2

From Electrodynamics to

SPECIAL RELATIVITY

Introduction. We have already had occasion to note that "Maxwell's trick" implied—tacitly but inevitably—the *abandonment of Galilean relativity*. We have seen how this development came about (it was born of Maxwell's desire to preserve charge conservation), and can readily appreciate its revolutionary significance, for

> To the extent that Maxwellean electrodynamics is physically correct, Newtonian dynamics—which *is* Galilean covariant—must be physically in error.

. . . but have now to examine the more detailed ramifications of this formal development. The issue leads, of course, to *special relativity*.

That special relativity is—though born of electrodynamics—"bigger" than electrodynamics (*i.e.*, that it has non-electrodynamic implications, applications —and roots) is a point clearly appreciated by Einstein himself (1905). Readers should understand, therefore, that my intent here is a limited one: my goal is not to produce a "complete account of the special theory of relativity" but only *to develop those aspects of special relativity which are* <u>*specifically relevant to our electrodynamical needs*</u> . . . and, conversely, to underscore those aspects of electrodynamics which are of a peculiarly "relativistic" nature.

In relativistic physics c —born of electrodynamics and called (not quite appropriately) the "velocity of light"—is recognized for what it is: a *constant* *of Nature* which would retain its relevance and more fundamental meaning "even if electrodynamics—light—did not exist." From

$$[\,c\,] = \text{velocity} = LT^{-1}$$

we see that in "c-physics" we can, if we wish, *measure temporal intervals in the units of spatial length*. It is in this spirit—and because it proves formally to be very convenient—that we agree henceforth to write

$$x^0 \equiv ct \quad \text{and} \quad \begin{cases} x^1 \equiv x \\ x^2 \equiv y \\ x^3 \equiv z \end{cases}$$

To indicate that he has used his "good clock and Cartesian frame" to assign coordinates to an "event" (*i.e.*, to a point in space at a moment in time: briefly, to a point in spacetime) an inertial observer O may write x^μ with $\mu \in \{0, 1, 2, 3\}$. Or he may (responding to the convenience of the moment) write one of the following:

$$x \equiv \begin{pmatrix} x^0 \\ \boldsymbol{x} \end{pmatrix} \equiv \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

We agree also to write

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu}\,, \quad \text{and also} \quad \partial \equiv \begin{pmatrix} \partial_0 \\ \boldsymbol{\nabla} \end{pmatrix} \equiv \begin{pmatrix} \partial_0 \\ \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix}$$

Note particularly that $\partial_0 = \frac{1}{c}\partial_t$. We superscript x's but subscript ∂'s in anticipation of a fundamental transformation-theoretic distinction that will be discussed in §2.

It is upon this notational base—simple though it is—that we will build.

1. Notational reexpression of Maxwell's equations. Even simple thoughts can be rendered unintelligible if awkwardly expressed ... and Maxwell's was hardly a "simple thought." It took physicists the better part of 40 years to gain a clear sense of the essentials of the theory that Maxwell had brought into being (and which he himself imagined to be descriptive of the mechanical properties of an imagined but elusive "æther"). Running parallel to the ever-deepening physical insight were certain notational adjustments/simplifications inspired by developments in the world of pure mathematics.[77]

During the last decade of that formative era increasing urgency attached to a question

[77] See "Theories of Maxwellian design" (1998).

> *What are the (evidently non-Galilean) transformations which preserve the form of Maxwell's equations?*

was first posed (1899) and resolved (1904) by H. A. Lorentz (1853–1928), who was motivated by a desire to avoid the *ad hoc* character of previous attempts to account for the results of the Michelson–Morley, Trouton–Noble and related experiments. Lorentz' original discussion[78] strikes the modern eye as excessively complex. The discussion which follows owes much to the mathematical insight of H. Minkowski (1864–1909),[79] whose work in this field was inspired by the accomplishments of one of his former students (A. Einstein), but which has roots also in Minkowski's youthful association with H. Hertz (1857–1894), and is distinguished by its notational modernism.

Here we look to the notational aspects of Minkowski's contribution, drawing tacitly (where Minkowski drew explicitly) upon the notational conventions and conceptual resources of tensor analysis. In a reversal of the historical order, I will in §2 let the pattern of our results serve to *motivate* a review of tensor algebra and calculus. We will be placed then in position to observe (in §3) the sense in which special relativity almost "invents itself." Now to work:

Let Maxwell's equations (65) be notated

$$\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{E} &= \rho \\ \boldsymbol{\nabla}\times\boldsymbol{B} - \tfrac{1}{c}\tfrac{\partial}{\partial t}\boldsymbol{E} &= \tfrac{1}{c}\boldsymbol{j} \\[6pt] \boldsymbol{\nabla}\cdot\boldsymbol{B} &= 0 \\ \boldsymbol{\nabla}\times\boldsymbol{E} + \tfrac{1}{c}\tfrac{\partial}{\partial t}\boldsymbol{B} &= \boldsymbol{0} \end{aligned}$$

where, after placing all fields on the left and sources on the right, we have grouped together the "sourcy" equations (Coulomb, Ampere), and formed a second quartet from their sourceless counterparts. Drawing now upon the notational conventions introduced on the preceding page we have

$$\left.\begin{aligned} \partial_1 E_1 + \partial_2 E_2 + \partial_3 E_3 &= \tfrac{1}{c}j^0 \equiv \rho \\ -\partial_0 E_1 \qquad\quad + \partial_2 B_3 - \partial_3 B_2 &= \tfrac{1}{c}j^1 \\ -\partial_0 E_2 - \partial_1 B_3 \qquad\quad + \partial_3 B_1 &= \tfrac{1}{c}j^2 \\ -\partial_0 E_3 + \partial_1 B_2 - \partial_2 B_1 \qquad\quad &= \tfrac{1}{c}j^3 \end{aligned}\right\} \tag{157.1}$$

[78] Reprinted in English translation under the title "Electromagnetic phenomena in a system moving with any velocity less than that of light" in *The Principle of Relativity* (1923), a valuable collection reprinted classic papers which is still available in paperback (published by Dover).

[79] See §7 of "Die Grundgleichungen für die elektromagnetischen Vorgänge in bewegten Körpen" (1907) in Minkowski's *Collected Works*.

$$\left.\begin{aligned}
-\partial_1 B_1 - \partial_2 B_2 - \partial_3 B_3 &= 0\\
+\partial_0 B_1 \qquad\qquad + \partial_2 E_3 - \partial_3 E_2 &= 0\\
+\partial_0 B_2 - \partial_1 E_3 \qquad\qquad + \partial_3 E_1 &= 0\\
+\partial_0 B_3 + \partial_1 E_2 - \partial_2 E_1 \qquad\qquad &= 0
\end{aligned}\right\} \tag{157.2}$$

where we have found it formally convenient to write

$$j \equiv \begin{pmatrix} j^0 \\ \boldsymbol{j} \end{pmatrix} = \begin{pmatrix} j^0 \\ j^1 \\ j^2 \\ j^3 \end{pmatrix} \quad \text{with} \quad j^0 \equiv c\rho \tag{158}$$

It is evident that (157.1) could be written in the following remarkably compact and simple form

$$\partial_\mu F^{\mu\nu} = \tfrac{1}{c} j^\nu$$

NOTE: Here as always, *summation* $\sum\limits_0^3$ *on the repeated index is understood.*

provided the $F^{\mu\nu}$ are defined by the following scheme:

$$\mathbb{F} \equiv \begin{pmatrix} F^{00} & F^{01} & F^{02} & F^{03} \\ F^{10} & F^{11} & F^{12} & F^{13} \\ F^{20} & F^{21} & F^{22} & F^{23} \\ F^{30} & F^{31} & F^{32} & F^{33} \end{pmatrix} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}$$

$$\equiv \mathbb{A}(\boldsymbol{E}, \boldsymbol{B}) \tag{159}$$

Here the $\mathbb{A}$-notation is intended to emphasize that the 4×4 matrix in question is <u>antisymmetric</u>; as such, it has $\begin{smallmatrix}\cdot&\cdot&\cdot\\ &\cdot&\cdot\\ & &\cdot\end{smallmatrix}$ or 6 independently-specifiable components, which at (159) we have been motivated to identify in a specific way with the six components of a pair of 3-vectors. The statement

$$F^{\nu\mu} = -F^{\mu\nu} \quad : \quad \text{more compactly} \quad \mathbb{F}^{\mathsf{T}} = -\,\mathbb{F} \tag{160}$$

evidently holds *at every spacetime point*, and will play a central role in our work henceforth.

