Special Relativistic Metric

by Alec Johnson, February 2009 revised February 2011

1 Axioms of inertial frames.

Special relativity begins with a set of axioms that constrain physical laws to satisfy certain invariance properties (called inertial invariants) under transformations between inertial coordinate frames (which we will also refer to as *inertial transformations* or *Lorentz transformations*).

1. **Relativity.** *Physical laws are inertially invariant.*

   This means that inertial coordinate transformation commutes with writing down (evaluating) physical laws.

2. **Light speed.** *There exists a speed (called c, the speed of light) which is an inertial invariant.*

   This means that for a path $\mathbf{x}(t)$, $(c\,dt)^2 - (d\mathbf{x})^2 = 0$ is an invariant condition under inertial coordinate transformation.

3. **Uniform motion.** *Uniform motion is an inertial invariant.*

   This means that under inertial transformations straight lines in space-time are mapped to straight lines, which implies (by considering linear meshes, for example) that the transformation between inertial coordinate systems is affine. (We assume here that inertial coordinate transformations are invertible.)

4. **Classical limit.** *Classical mechanics holds in the limit of slow velocities.*

   In particular, the force must equal the derivative of the (classical) momentum. (We derive special relativistic dynamics by writing down laws that are invariant under relativistic inertial transformations and that agree with classical mechanics in the slow-speed limit.)

Axioms 2 and 3 imply relativistic kinematics; i.e., they constrain the set of allowed coordinate transformations.

Axiom 1 is a meta-principle. We take it to imply that the laws of mechanics (relativistic Newtonian force laws and Maxwell's electromagnetism laws) must be invariant with respect to the class of inertial transformations selected by the kinematic axioms 2 and 3. We derive relativistic mechanics from axioms 1, 4, and relativistic kinematics.

These axioms admit rescaling of space and time by an arbitrary positive or negative scalar. To restrict the class of inertial transformations so that the relative velocity of one frame with respect to another uniquely determines the scale of space and time, it is necessary to impose an additional symmetry assumption such as:

5′. **velocity symmetry.** *If (the origin of) inertial system A has velocity* $\mathbf{v}$ *in inertial system B then (the origin of) B has velocity* $-\mathbf{v}$ *in A.*

Axiom 5′ ensures that the "rulers" (or "stop-watches") used in system A and B have equivalent scales (or rates).

2 Minkowski 4-space

Minkowski gave an elegant formulation of Einstein's theory of relativity in terms of space and (rescaled) time coordinates.

Definition 2.1. The **four-vector** coordinates of a point in space-time at time t and position $\mathbf{x}$ are defined to be

$$x^\mu := (x^0, x^1, x^2, x^3)^T, \text{ where } x_0 := ct.$$

Proposition 2.2. *The linear part of the coordinate transformation from an unprimed coordinate system to a primed system is given by* $dx^{\nu'} = \Sigma_\mu \Lambda_\mu^{\nu'} dx^\mu$, *where*

$$\Lambda_\mu^{\nu'} := \frac{\partial x'^\nu}{\partial x^\mu}.$$

Convention 2.3 (Einstein summation)**.** An index appearing exactly twice in a product once as a subscript and once as a superscript implies summation over that index: $\Lambda_\mu^{\nu'} dx^\mu := \Sigma_\mu \Lambda_\mu^{\nu'} dx^\mu$.

Definition 2.4 (Lorentz-invariant)**.** We say that an equation in 4-space is **Lorentz-invariant** if transforming it to another inertial reference frame gives an equivalent equation. (For example, if $d\tau$ (the elapse of "proper time") is independent of coordinates, then multiplying $\Lambda_\mu^{\nu'}$ by the equation $\tilde{v}^\mu := \frac{dx^\mu}{d\tau}$ defining proper velocity gives $\tilde{v}^{\mu'} = \frac{dx^{\mu'}}{d\tau}$, i.e., proper velocity is Lorentz-invariant (i.e. a well-defined physical quantity independent of coordinates)).

