*. Addison Wesley, (1965).

[4] Paul Tipler and Ralph Llewellyn. *Modern Physics, 5th ed.* W. H. Freeman, (2007).

[5] Andrew M. Steane. *Relativity made Relatively Easy*. Oxford University Press, (2012).

[6] James F. Woodward. Killing Time. *Foundations of Physics Letters.* ***9*** *1–23*, (1996).

[7] Jonathan W. Tooker. Quick Disproof of the Riemann Hypothesis. *viXra:1906.0236*, (2019).

[8] Jonathan W. Tooker. Zeros of the Riemann Zeta Function within the Critical Strip and off the Critical Line. *viXra:1912.0030*, (2019).

[9] Jonathan W. Tooker. Fractional Distance: The Topology of the Real Number Line. *viXra:2111.0072*, (2019).

[10] Jonathan W. Tooker. Sixty-Six Theses: Next Steps and the Way Forward in the Modified Cosmological Model. *viXra:2206.0152*, (2022).

[11] Robert M. Wald. *General Relativity*. University of Chicago Press, (1984).

[12] John M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Springer, (1997).

[13] William M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry, Revised 2nd Ed.* World Scientific, (1998).

[14] John M. Lee. *Introduction To Topological Manifolds*. Springer, (2000).

[15] John Lee. *Introduction to Smooth Manifolds, 2nd Ed.* Springer, (2013).

[16] Chris J. Isham. *Modern Differential Geometry for Physicists, 2nd Ed.* World Scientific, (1999).

[17] Roger Penrose and Wolfgang Rindler. *Spinors and Space-Time: Volume 1, Two-Spinor Calculus and Relativistic Fields*. Cambridge University Press, (1987).

[18] David N. Williams. Introduction to Lorentz Spinors. `https://websites.umich.edu/ williams/notes/spinor.pdf`.

[19] E.C. Zeeman. The topology of minkowski space. *Topology*, 6(2):161–170, 1967.

[20] Jonathan W. Tooker. Ontological Physics. *viXra:1312.0168*, (2013).

[21] Jonathan W. Tooker. Time Arrow Spinors for the Modified Cosmological Model. *viXra:1807.0454*, (2018).

[22] Arthur Jaffe. Lorentz transformations, Rotations, and Boosts, (2013). `https://api.semanticscholar.org/CorpusID:51809955`.

[23] Steven Weinberg. *The Quantum Theory of Fields*. Cambridge University Press, (1995).

[24] Jonathan W. Tooker. Quantum Structure. *viXra:1302.0037*, (2013).