It follows by inspection from results now in hand that the sourceless field equations (157.2) can be formulated

$$\partial_\mu G^{\mu\nu} = 0$$

with

$$\mathbb{G} \equiv \|G^{\mu\nu}\| = \mathbb{A}(-\boldsymbol{B}, \boldsymbol{E}) = \begin{pmatrix} 0 & B_1 & B_2 & B_3 \\ -B_1 & 0 & -E_3 & E_2 \\ -B_2 & E_3 & 0 & -E_1 \\ -B_3 & -E_2 & E_1 & 0 \end{pmatrix} \tag{161}$$

...but with this step we have acquired an obligation to *develop the sense in which* $\mathbb{G}$ *is a "natural companion" of* $\mathbb{F}$. To that end:

Let the *Levi-Civita symbol* $\epsilon_{\mu\nu\rho\sigma}$ be defined

$$\epsilon_{\mu\nu\rho\sigma} \equiv \begin{cases} +1 & \text{if } (\mu\nu\rho\sigma) \text{ is an } \underline{\text{even}} \text{ permutation of } (0123) \\ -1 & \text{if } (\mu\nu\rho\sigma) \text{ is an } \underline{\text{odd}} \text{ permutation of } (0123) \\ \ \ 0 & \text{otherwise} \end{cases}$$

and let quantities $F^{\star}_{\mu\nu}$ be constructed with its aid:

$$F^{\star}_{\mu\nu} \equiv \tfrac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^{\alpha\beta} \quad \text{where } \sum_{\rho,\,\sigma} \text{ is understood} \tag{162}$$

By computation we readily establish that

$$\begin{aligned} \mathbb{F}^{\star} \equiv \|F^{\star}_{\mu\nu}\| &= \begin{pmatrix} 0 & F^{23} & -F^{13} & F^{12} \\ & 0 & F^{03} & -F^{02} \\ & (-) & 0 & F^{01} \\ & & & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & -E_3 & E_2 \\ B_2 & E_3 & 0 & -E_1 \\ B_3 & -E_2 & E_1 & 0 \end{pmatrix} = \mathbb{A}(\boldsymbol{B},\boldsymbol{E}) \end{aligned}$$

which would become $\mathbb{G}$ if we could change the sign of the B-entries, and this is readily accomplished: multiply $\mathbb{A}(\boldsymbol{B},\boldsymbol{E})$ by

$$g \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{163}$$

on the right (this leaves the 0^{th} column unchanged, but changes the sign of the 1^{st}, 2^{nd} and 3^{rd} columns), and again by another factor of g on the left (this leaves the 0^{th} row unchanged, but changes the sign of the 1^{st}, 2^{nd} and 3^{rd} rows, the 1^{st}, 2^{nd} and 3^{rd} elements of which have now been restored to their *original* signs). We are led thus to $g\mathbb{A}(\boldsymbol{B},\boldsymbol{E})g = \mathbb{A}(-\boldsymbol{B},\boldsymbol{E})$ which—because

$$g = \begin{cases} g^{\mathsf{T}} & : \ g \text{ is its own transpose } (i.e., \text{ is } symmetric) \\ g^{-1} & : \ g \text{ is its own inverse} \end{cases} \tag{164}$$

—can also be expressed $\mathbb{A}(\boldsymbol{B},\boldsymbol{E}) = g\mathbb{A}(-\boldsymbol{B},\boldsymbol{E})g$. In short,[80]

$$\mathbb{F}^{\star} = g\,\mathbb{G}g^{\mathsf{T}} \quad \text{equivalently} \quad \mathbb{G} = g^{-1}\mathbb{F}^{\star}(g^{-1})^{\mathsf{T}} \tag{165}$$

---

[80] PROBLEM 35.

Let the elements of $\mathcal{g}$ be called $g_{\mu\nu}$, and the elements of $\mathcal{g}^{-1}$ (though they happen to be numerically identical to the elements of $\mathcal{g}$) be called $g^{\mu\nu}$:

$$\mathcal{g} \equiv \|g_{\mu\nu}\| \quad \text{and} \quad \mathcal{g}^{-1} \equiv \|g^{\mu\nu}\| \quad \Rightarrow \quad g^{\mu\alpha}g_{\alpha\nu} = \delta^{\mu}{}_{\nu} \equiv \begin{cases} 1 & \text{if } \mu = \nu \\ 0 & \text{if } \mu \neq \nu \end{cases}$$

We then have

$$F^{\bigstar}_{\mu\nu} = g_{\mu\alpha}g_{\nu\beta}G^{\alpha\beta} \quad \text{or equivalently} \quad G^{\mu\nu} = g^{\mu\alpha}g^{\nu\beta}F^{\bigstar}_{\alpha\beta}$$

To summarize: we have

$$F^{\mu\nu} \xrightarrow[\bigstar]{\qquad\qquad} F^{\bigstar}_{\mu\nu} \xrightarrow{\text{lift indices with } \mathcal{g}^{-1}} F^{\bigstar\mu\nu} = G^{\mu\nu}$$

which in $(\boldsymbol{E}, \boldsymbol{B})$-notation reads

$$\mathbb{F} = \mathbb{A}(\boldsymbol{E}, \boldsymbol{B}) \longrightarrow \mathbb{A}(\boldsymbol{B}, \boldsymbol{E}) \longrightarrow \mathbb{A}(-\boldsymbol{B}, \boldsymbol{E}) = \mathbb{G}$$

Repetition of the process gives

$$\mathbb{G} = \mathbb{A}(-\boldsymbol{B}, \boldsymbol{E}) \longrightarrow \mathbb{A}(\boldsymbol{E}, -\boldsymbol{B}) \longrightarrow \mathbb{A}(-\boldsymbol{E}, -\boldsymbol{B}) = -\mathbb{F}$$

$G^{\mu\nu}$ is said to be the "dual" of $F^{\mu\nu}$, and the process $F^{\mu\nu} \longrightarrow G^{\mu\nu}$ is called "dualization;" it amounts to a kind of "*rotation* in $(\boldsymbol{E}, \boldsymbol{B})$-space," in the sense illustrated below:

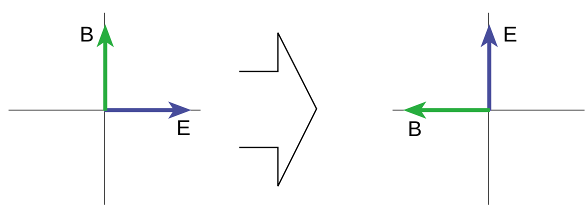

FIGURE 45: *The "rotational" effect of "dualization" on* $\boldsymbol{E}$ *and* $\boldsymbol{B}$.

Preceding remarks lend precise support and meaning to the claim that $F^{\mu\nu}$ and $G^{\mu\nu}$ are "natural companions," and very closely related.

We shall—as above, but more generally (and for the good tensor-theoretic reasons that will soon emerge) *use* $g^{\mu\nu}$ *and* $g_{\mu\nu}$ *to raise and lower—in short, to "manipulate"—indices*, writing (for example)[81]

$$\partial^{\mu} = g^{\mu\alpha}\partial_{\alpha}\,, \quad \partial_{\mu} = g_{\mu\alpha}\partial^{\alpha}$$
$$j^{\mu} = g^{\mu\alpha}j_{\alpha}\,, \quad j_{\mu} = g_{\mu\alpha}j^{\alpha}$$
$$F_{\mu\nu} = g_{\mu\alpha}F^{\alpha}{}_{\nu} = g_{\mu\alpha}g_{\nu\beta}F^{\alpha\beta}$$

[81] PROBLEM 36.

We are placed thus in position to notice that the sourceless Maxwell equations (157.2) can be formulated[82]

$$\left.\begin{aligned}\partial_1 F_{23} + \partial_2 F_{31} + \partial_3 F_{12} &= 0\\ \partial_0 F_{23} + \partial_2 F_{30} + \partial_3 F_{02} &= 0\\ \partial_0 F_{13} + \partial_1 F_{30} + \partial_3 F_{01} &= 0\\ \partial_0 F_{12} + \partial_1 F_{20} + \partial_2 F_{01} &= 0\end{aligned}\right\}\tag{166.1}$$

where the *sums over cyclic permutations* are sometimes called "windmill sums." More compactly, we have[83]

$$\epsilon_{\mu\alpha\rho\sigma}\partial^\alpha F^{\rho\sigma} = 0 \tag{166.2}$$

There is <u>no new physics</u> in the material presented thus far: our work has been merely reformulational, notational—old wine in new bottles. Proceeding in response mainly to the *linearity* of Maxwell's equations, we have allowed ourselves to play linear-algebraic and notational games intended to <u>maximize the formal symmetry/simplicity</u> of Maxwell's equations . . . so that the transformation-theoretic problem which is our real concern can be posed in the simplest possible terms. Maxwell himself[84] construed the electromagnetic field to involve a *pair of 3-vector fields:* $\boldsymbol{E}$ *and* $\boldsymbol{B}$. We have seen, however, that

- one can construe the components of $\boldsymbol{E}$ and $\boldsymbol{B}$ to be the accidentally distinguished names given to the six independently-specifiable non-zero components of an *antisymmetric tensor*[85]*field* $F^{\mu\nu}$. The field equations then read
$$\partial_\mu F^{\mu\nu} = \tfrac{1}{c}j^\nu \quad\text{and}\quad \epsilon_{\mu\alpha\rho\sigma}\partial^\alpha F^{\rho\sigma} = 0 \tag{167}$$
provided the $g^{\alpha\beta}$ that enter into the definition $\partial^\alpha \equiv g^{\alpha\beta}\partial_\beta$ are given by (163). Alternatively . . .
- one can adopt the view that the electromagnetic field to involves a *pair* of antisymmetric tensor fields $F^{\mu\nu}$ and $G^{\mu\nu}$ which are constrained to satisfy not only the field equations
$$\partial_\mu F^{\mu\nu} = \tfrac{1}{c}j^\nu \quad\text{and}\quad \partial_\mu G^{\mu\nu} = 0 \tag{168.1}$$
but also the algebraic condition
$$G^{\mu\nu} = \tfrac{1}{2}g^{\mu\alpha}g^{\nu\beta}\epsilon_{\alpha\beta\rho\sigma}F^{\rho\sigma} \tag{168.2}$$
Here again, the "index manipulators" $g_{\mu\nu}$ and $g^{\mu\nu}$ must be assigned the specific meanings implicit in (163).

---

[82] PROBLEM 37.

[83] PROBLEM 38.

[84] Here I take some liberty with the complicated historical facts of the matter: see again the fragmentary essay[77] cited earlier.

[85] For the moment "tensor" simply means "doubly indexed."