Definition 2.5. The **interval** between two events separated by a displacement (dt, dx) is defined to be $(c\,dt)^2 - (dx)^2$.

Definition 2.6. The **4-vector scalar product** of two vectors v^μ and u^μ is $v^0u^0 - v^1u^1 - v^2u^2 - v^3u^3$. The matrix of this bilinear form,

$$\eta^{\mu\nu} := \eta_{\mu\nu} := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

is called the matrix of contravariant (or covariant) components of the Lorentz **metric**.

Definition 2.7 (Dual vectors)**.** Observe that multiplying the Lorentz metric by a vector negates the spatial components. We adopt the convention that the vectors $v^\mu = (v^0, v^1, v^2, v^3)^T$ and $v_\mu = (v_0, v_1, v_2, v_3) := (v^0, -v^1, -v^2, -v^3)^T$ are **dual vectors** of one another. Then we can write

$$v^\mu = \eta^{\mu\nu} v_\nu \qquad \text{and} \qquad v_\mu = \eta_{\mu\nu} v^\nu.$$

Thus by the Einstein summation convention the 4-vector scalar product of v^μ with u^μ may be be written

$$v^\mu u_\mu = v^\mu \eta_{\mu\nu} u^\nu = v_\mu \eta^{\mu\nu} u_\nu.$$

In the next section we argue that Lorentz transformations are characterized by the requirement that they preserve the interval between two events. This is equivalent to the requirement that they preserve the Lorentz metric. Indeed,

$$\mathrm{d}x^{\mu'}\mathrm{d}x_{\mu'} := \mathrm{d}x^{\mu'}\eta_{\mu'\nu'}\mathrm{d}x^{\nu'} = \mathrm{d}x^{\mu}\Lambda_{\mu}^{\mu'}\eta_{\mu'\nu'}\Lambda_{\nu}^{\nu'}\mathrm{d}x^{\nu},$$

which equals $\mathrm{d}x^{\mu}\eta_{\mu\nu}\mathrm{d}x^{\nu} =: \mathrm{d}x^{\mu}\mathrm{d}x_{\mu}$ if (and in general only if) $\Lambda_{\mu}^{\mu'}\eta_{\mu'\nu'}\Lambda_{\nu}^{\nu'} = \eta_{\mu\nu}$.

3 Interval invariance

The purpose of this section is to show that light cone invariance plus velocity symmetry implies interval invariance. We argue that any inertial transformation is a Lorentz "boost" (velocity shift) in the x-direction preceeded and followed by a rotation.

Definition 3.1. The **light cone** of a reference frame is the set of all displacement vectors whose interval is zero, i.e., whose speed is the speed of light.

Proposition 3.2. *Axiom 2 says that inertial transformations map the light cone to itself.*

Definition 3.3. We call positive intervals **timelike**, negative intervals **spacelike**, and zero intervals **lightlike**. These terms denote properties that are well-defined (invariant under inertial transformation) due to the invariance of the light cone, in accordance with the following theorem.

Definition 3.4. The rescaling factor for time is called the **Lorentz** factor: $\gamma = \frac{dt}{d\tau}$, where $d\tau$ denotes a time interval between events in a frame of reference where they occur in the same position.

Theorem 3.5. *For one dimension of space the coordinate transformation from a primed frame moving at velocity v to a stationary frame is given by*

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{bmatrix} \gamma & \gamma v/c \\ \gamma v/c & \gamma \end{bmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix}. \tag{1}$$

Proof. Without loss of generality $c = 1$. Call the linear transformation Λ. Light speed is invariant, so $(1,1)$ and $(1,-1)$ are orthogonal eigenvectors, with eigenvalues, say, λ_1, λ_2. So Λ is symmetric. (Indeed, $\Lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda_1+\lambda_2 & \lambda_1-\lambda_2 \\ \lambda_1-\lambda_2 & \lambda_1+\lambda_2 \end{bmatrix} \Big/ 2$.) By Definition 3.4 and the definition of v, $\Lambda \cdot (1,0) = (\gamma, v\gamma)$, as needed. Remark: Observe that $\lambda_1 = \gamma(1+v/c)$, $\lambda_2 = \gamma(1-v/c)$, and $\lambda_1\lambda_2 = \gamma^2(1-(v/c)^2)$. □