It will emerge that Lorentz' question (page 107), if phrased in the terms natural to either of those descriptions of Maxwellian electrodynamics, virtually "answers itself." But to see how this comes about one must possess a command of the basic elements of tensor analysis—a subject with which Minkowski (mathematician that he was) enjoyed a familiarity not shared by any of his electrodynamical predecessors or contemporaries.[86]

2. Introduction to the algebra and calculus of tensors. Let P be a *point in an N-dimensional manifold* $\mathcal{M}$.[87] Let $(x^1, x^2, \ldots, x^N)$ be coordinates assigned to P by a coordinate system $\mathcal{X}$ inscribed on a neighborhood[88] containing P, and

[86] Though (167) and (168) serve optimally my immediate purposes, the reader should be aware that there exist also many alternative formulations of the Maxwellian theory, and that these may afford advantages in specialized contexts. We will have much to say about the formalism that proceeds from writing

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

and considering the fundamental object of electrodynamic analysis to be a *single 4-vector field*. Alternatively, one might construct and study the "6-vector"

$$f = \begin{pmatrix} f^1 \\ f^2 \\ f^3 \\ f^4 \\ f^5 \\ f^6 \end{pmatrix} \equiv \begin{pmatrix} E_1 \\ E_2 \\ E_3 \\ B_1 \\ B_2 \\ B_3 \end{pmatrix}$$

(see §26 in Arnold Sommerfeld's *Electrodynamics* (English translation 1964) or my *Classical Field Theory* (1999), Chapter 2, pages 4–6). Or one might consider electrodynamics to be concerned with the properties of a *single* <u>*complex*</u> *3-vector*

$$\boldsymbol{V} \equiv \boldsymbol{E} + i\boldsymbol{B}$$

(see Appendix B in my "On some recent electrodynamical work by Thomas Wieting" (2001)). And there exist yet many other formalisms. Maxwell himself gave passing attention to a "quaternionic" formulation of his theory.

[87] Think "surface of a sphere," "surface of a torus," *etc.* or of their higher-dimensional counterparts. Or of N-dimensional Euclidean space itself. Or—as soon as you can—4-dimensional spacetime. I intend to proceed quite informally, and to defer questions of the nature "What *is* a manifold?" until such time as we are able to look back and ask "What properties *should* we fold into our definitions? What did we need to make our arguments work?"

[88] I say "neighborhood" because it may happen that every coordinate system inscribed on $\mathcal{M}$ necessarily displays one or more singularities (think of the longitude of the North Pole). It is our announced intention to stay away from such points.

let $(x^1, x^2, \ldots, x^N)$ be the coordinates assigned to that same point by a second coordinate system $\mathcal{X}$. We seek to develop rules according to which objects defined in the neighborhood of P respond to coordinate transformations: $\mathcal{X} \rightarrow \mathcal{X}$.

The statement that "$\phi(x)$ transforms as a *scalar field*" carries this familiar meaning:

$$\phi(x) \longrightarrow \phi(x) \equiv \phi(x(x)) \tag{169}$$

Here and henceforth: $x(x)$ alludes to the functional statements

$$x^m = x^m(x^1, x^2, \ldots, x^N) \quad : \quad m = 1, 2, \ldots N \tag{170}$$

that describe how $\mathcal{X}$ and $\mathcal{X}$ are, in the instance at hand, specifically related. How do the partial *derivatives* of ϕ transform? By calculus

$$\frac{\partial \phi}{\partial x^m} = \frac{\partial x^a}{\partial x^m}\frac{\partial \phi}{\partial x^a} \tag{171.1}$$

where (as always) $\sum\limits_a$ is understood. Looking to the 2nd derivatives, we have

$$\frac{\partial^2 \phi}{\partial x^m \partial x^n} = \frac{\partial x^a}{\partial x^m}\frac{\partial x^b}{\partial x^n}\frac{\partial^2 \phi}{\partial x^a \partial x^b} + \frac{\partial^2 x^a}{\partial x^m \partial x^n}\frac{\partial \phi}{\partial x^a} \tag{171.2}$$

Et cetera. Such are the "objects" we encounter in routine work, and the transformation rules which we want to be able to manipulate in a simple manner.

The quantities $\partial x^a / \partial x^m$ arise directly and exclusively from the equations (170) that describe $\mathcal{X} \leftarrow \mathcal{X}$. They constitute the elements of the "transformation matrix"

$$\begin{aligned} \mathbb{W} &\equiv \|W^n{}_m\| \\ W^n{}_m &\equiv \partial x^n / \partial x^m \end{aligned} \tag{172.1}$$

—the value of which will in general vary from point to point. Function theory teaches us that the coordinate transformation will be *invertible* (*i.e.*, that we can proceed from $x^n = x^n(x)$ to equations of the form $x^n = x^n(x)$) if and only if $\mathbb{W}$ is non-singular: $\det \mathbb{W} \neq 0$, which we always assume to be the case (in the neighborhood of P). The inverse $\mathcal{X} \rightarrow \mathcal{X}$ of $\mathcal{X} \leftarrow \mathcal{X}$ gives rise to

$$\begin{aligned} \mathbb{M} &\equiv \|M^m{}_n\| \\ M^m{}_n &\equiv \partial x^m / \partial x^n \end{aligned} \tag{172.2}$$

It is important to notice that

$$\mathbb{W}\mathbb{M} = \left\| \sum_a \frac{\partial x^n}{\partial x^a}\frac{\partial x^a}{\partial x^m} \right\| = \left\| \frac{\partial x^n}{\partial x^m} \right\| = \left\| \delta^n{}_m \right\| = \mathbb{I} \tag{173}$$

i.e., that the matrices $\mathbb{M}$ and $\mathbb{W}$ are inverses of each other.

Objects $X^{m_1\dots m_r}{}_{n_1\dots n_s}$ are said to comprise the "components of a (mixed) tensor of contravariant rank r and covariant rank s if and only if they respond to $\mathcal{X} \rightarrow \mathcal{X}$ by the following multilinear rule:

$$\begin{aligned}
&X^{m_1\dots m_r}{}_{n_1\dots n_s}\\
&\downarrow\\
&X^{m_1\dots m_r}{}_{n_1\dots n_s} = M^{m_1}{}_{a_1}\cdots M^{m_r}{}_{a_r}W^{b_1}{}_{n_1}\cdots W^{b_s}{}_{n_s}X^{a_1\dots a_r}{}_{b_1\dots b_s}
\end{aligned}\tag{174}$$

All indices range on $\{1, 2, \dots, N\}$, N is called the "dimension" of the tensor, and summation on repeated indices is (by the "Einstein summation convention") understood. The covariant/contravariant distinction is signaled notationally as a subscript/superscript distinction, and alludes to whether it is $\mathbb{W}$ or $\mathbb{M}$ that transports the components in question "across the street, from the $\mathcal{X}$-side to the $\mathcal{X}$-side."

If

$$X^m \longrightarrow X^m = M^m{}_a X^a$$

then the X^m are said to be "components of a contravariant *vector*." Coordinate differentials provide the classic prototype:

$$dx^m \longrightarrow dx^m = \sum_a \frac{\partial x^m}{\partial x^a} dx^a \tag{175}$$

If, on the other hand,

$$X_n \longrightarrow X_n = W^b{}_n X_b$$

then the X_n are said to be "components of a *co*variant vector." Here the first partials $\phi_{,n} \equiv \partial_n \phi$ of a scalar field (components of the **gradient**) provide the classic prototype:

$$\phi_{,n} \longrightarrow \phi_{,n} = \sum_b \phi_{,b} \frac{\partial x^b}{\partial x^n} \tag{176}$$

That was the lesson of (171.1).

Look, however, to the lesson of (171.2), where we found that

$$\phi_{,mn} \longrightarrow \phi_{,mn} = \sum_b \phi_{,ab} \frac{\partial x^a}{\partial x^m}\frac{\partial x^b}{\partial x^n} + \text{extraneous term}$$

The intrusion of the "extraneous term" is typical of the *differential calculus of tensors*, and arises from an elementary circumstance: hitting

$$X^m{}_n = M^m{}_a W^b{}_n X^a{}_b \quad \text{(say)}$$

with $\boldsymbol{\partial}_p = W^q{}_p \partial_q$ gives

$$\begin{aligned}\boldsymbol{X}^m{}_{n,p} &= M^m{}_a W^b{}_n X^a{}_{b,q} W^q{}_p + W^q{}_p \frac{\partial (M^m{}_a W^b{}_n)}{\partial x^q} X^a{}_b \\ &= (\text{term with covariant rank increased by one}) \\ &\qquad\qquad + (\text{extraneous term})\end{aligned}$$

The "extraneous term" vanishes if the M's and W's are *constant*; *i.e.*, if the functions $\boldsymbol{x}^n(x)$ depend at most *linearly* upon their arguments $\boldsymbol{x}^n = M^n{}_a x^a + \xi^a$. And in a small number of (electrodynamically important!) cases the extraneous terms *cancel* when derivatives are *combined* in certain ways ... as we will soon have occasion to see. But in general, effective management of the extraneous term must await the introduction of some powerful new ideas—ideas that belong not to the algebra of tensors (my present concern) but to the calculus of tensors. For the moment I must be content to emphasize that, on the basis of evidence now in hand,

Not every multiply-indexed object transforms tensorially!

In particular, the x^n themselves do not transform tensorially except in the linear case $\boldsymbol{x}^n = M^n{}_a x^a$.

A conceptual point of major importance: the $X^{m_1 \cdots m_r}{}_{n_1 \ldots n_s}$ *refer* to a tensor, but do not themselves *comprise* the tensor: they are the components of the tensor X with respect to the coordinate system $\mathfrak{X}$, and collectively serve to *describe* X. Similarly $\boldsymbol{X}^{m_1 \cdots m_r}{}_{n_1 \ldots n_s}$ with respect to $\boldsymbol{\mathfrak{X}}$. The tensor itself is a coordinate-independent object that lives "behind the scene." The situation is illustrated in Figure 46.

To lend substance to a remark made near the top of the page: Let X_m transform as a covariant vector. Look to the transformation properties of $X_{m,n}$ and obtain

$$\boldsymbol{X}_{m,n} = W^a{}_m W^b{}_n X_{a,b} + \underbrace{\frac{\partial^2 x^a}{\partial \boldsymbol{x}^n \partial \boldsymbol{x}^m} X_a}_{\text{extraneous term, therefore non-tensorial}}$$

Now construct $A_{mn} \equiv X_{m,n} - X_{n,m} = -A_{nm}$ and obtain

$$\boldsymbol{A}_{mn} = W^a{}_m W^b{}_n A_{ab} \quad \text{because the \underline{extraneous terms cancel}}$$

We conclude that the antisymmetric construction A_{mn} (which we might call the **curl** of the covariant vector field $X_m(x)$) *does*—"accidentally"—transform tensorially.

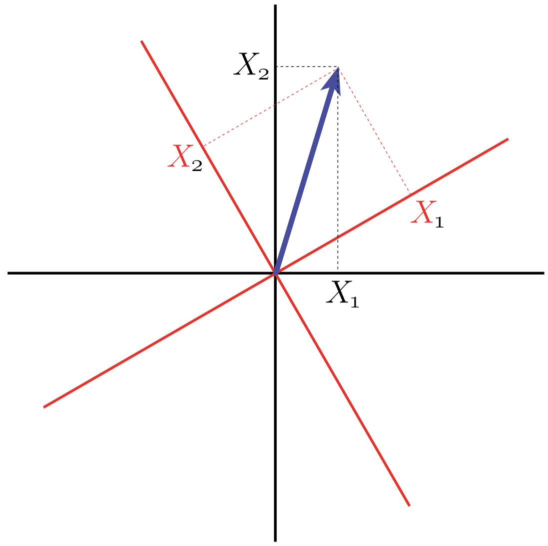

FIGURE 46: *The X_m serve to describe the blue arrow with respect to the black coordinate system $\mathfrak{X}$, as the X_m serve to describe the blue arrow with respect to the red coordinate system $\mathfrak{X}$. But neither X_m nor X_m will be confused with the blue arrow itself: to do so would be to confuse descriptors with the thing described. So it is with tensors in general. Tensor analysis is concerned with* relationships among alternative descriptors, not with "things in themselves."

The following points are elementary, but fundamental to *applications* of the tensor concept:

1) If the components $X^{\cdots}{}_{\ldots}$ of a tensor (all) vanish one coordinate system, then they vanish in *all* coordinate systems—this by the *homogeneity* of the defining statement (174).
2) Tensors can be added/subtracted if and only if $X^{\cdots}{}_{\ldots}$ and $Y^{\cdots}{}_{\ldots}$ are of the same covariant/contravariant rank and dimension. Constructions of (say) the form $A^m + B_m$ "come unstuck" when transformed; for that same reason, statements of (say) the form $A^m = B_m$—while they may be valid in some given coordinate system—do not entail $A^m = B_m$. But . . .
3) If $X^{\cdots}{}_{\ldots}$ and $Y^{\cdots}{}_{\ldots}$ are of the *same* rank and dimension, then
$$X^{\cdots}{}_{\ldots} = Y^{\cdots}{}_{\ldots} \quad \Longrightarrow X^{\cdots}{}_{\ldots} = Y^{\cdots}{}_{\ldots}$$
It is, in fact, because of the remarkable transformational stability of tensorial equations that we study this subject, and try to formulate our physics in tensorial terms.
4) If $X^{\cdots}{}_{\ldots}$ and $Y^{\cdots}{}_{\ldots}$ are co-dimensional tensors of ranks $\{r', s'\}$ and $\{r'', s''\}$ then their product $X^{\cdots}{}_{\ldots}Y^{\cdots}{}_{\ldots}$ is tensorial with rank $\{r' + r'', s' + s''\}$: tensors of the same dimension can be multiplied irrespective of their ranks.

If $X^{\cdots}{}_{\cdots}$ is tensorial of rank $\{r, s\}$ then a the operation of

> CONTRACTION: Set a superscript equal to a subscript, and add

yields components of a tensor of rank $\{r-1, s-1\}$. The mechanism is exposed most simply by example: start from (say)

$$X^{jk}{}_\ell = M^j{}_a M^k{}_b W^c{}_\ell X^{ab}{}_c$$

Set (say) $k = \ell$ and obtain

$$\begin{aligned} X^{jk}{}_k &= M^j{}_a M^k{}_b W^c{}_k X^{ab}{}_c \\ &= M^j{}_a \quad \delta^c{}_b \quad X^{ab}{}_c \qquad \text{by} \quad \mathbb{M}\,\mathbb{W} = \mathbb{I} \\ &= M^j{}_a X^{ab}{}_b \end{aligned}$$

according to which $X^j \equiv X^{jk}{}_k$ transforms as a contravariant vector. Similarly, the twice-contracted objects $X^{jk}{}_{jk}$ and $X^{jk}{}_{kj}$ transform as (generally distinct) *invariants*.[89] Mixed tensors of high rank can be singly/multiply contracted in many distinct ways. It is also possible to "contract one tensor into another;" a simple example:

$$A_k B^k \quad : \quad \begin{cases} \text{invariant formed by contracting a covariant} \\ \text{vector into a contravariant vector} \end{cases}$$

The "Kronecker symbol" $\delta^m{}_n$ is a number-valued object[90] with which all readers are familiar. If "transformed tensorially" it gives

$$\begin{aligned} \delta^m{}_n \longrightarrow \delta^m{}_n &= M^m{}_a W^b{}_n \delta^a{}_b \\ &= M^m{}_a W^a{}_n \\ &= \delta^m{}_n \qquad \text{by} \quad \mathbb{M}\,\mathbb{W} = \mathbb{I} \end{aligned}$$

and we are brought to the remarkable conclusion that *the components $\delta^m{}_n$ of the Kronecker tensor have the <u>same numerical values in every coordinate system</u>*. Thus does $\delta^m{}_n$ become what I will call a "universally available object"—to be joined soon by a few others. With this ...

We are placed in position to observe that if the quantities g_{mn} transform as the components of a 2nd rank covariant tensor

$$g_{mn} \longrightarrow g_{mn} = W^a{}_m W^b{}_n g_{ab} \tag{177}$$

---

[89] The "theory of invariants" was a favorite topic among 19th Century mathematicians, and provided the founding fathers of tensor analysis with a source of motivation (see pages 206–211 in E. T. Bell's *The Development of Mathematics* (1945)).

[90] See again the top of page 110.

then

1) the equation $g^{ma}g_{an} = \delta^m{}_n$, if taken as (compare page 110) a definition of the contravariant tensor g^{mn}, makes good coordinate-independent tensor-theoretic sense, and
2) so do the equations

$$\begin{aligned} X^{\cdots m \cdots}_{\cdots\ \ \cdots} &\equiv g^{ma}X^{\cdots\ \ \cdots}_{\cdots\, a\, \cdots} \\ X^{\cdots\ \ \cdots}_{\cdots m \cdots} &\equiv g_{ma}X^{\cdots\, a\, \cdots}_{\cdots\ \ \cdots} \end{aligned}$$

by means of which we have proposed already on page 110 to raise and lower indices.[91] To insure that $g^{ma}X^{\cdots\ \ \cdots}_{\cdots\, a\, \cdots}$ and $g^{am}X^{\cdots\ \ \cdots}_{\cdots\, a\, \cdots}$ are identical we will require that

$$g_{mn} = g_{nm} \quad : \quad \text{implies the symmetry also of } g^{mn}$$

The transformation equation (177) admits—uncharacteristically–of matrix formulation

$$\mathbb{g} \longrightarrow \mathbb{g} = \mathbb{W}^{\mathsf{T}}\mathbb{g}\,\mathbb{W}$$

Taking determinant of both sides, and writing

$$g \equiv \det \mathbb{g}\,, \quad W \equiv \det \mathbb{W} = 1/\det \mathbb{M} = M^{-1}$$

we have

$$g \longrightarrow g = W^2 g \tag{178.1}$$

The statement that $\phi(x)$ transforms as a *scalar density of weight w* carries this meaning:

$$\phi(x) \longrightarrow \phi(x) = W^w \cdot \phi(x(x))$$

We recover (169) in the "weightless" case $w = 0$ (and for arbitrary values of w when it happens that $W = 1$). Evidently

$$g \equiv \det \mathbb{g} \text{ transforms as a scalar density of weight } w = 2 \tag{178.2}$$

The more general statement that $X^{m_1\ldots m_r}{}_{n_1\ldots n_s}$ transforms as a *tensor density of weight w* means that

$$X^{m_1\ldots m_r}{}_{n_1\ldots n_s} = W^w \cdot M^{m_1}{}_{a_1}\cdots M^{m_r}{}_{a_r}W^{b_1}{}_{n_1}\cdots W^{b_s}{}_{n_s}X^{a_1\ldots a_r}{}_{b_1\ldots b_s}$$

We can multiply/contract tensors of dissimilar weight, but must be careful not to try to add them or set them equal. The "tensor/tensor density distinction" becomes significant only in contexts where $W \neq 1$.