Corollary 3.6. *For one dimension of space the coordinate transformation from a stationary frame to a primed frame moving at velocity v is called a* boost *and is given by*

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{bmatrix} \gamma & -\gamma v/c \\ -\gamma v/c & \gamma \end{bmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}. \tag{2}$$

Proof. Use assumption $5'$. □

Theorem 3.7. $\gamma = \frac{1}{\sqrt{1-(v/c)^2}}$, *where v is the relative speed of the one reference frame with respect to the other.*

Proof. Compose the Lorentz transformation for a boost with its inverse:

$$\begin{aligned} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} \gamma & -\gamma v/c \\ -\gamma v/c & \gamma \end{bmatrix} \begin{bmatrix} \gamma & \gamma v/c \\ \gamma v/c & \gamma \end{bmatrix} \\ &= \begin{bmatrix} \gamma^2(1-(v/c)^2) & 0 \\ 0 & \gamma^2(1-(v/c)^2) \end{bmatrix} \end{aligned}$$

□

Corollary 3.8 (formula for 3D boost)**.** *For 3D space the coordinate transformation from a stationary frame to a primed frame moving at velocity* $\mathbf{v}$ *is given by*

$$\begin{pmatrix} ct' \\ \mathbf{x}' \end{pmatrix} = \begin{bmatrix} \gamma & \mathbf{u}/c \\ \mathbf{u}/c & \mathbb{I} + \frac{\mathbf{uu}}{c^2(\gamma+1)} \end{bmatrix} \begin{pmatrix} ct \\ \mathbf{x} \end{pmatrix}, \tag{3}$$

where $\mathbf{u} := \gamma\mathbf{v}$ *and* $\gamma^2 = 1 + |\mathbf{u}/c|^2$ *and note that* $\frac{\mathbf{uu}}{c^2(\gamma+1)} = (\gamma-1)\frac{\mathbf{uu}}{c^2(\gamma^2-1)} = (\gamma-1)\frac{\mathbf{uu}}{|\mathbf{u}|^2}$.

Proof. By tensoriality it is enough to consider the special case where $\mathbf{u} = (u,0,0)$, for which:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{bmatrix} \gamma & u/c & 0 & 0 \\ u/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}. \tag{4}$$

□

Theorem 3.9. *The interval between two events is an inertial invariant (and thus well-defined).*

Proof. This takes some work. The way I know how to do this is to explicitly deduce the coordinate transformation for a *boost* (an inertial transformation between coordinate systems with aligned spatial axes), and a *rotation*, show that they preserve intervals, and argue that all inertial coordinate transformations are compositions of these two types. □

Theorem 3.10. *Invariance of intervals characterizes inertial transformations.*

Proof. The proof is analogous to proving that distance-preserving mappings preserve the dot product and hence are linear orthonormal affine maps. (In this case a non-positive-definite bilinear form takes the place of the dot product). □

Remark 3.11 (kinematic invariant)**.** The set of linear transformations Λ that map the light cone to itself is characterized by the property that the interval between any two events is multiplied by $\sqrt[n]{(\det\Lambda)^2}$.

References

[1] David J. Griffiths. Chapter 10 of *Introduction to Electrodynamics.* Springer-Verlag, ©1995.

[2] J. Foster and J.D. Nightingale. Appendix A of *A Short Course in General Relativity.* Springer-Verlag, ©1995.

[3] David McMahon. *Relativity demystified.* McGraw-Hill, ©2006.

[4] Patricia M. Schwarz and John H. Schwarz. *Special Relativity: From Einstein to Strings.* Cambridge, ©2004.