Familiarity with the tensor density concept places us in position to consider the tensor-theoretic significance of the Levi-Civita symbol

---

[91] Note, however, that we work now N-dimensionally, and have stripped g_{mn} of its formerly specialized (Lorentzian) construction (163): it has become "generic."

$$\epsilon_{n_1 n_2 \ldots n_N} \equiv \text{sgn}\begin{pmatrix} 1 & 2 & \cdots & N \\ n_1 & n_2 & \cdots & n_N \end{pmatrix}$$

where "sgn" refers to the "signum," which reports (see again page 109) whether $\{n_1, n_2, \ldots, n_N\}$ is an even/odd permutation of $\{1, 2, \ldots, N\}$ or no permutation at all. The tentative assumption that $\epsilon_{n_1 n_2 \ldots n_N}$ transforms as a (totally antisymmetric) tensor density of unspecified weight w

$$\begin{aligned}\epsilon_{n_1 n_2 \ldots n_N} &= W^w \cdot \underbrace{W^{a_1}{}_{n_1} W^{a_2}{}_{n_2} \cdots W^{a_N}{}_{n_N} \epsilon_{a_1 a_2 \ldots a_N}}_{\substack{= \epsilon_{n_1 n_2 \ldots n_N} \det \mathbb{W} \\ \text{by definition of the determinant!}}} \\ &= W^{w+1} \cdot \epsilon_{n_1 n_2 \ldots n_N}\end{aligned}$$

brings us to the remarkable conclusion that *the components of the Levi-Civita tensor will have the <u>same numerical values in every coordinate system</u> provided $\epsilon_{n_1 n_2 \ldots n_N}$ is assumed to transform as a <u>density of weight $w = -1$</u>*. The Levi-Civita tensor thus joins our short list of "universally available objects."[92]

I have remarked that $\epsilon_{n_1 n_2 \ldots n_N}$ is "totally antisymmetric." It is of importance to notice in this connection that—more generally—statements of the forms

$$X^{\cdots m \cdots n \cdots}{}_{\cdots} = \pm X^{\cdots n \cdots m \cdots}{}_{\cdots}$$

and

$$X^{\cdots}{}_{\cdots m \cdots n \cdots} = \pm X^{\cdots}{}_{\cdots n \cdots m \cdots}$$

have tensorial (or coordinate system independent) significance, while symmetry statements of the hybrid form

$$X^{\cdots m \cdots}{}_{\cdots n \cdots} = \pm X^{\cdots n \cdots}{}_{\cdots m \cdots}$$

—while they might be valid in some particular coordinate system—"become unstuck" when transformed. Note also that

$$X^{mn} = \tfrac{1}{2}(X^{mn} + X^{nm}) + \tfrac{1}{2}(X^{mn} - X^{nm})$$

serves to resolve X^{mn} tensorially into its *symmetric and antisymmetric parts*.[93]

---

[92] The (weightless) "metric tensor" g_{mn} is *not* "universally available," but must be introduced "by hand." In contexts where g_{mn} *is* available (has been introduced to facilitate index manipulation) it becomes natural to construct

$$\sqrt{g}\,\epsilon_{n_1 n_2 \ldots n_N} \quad : \quad \underline{\text{weightless}} \text{ totally antisymmetric tensor}$$

—the values of which range on $\{0, \pm\sqrt{g}\}$ in all coordinate systems.

[93] PROBLEM 39.

We have now in our possession a command of **tensor algebra** which is sufficient to serve our immediate needs, but must sharpen our command of the **differential calculus of tensors**. This is a more intricate subject, but one into which—surprisingly—we need not enter very deeply to acquire the tools needed to achieve our electrodynamical objectives. I will be concerned mainly with the development of a short list of "accidentally tensorial derivative constructions,"[94] and will glance only cursorily at what might be called the "non-accidental aspects" of the tensor calculus.

CATALOG OF ACCIDENTALLY TENSORIAL DERIVATIVE CONSTRUCTIONS

1. We established already at (171.1) that if ϕ transforms as a weightless scalar field then the components of the *gradient* of ϕ

$$\partial_m\phi \text{ transform tensorially} \tag{179.1}$$

2. And we observed on page 115 that if X_m transforms as a weightless covariant vector field then the components of the *curl* of X_m transform tensorially.

$$\partial_n X_m - \partial_m X_n \text{ transform tensorially} \tag{179.2}$$

3. If X_{jk} is a weightless tensor field, how do the $\partial_i X_{jk}$ transform? Immediately

$$\begin{aligned}\partial_i X_{jk} &= W^b{}_j W^c{}_k \cdot W^a{}_i\,\partial_a X_{bc} + X_{bc}\partial_i\big\{W^b{}_j W^c{}_k\big\}\\ &= W^a{}_i W^b{}_j W^c{}_k\,\partial_a X_{bc} + \underbrace{X_{bc}\Big\{\frac{\partial^2 x^b}{\partial x^i\partial x^j}\frac{\partial x^c}{\partial x^k} + \frac{\partial x^b}{\partial x^j}\frac{\partial^2 x^c}{\partial x^k\partial x^i}\Big\}}_{\text{extraneous term}}\end{aligned}$$

so $\partial_i X_{jk}$ transforms tensorially only under such circumstances as cause the "extraneous term" to vanish: this happens when $\mathcal{X} \to \mathcal{X}$ is "affine;" *i.e.*, when the $\mathbb{W}$-matrix is x-independent. Notice, however, that we now have

$$\begin{aligned}\partial_i X_{jk} + \partial_j X_{ki} + \partial_k X_{ij} = W^a{}_i W^b{}_j W^c{}_k(\partial_a X_{bc} + \partial_a X_{bc} + \partial_a X_{bc})&\\ + X_{bc}\Big\{\frac{\partial^2 x^b}{\partial x^i\partial x^j}\frac{\partial x^c}{\partial x^k} + \frac{\partial x^b}{\partial x^j}\frac{\partial^2 x^c}{\partial x^k\partial x^i}&\\ + \frac{\partial^2 x^b}{\partial x^j\partial x^k}\frac{\partial x^c}{\partial x^i} + \frac{\partial x^b}{\partial x^k}\frac{\partial^2 x^c}{\partial x^i\partial x^j}&\\ + \frac{\partial^2 x^b}{\partial x^k\partial x^i}\frac{\partial x^c}{\partial x^j} + \frac{\partial x^b}{\partial x^i}\frac{\partial^2 x^c}{\partial x^j\partial x^k}\Big\}&\end{aligned}$$

in which $\{\text{etc.}\}$ is bc-symmetric; if X_{bc} were *anti*symmetric the extraneous term would therefore drop away. We conclude that if X_{jk} is an *antisymmetric* weightless covariant tensor field then the components of the windmill sum

$$\partial_i X_{jk} + \partial_j X_{ki} + \partial_k X_{ij} \text{ transform tensorially} \tag{179.3}$$

---

[94] The possibility and electrodynamical utility of such a list was brought first to my attention when, as a student, I happened upon the discussion which appears on pages 22–24 of E. Schrödinger's *Space-time Structure* (1954). This elegant little volume (which runs to only 119 pages) provides physicists with an elegantly succinct introduction to tensor analysis. I recommend it to your attention.

4. If X^m is a vector *density of unspecified weight* w how does $\partial_m X^m$ transform? Immediately

$$\partial_m X^m = W^w \cdot \underbrace{M^m{}_a \partial_m}_{\partial_a} X^a + X^a \partial_m \{W^w \cdot M^m{}_a\}$$

$$= W^w \cdot \partial_a X^a + X^a \Big\{ W^w \frac{\partial}{\partial x^m}\frac{\partial x^m}{\partial x^a} + wW^{w-1}\frac{\partial W}{\partial x^a}\Big\}$$

An important LEMMA[95] asserts that

$$\begin{aligned}\frac{\partial}{\partial x^m}\frac{\partial x^m}{\partial x^a} &= \frac{\partial}{\partial x^a}\log\det\left\|\frac{\partial x^m}{\partial x^n}\right\| \\ &= \partial_a \log M = -\partial_a \log W \\ &= -W^{-1}\partial_a W\end{aligned}$$

so

$$= W^w \cdot \partial_a X^a + \underbrace{X^a (w-1) W^{w-1}\frac{\partial W}{\partial x^a}}_{\text{extraneous term}}$$

The extraneous term vanishes (for all w) when $\mathfrak{X} \to \mathfrak{X}$ has the property that W is x-independent,[96] and it vanishes unrestrictedly if $w = 1$. We conclude that if X^m is a contravariant vector density of unit weight then its *divergence*

$$\partial_m X^m \text{ transforms tensorially (by invariance)} \tag{179.4}$$

5. If X^{mn} is a vector density of unspecified weight w how does $\partial_m X^{mn}$ transform? Immediately

$$\partial_m X^{mn} = \underbrace{W^w \cdot M^m{}_a M^n{}_b (W^c{}_m \partial_c X^{ab})}_{= W^w \cdot M^n{}_b\, \partial_a X^{ab} \quad \text{by } M^m{}_a W^c{}_m = \delta^c{}_a} + \underbrace{X^{ab}\partial_m (W^w \cdot M^m{}_a M^n{}_b)}_{\text{extraneous term}}$$

The extraneous term can be developed

$$X^{ab}\Big\{ M^n{}_b w W^{w-1}(M^m{}_a \partial_m) W + W^w \Big[M^n{}_b \underbrace{\partial_m M^m{}_a}_{= -W^{-1}\partial_a W \quad \text{by the LEMMA}} + (M^m{}_a \partial_m) M^n{}_b \Big] \Big\}$$

so by $M^m{}_a \partial_m = \partial_a$ we have

$$\text{extraneous term} = X^{ab}\Big\{ M^n{}_b (w-1) W^{w-1}\partial_a W + W^w \frac{\partial^2 x^n}{\partial x^a \partial x^b}\Big\}$$

---

[95] For the interesting but somewhat intricate proof, see CLASSICAL DYNAMICS (1964/65), Chapter 2, page 49.
[96] This is weaker than the requirement that $\mathbb{W}$ be x-independent.

The second partial is ab-symmetric, and makes no net contribution if we assume X^{ab} to be ab-*anti*symmetric. The surviving fragment of the extraneous term vanishes (all w) if W is constant, and vanishes unrestrictedly if $w = 1$. We are brought thus to the conclusion that if X^{mn} is an antisymmetric density of unit weight then

$$\partial_m X^{mn} \text{ transforms tensorially} \tag{179.5}$$

"Generalized divergences" $\partial_m X^{mn_1\cdots n_p}$ yield to a similar analysis, but will not be needed.

6. Taking (179.5) and (179.4) in combination we find that under those same conditons (*i.e.*, if X^{mn} is an antisymmetric density of unit weight) then

$$\partial_m \partial_n X^{mn} \text{ transforms tensorially}$$

but this is hardly news: the postulated antisymmetry fo X^{mn} combines with the manifest symmetry of $\partial_m \partial_n$ to give

$$\partial_m \partial_n X^{mn} = 0 \text{ automatically}$$

The evidence now in hand suggests—accurately—that *antisymmetry has a marvelous power to dispose of what we have called "extraneous terms."* The calculus of antisymmetric tensors is in fact much easier than the calculus of tensors-in-general, and is known as the **exterior calculus**. That independently developed sub-branch of the tensor calculus supports not only a differential calculus of tensors but also—uniquely—an integral calculus, which radiates from the theory of determinants (which are antisymmetry-infested) and in which the fundamental statement is a vast generalization of Stokes' theorem.[97]

> REMARK: Readers will be placed at no immediate disadvantage if, on a first reading, they skip the following descriptive comments, which have been inserted only in the interest of a kind of "sketchy completeness" and which refer to material which is—remarkably!—inessential to our electrodynamical progress (though indispensable in many other physical contexts).

In more general (antisymmetry-free) contexts one deals with the non-tensoriality of $\partial_m X^{\cdots}{}_{\cdots}$ by <u>modifying the concept of differentiation</u>, writing (for example)

$$\begin{aligned} D_j X_k &\equiv \underbrace{W^b{}_j W^c{}_k\, \partial_b X_c}_{\text{tensorial transform of } \partial_j X_k} \\ &\equiv \text{components of the \textbf{covariant derivative} of } X_k \end{aligned}$$

---

[97] See again the MATHEMATICAL DIGRESSION that culminates on page 50. A fairly complete and detailed account of the exterior calculus can be found in "Electrodynamical applications of the exterior calculus" (1996).

where by computation

$$= \partial_j X_k - X_i \Gamma^i{}_{jk}$$

with

$$\Gamma^i{}_{jk} \equiv \frac{\partial x^i}{\partial x^p}\frac{\partial^2 x^p}{\partial x^j \partial x^k}$$

By extension of the notational convention $X_{k,\,j} \equiv \partial_j X_k$ one writes $X_{k;\,j} \equiv D_j X_k$. It is a clear that $X_{j;\,k}$—since created by "tensorial continuation" from the "seed" $\partial_j X_k$—transforms tensorially, and that it has something to do with familiar differentiation (*is* differentiation, but with built-in compensation for the familiar "extraneous term," and reduces to ordinary differentiation in the root coordinate system $\mathcal{X}$). The quantities $\Gamma^i{}_{jk}$ turn out not to transform tensorially, but by the rule

$$= M^i{}_a W^b{}_j W^c{}_k \Gamma^a{}_{bc} + \frac{\partial x^i}{\partial x^p}\frac{\partial^2 x^p}{\partial x^j \partial x^k}$$

characteristic of "affine connections." Finally, one gives up the assumption that there exists a coordinate system (the $\mathcal{X}$-system of prior discussion) in which D_j and ∂_j have coincident (*i.e.*, in which $\Gamma^i{}_{jk}$ vanishes globally). The affine connection $\Gamma^i{}_{jk}(x)$ becomes an object that we are free to *deposit* on the manifold $\mathcal{M}$, to create an "affinely connected manifold"...just as by deposition of $g_{ij}(x)$ we create a "metrically connected manifold." But when we do *both* things[98] a compatability condition arises, for we expect

- index manipulation followed by covariant differentiation, and
- covariant differentiation followed by index manipulation

to yield the same result. This is readily shown to entail $g_{ij;k} = 0$, which in turn entails

$$\Gamma^i{}_{jk} = \tfrac{1}{2} g^{ia}\Big(\frac{\partial g_{aj}}{\partial x^k} + \frac{\partial g_{ak}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^a}\Big)$$

The affine connection has become implicit in the metric connection—it has become the "Christoffel connection," which plays a central role in Riemannian geometry and its applications (general relativity): down the road just a short way lies the Riemann-Christoffel curvature tensor

$$R^m{}_{nij} = \frac{\partial \Gamma^m{}_{nj}}{\partial x^i} - \frac{\partial \Gamma^m{}_{ni}}{\partial x^j} + \Gamma^m{}_{ai}\Gamma^a{}_{nj} - \Gamma^m{}_{aj}\Gamma^a{}_{ni}$$

which enters into statements such as the following

$$X_{n;ij} - X_{n;ji} = X_a R^a{}_{nij}$$

which describes the typical *inequality of crossed covariant derivatives*. The "covariant derivative" was invented by Elwin Christoffel (1829–1900) in 1869.

[98] Notice that we *need* both if we want to construct such things as the

$$\text{covariant Laplacian of } \phi \equiv g^{mn}\phi_{;mn}$$

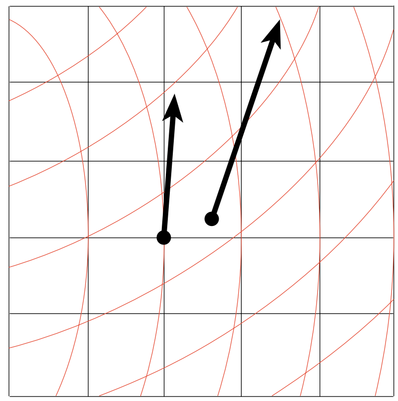

Figure 47: *Any attempt to construct a transformationally coherent theory of differentiation by comparing such neighboring vectors is doomed unless* $\mathcal{X} \to \mathcal{X}$ *gives rise to a transformation matrix that is* constant *on the neighborhood.*

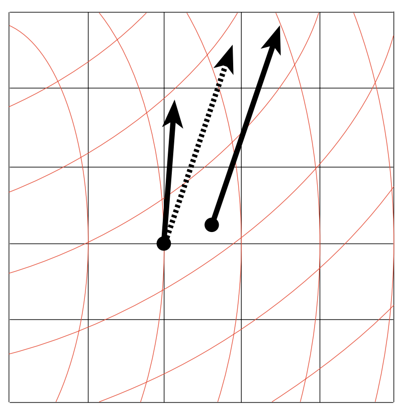

Figure 48: *The problem just noted is resolved if one compares one vector* with the local parallel transport *of the other—a "stand-in" rooted to the same point as the original vector. For then only a single transformation matrix enters into the discussion.*

Sharp insight into the meaning of the covariant derivative was provided in 1917 by Levi-Civita,[99] who pointed out that when one works from Figure 47 one *cannot realistically expect* to obtain a transformationally sensible result, for the

[99] The fundamental importance of Levi-Civita's idea was immediately appreciated and broadcast by Hermann Weyl. See §14 in his classic *Space, Time & Matter* (4th edition 1920, the English translation of which has been reprinted by Dover).

transformation matrices $\mathbb{W}(x)$ and $\mathbb{W}(x+dx)$ that act upon (say) $X_m(x)$ and $X_m(x+dx)$ are, in general, distinct. Levi-Civita observed that a workable procedure does, however, result if one looks not $X_m(x+dx) - X_m(x)$ but to $\mathfrak{X}_m(x) - X_m(x)$, where

$\mathfrak{X}_m(x)$ results from **parallel transport**
of $X_m(x+dx)$ from $x+dx$ back to x

He endowed the intuitive concept "parallel transport" (Figure 48) with a precise (natural) meaning, and immediately recovered the standard theory of covariant differentiation. But he obtained also much else: he showed, for example, that "geodesics" can be considered to arise not as "shortest" curves—curves produced by minimization of arc length

$$\int ds \quad \text{with} \quad (ds)^2 = g_{mn}dx^m dx^n$$

—but as curves whose tangents can be got one from another by parallel transportation: head off in some direction and "follow your nose" was the idea. Levi-Civita's idea so enriched a subject previously known as the "absolute differential calculus" that its name was changed . . . to "tensor analysis."

Our CATALOG (pages 120–122) can be looked upon as an ennumeration of circumstances in which—"by accident"—the $\varGamma$-apparatus falls away. Look, for example, to the "covariant curl," where we have

$$\begin{aligned} X_{m;n} - X_{n;m} &= (X_{m,n} - X_a\varGamma^a{}_{nm}) - (X_{n,m} - X_a\varGamma^a{}_{mn}) \\ &= X_{m,n} - X_{n,m} \qquad \text{by } \varGamma^a{}_{mn} = \varGamma^a{}_{nm} \end{aligned}$$

The basic principles of the "absolute differential calculus" were developed between 1884 and 1894 by Gregorio Ricci-Curbastro (1853–1925), who was a mathematician in the tradition of Riemann and Christoffel.[100] In 1896 his student, Tullio Levi-Civita (1873–1941), published "Sulle transformazioni della eqazioni dinamiche" to demonstrate the *physical* utility of the methods which Ricci himself had applied only to differential geometry. In 1900—at the urging of Felix Klein, in Göttingen—Ricci and Levi-Civita co-authored "Méthodes de calcul différentiel absolus et leurs applications," a lengthy review of the subject . . . but they were Italians writing in French, and published in a German periodical (Mathematische Annalen), and their work was largely ignored: for nearly twenty years the subject was known to only a few cognoscente (who included Minkowski at Göttingen), and cultivated by fewer. General interest in the subject developed—explosively!—only in the wake of Einstein's *general theory of relativity* (1916). Tensor methods had been brought to the reluctant attention of Einstein by Marcel Grossmann, a geometer who had been a classmate of Einstein's at the ETH in Zürich (Einstein reportedly used to study

[100] Ricci had interest also in physics, and as a young man published (in Nuovo Cimento) the first Italian account of Maxwellian electrodynamics.

Grossmann's class notes instead of attending Minkowski's lectures) and whose father had been instrumental in obtaining for the young and unknown Einstein a position in the Swiss patent office.

Acceptence of the tensor calculus was impeded for a while by those (mainly mathematicians) who perceived it to be in competition with the exterior calculus—an elegant French creation (Poincaré, Goursat, Cartan, ...) which treats (but more deeply) a narrower set of issues, but (for that very reason) supports also a robust *integral* calculus. The exterior calculus shares the Germanic pre-history of tensor analysis (Gauss, Grassmann, Riemann, ...) but was developed semi-independently (and somewhat later), and has only fairly recently begun to be included among the work-a-day tools of mathematical physicists. Every physicist can be expected today to have some knowledge of the tensor calculus, but the exterior calculus has yet to find a secure place in the pedagogical literature of physics, and for that (self-defeating) reason physicists who wish to be understood still tend to avoid the subject ... in their writing and (at greater hazard) in their creative thought.

3. Transformation properties of the electromagnetic field equations. We will be led in the following discussion from Maxwell's equations to—first and most easily—the group of "Lorentz transformations," which by some fairly natural interpretive enlargement detach from their electrodynamic birthplace to provide the foundation of Einstein's *Principle of Relativity*. But it will emerge that

> The covariance group of a theory depends
> in part upon *how the theory is expressed*:

slight adjustments in the formal rendition of Maxwell's equations will lead to transformation groups that differ radically from the Lorentz group (but that contain the Lorentz group as a subgroup) ... and that also is a lesson that admits of "enlargement"—that pertains to fields far removed from electrodynamics. The point merits explicit acknowledgement because it relates to how casually accepted conventions can exert unwitting control on the *development* of physics.

FIRST POINT OF VIEW Let Maxwell's equations be notated[101]

$$\partial_\mu F^{\mu\nu} = \tfrac{1}{c} j^\nu \tag{180.1}$$

$$\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \tag{180.2}$$

where $F^{\mu\nu}$ is antisymmetric and where

$$F_{\mu\nu} \equiv g_{\mu\alpha} g_{\nu\beta} F^{\alpha\beta} \quad \text{with} \quad \mathbb{g} \equiv \|g_{\mu\nu}\| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{181}$$

is (automatically) also antisymmetric. From $F_{\mu\nu} = -F_{\nu\mu}$ it follows, by the way,

[101] Compare (167).

that (180.2) reduces to the triviality $0 = 0$ unless μ, ν and λ are distinct, so the equation in question is just a condensed version of the sourceless Maxwell equations as they were encountered on page 111.[102] In view of entry (179.5) in our CATALOG it becomes natural to

> **assume** that $F^{\mu\nu}$ and j^{μ} transform as the components of tensor <u>densities of unit weight</u>:
>
> $$F^{\mu\nu} \longrightarrow {\color{red}F}^{\mu\nu} = W \cdot M^{\mu}{}_{\alpha} M^{\nu}{}_{\beta} F^{\alpha\beta} \qquad \mathbf{A}_1$$
>
> $$j^{\mu} \longrightarrow {\color{red}j}^{\mu} = W \cdot M^{\mu}{}_{\alpha} j^{\alpha} \qquad \mathbf{A}_2$$
>
> We note[103] that it makes coordinate-independent good sense to assume of the field tensor that it is antisymmetric:
>
> $$F^{\mu\nu} \text{ antisymmetric} \Longrightarrow {\color{red}F}^{\mu\nu} \text{ antisymmetric}$$

The unrestricted covariance (in the sense "form-invariance under coordinate transformation") of (180.1) is then assured

$$\partial_{\mu} F^{\mu\nu} = \tfrac{1}{c} j^{\nu} \longrightarrow {\color{red}\partial}_{\mu} {\color{red}F}^{\mu\nu} = \tfrac{1}{c} {\color{red}j}^{\nu}$$

On grounds that it would be intolerable for the description (181) of $\mathcal{g}$ to be "special to the coordinate system $\mathcal{X}$" we

> **assume** $g_{\mu\nu}$ to transform as a symmetric tensor of zero weight
>
> $$g_{\mu\nu} \longrightarrow {\color{red}g}_{\mu\nu} = W^{\alpha}{}_{\mu} W^{\beta}{}_{\nu} g_{\alpha\beta} \qquad \mathbf{B}_1$$
>
> but <u>impose upon $\mathcal{X} \to {\color{red}\mathcal{X}}$ the *constraint*</u> that
>
> $$= g_{\mu\nu} \qquad \mathbf{B}_2$$

This amounts in effect to imposition of the requirement that $\mathcal{X} \to {\color{red}\mathcal{X}}$ be of such a nature that

$$\mathbb{W}^{\mathsf{T}} \mathcal{g} \mathbb{W} = \mathcal{g} \quad \text{everywhere} \tag{182}$$

---

[102] We might write

$$\|F^{\mu\nu}\| \equiv \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}$$

$$\therefore \|F_{\mu\nu}\| = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix}$$

to establish explicit contact with orthodox 3-vector notation and terminology (and at the same time to make antisymmetry manifest), but such a step would be extraneous to the present line of argument.

[103] See again page 119.

Looking to the determinant of the preceding equation we obtain

$$W^2 = 1$$

from which (arguing from continuity) we conclude that

$$W \text{ is } \begin{cases} \text{everywhere equal to } +1, \text{ else} \\ \text{everywhere equal to } -1. \end{cases} \tag{183}$$

This result protects us from a certain embarrassment: assumptions $\mathbf{A}_1$ and $\mathbf{B}_1$ jointly imply that $F_{\mu\nu}$ transforms as a tensor of unit weight, while covariance of the windmill sum in (180.2) was seen at (179.3) to require $F_{\mu\nu}$ to transform as a weightless tensor. But (183) reduces all weight distinctions to empty trivialities. Thus does $\mathbf{B}_2$ insure the covariance of (180.2):

$$\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \longrightarrow \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0$$

From (182) we will extract the statement that

$$\mathcal{X} \rightarrow \mathcal{X} \text{ is a } \textbf{Lorentz transformation} \tag{184}$$

and come to the conclusion that *Maxwellian electrodynamics—as formulated above—is Lorentz covariant*. Lorentz (1904) and Einstein (1905) were the independent co-discoverers of this fundamental fact, which they established by two alternative (and quite distinct) lines of argument.

SECOND POINT OF VIEW Retain both the field equations (180) and the assumptions **A** but—in order to escape from the above-mentioned "point of embarrassment"—agree in place of $\mathbf{B}_1$ to

> **assume** that $g_{\mu\nu}$ transforms as a symmetric tensor *density* of weight $w = -\frac{1}{2}$

$$g_{\mu\nu} \longrightarrow g_{\mu\nu} = W^{-\frac{1}{2}} \cdot W^\alpha{}_\mu W^\beta{}_\nu g_{\alpha\beta} \qquad \mathbf{B}_1^*$$

for then $F_{\mu\nu}$ becomes weightless, as (179.3) requires. Retaining

$$= g_{\mu\nu} \qquad \mathbf{B}_2$$

we obtain

$$W^{-\frac{1}{2}} \cdot \mathbb{W}^{\mathsf{T}} \mathbb{g} \mathbb{W} = \mathbb{g} \quad \text{everywhere} \tag{185.1}$$

If spacetime were N-dimensional the determinantal argument would now give

$$W^{2-\frac{N}{2}} = 1$$

which (uniquely) in the physical case ($N = 4$) reduces to a triviality: $W^0 = 1$. The constraint (183) therefore drops away, with consequences which I will discuss in a moment.

THIRD POINT OF VIEW This differs only superficially from the viewpoint just considered. *Retain* $\mathbf{B}_1$ but in place of $\mathbf{B}_2$

assume that

$$g_{\mu\nu} = \Omega g_{\mu\nu} \qquad \mathbf{B}_2^*$$

Then

$$\mathbb{W}^{\mathsf{T}} g\, \mathbb{W} = \Omega g \tag{185.2}$$

and the determinantal argument supplies

$$\begin{aligned} \Omega &= W^{\frac{2}{N}} \\ &\downarrow \\ &= W^{\frac{1}{2}} \quad \text{in the physical case } N = 4 \end{aligned}$$

Equations (185.1) and (185.2) evidently say the same thing: the Lorentzian constraint (183) drops away and in place of (184) we have

$$\mathcal{X} \rightarrow \mathcal{X} \text{ is a } \textbf{conformal transformation} \tag{186}$$

The conformal covariance of Maxwellian electrodynamics was discovered independently by Cunningham[104] and Bateman.[105] It gives rise to ideas which have a curious past[106] and which have assumed a central place in elementary particle physics at high energy. Some of the electrodynamical implications of conformal covariance are so surprising that they have given rise to vigorous controversy.[107] A transformation is said (irrespective of the specific context) to be "conformal" if it <u>preserves angles locally</u> . . . though such transformations do not (in general) preserve non-local angles, nor do they (even locally) preserve length. Engineers make heavy use of the conformal recoordinatizations of the plane that arise from the theory of complex variables *via* the statement

$$z \rightarrow z = f(z) \quad : \quad f(z) \text{ analytic}$$

The bare bones of the argument: write $z = x + iy$, $z = u + iv$ and obtain

$$\begin{matrix} u = u(x,y) \\ v = v(x,y) \end{matrix} \quad \text{giving} \quad \begin{matrix} du = u_x dx + u_y dy \\ dv = v_x dx + v_y dy \end{matrix}$$

[104] E. Cunningham, "The principle of relativity in electrodynamics and an extension thereof," Proc. London Math. Soc. **8**, 223 (1910).

[105] H. Bateman, "The transformation of the electrodynamical equations," Proc. London Math. Soc. **8**, 223 (1910).

[106] T. Fulton, F. Rohrlich & L. Witten, "Conformal invariance in physics," Rev. Mod. Phys. **34**, 442 (1962).

[107] See "Radiation in hyperbolic motion" in R. Peierls, *Surprises in Theoretical Physics* (1979), page 160.

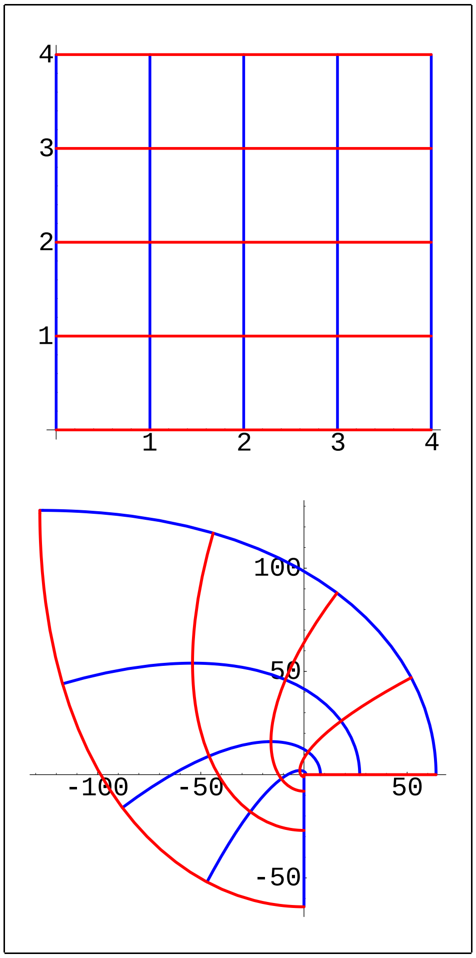

FIGURE 49: *Cartesian grid (above) and its conformal image (below) in the case $f(z) = z^3$, which supplies*

$$\begin{aligned} u(x,y) &= x^3 - 3xy^2 \\ v(x,y) &= 3x^2y - y^3 \end{aligned}$$

The command `ParametricPlot` *was used to construct the figure.*

But

$$\text{analyticity of } f(z) \iff \text{CAUCHY-RIEMANN CONDITIONS}: \begin{cases} u_x = +v_y \\ u_y = -v_x \end{cases}$$

so

$$\begin{pmatrix} u_x \\ u_y \end{pmatrix} \cdot \begin{pmatrix} v_x \\ v_y \end{pmatrix} = u_x v_x + u_y v_y = -u_x u_y + u_y u_x = 0$$

which is to say: curves of constant u are everywhere $\perp$ to curves of constant v, just as curves of constant x were everywhere normal to curves of constant y. The situation is illustrated in the preceding figure. The 2-dimensional case—in which one can conformally transform in as infinitely many ways as one can select $f(z)$—is, however, exceptional:[108] in the cases $N > 2$ conformality arises from a less esoteric circumstance, and the possibilities are described by a finite set of parameters. Let A^m and B^m be weightless vectors, let the inner product be defined $(A,B) \equiv g_{mn}A^m B^n$, and suppose g_{mn} to transform as a symmetric tensor *density of weight* w. Then (A,B) and the "squared lengths" (A,A) and (B,B) of all transform (not as invariants but) as *scalar densities*. But the

$$\text{angle between } A^m \text{ and } B^m \equiv \arccos\left\{\frac{(A,B)}{\sqrt{(A,A)(B,B)}}\right\}$$

clearly *does* transform by invariance. Analysis of (185.2) gives rise in the physical case $(N = 4)$ to a 15-parameter *conformal group* that contains the 6-parameter *Lorentz group* as a subgroup.

FOURTH POINT OF VIEW Adopt the (unique) affine connection $\Gamma^\lambda{}_{\mu\nu}$ which vanishes here in our inertial $\mathcal{X}$-coordinate system. For us there is then no distinction between ordinary differentiation and covariant differentiation. So in place of (180) we can, if we wish, write

$$F^{\mu\nu}{}_{;\mu} = \tfrac{1}{c}j^\nu \tag{187.1}$$

$$F_{\nu\lambda;\mu} + F_{\lambda\mu;\nu} + F_{\mu\nu;\lambda} = 0 \tag{187.2}$$

Which is to say: we can elect to "tensorially continuate" our Maxwell equations to other coordinate systems or arbitrary (moving curvilinear) design. We retain the description (181) of $g_{\mu\nu}$, and we retain

$$g_{\mu\nu} \longrightarrow g_{\mu\nu} = W^\alpha{}_\mu W^\beta{}_\nu g_{\alpha\beta} \tag*{$\mathbf{B}_1$}$$

But we have no longer any reason to retain $\mathbf{B}_2$, no longer any reason to impose any specific constraint upon the design of $g_{\mu\nu}$. We arrive thus at a formalism in which

$$F^{\mu\nu}{}_{;\mu} = \tfrac{1}{c}j^\nu \longrightarrow F^{\mu\nu}{}_{;\mu} = \tfrac{1}{c}j^\nu$$

$$F_{\nu\lambda;\mu} + F_{\lambda\mu;\nu} + F_{\mu\nu;\lambda} = 0 \longrightarrow F_{\nu\lambda;\mu} + F_{\lambda\mu;\nu} + F_{\mu\nu;\lambda} = 0$$

and in which

$$\mathcal{X} \to \mathcal{X} \text{ is } \textbf{unrestricted} \tag{188}$$

No "natural weights" are assigned within this formalism to $F^{\mu\nu}$, j^μ and $g_{\mu\nu}$, but formal continuity with the conformally-covariant formalism (whence with the Lorentz-covariant formalism) seems to require that we assign weights $w = 1$ to $F^{\mu\nu}$ and j^μ, weight $w = -\frac{1}{2}$ to $g_{\mu\nu}$.

[108] See page 55 of "The transformations which preserve wave equations" (1979) in TRANSFORMTIONAL PHYSICS OF WAVES (1979–1981).

Still other points of view are possible,[109] but I have carried this discussion already far enough to establish the validity of a claim made at the outset: the only proper answer to the question "What transformations $\mathcal{X} \to \mathcal{X}$ preserve the structure of Maxwell's equations?" is "It depends—depends on how you have chosen to *write* Maxwell's equations."

We have here touched, in a physical setting, upon an idea—look at "objects," and the groups of transformations which preserve relationships among those objects—which Felix Klein, in the lecture given when (in 1872, at the age of 23) he assumed the mathematical professorship at the University of Erlangen, proposed might be looked upon as *the* organizing principle of *all* pure/applied mathematics—a proposal which has come down to us as the "Erlangen Program." It has been supplanted in the world of pure mathematics, but continues to illuminate the historical and present development of physics.[110]

4. Lorentz transformations, and some of their implications. To state that $\mathcal{X} \leftarrow \mathcal{X}$ is a *Lorentz transformation* is, by definition, to state that the associated transformation matrix $\mathbb{M} \equiv \|M^\mu{}_\nu\| \equiv \|\partial x^\mu/\partial x^\nu\|$ has (see again page 127) the property that

$$\mathbb{M}^\mathsf{T} g\, \mathbb{M} = g \quad \text{everywhere} \tag{182}$$

where by fundamental assumption $g = g^\mathsf{T} = g^{-1}$ possesses at each point in spacetime the specific structure indicated at (181).

I begin with the observation that $\mathbb{M}$ must necessarily be a *constant* matrix. The argument is elementary: hit (182) with ∂_λ and obtain

$$(\partial_\lambda \mathbb{M})^\mathsf{T} g \mathbb{M} + \mathbb{M}^\mathsf{T} g (\partial_\lambda \mathbb{M}) = \mathbb{O} \quad \text{because } g \text{ is constant}$$

This can be rendered

$$g_{\alpha\beta} M^\alpha{}_{\lambda\mu} M^\beta{}_\nu + g_{\alpha\beta} M^\alpha{}_\mu M^\beta{}_{\nu\lambda} = 0$$

where $M^\alpha{}_{\lambda\mu} \equiv \partial^2 x^\alpha/\partial x^\lambda \partial x^\mu = M^\alpha{}_{\mu\lambda}$. More compactly

$$\Gamma_{\mu\nu\lambda} + \Gamma_{\nu\lambda\mu} = 0$$

where $\Gamma_{\mu\nu\lambda} \equiv g_{\alpha\beta} M^\alpha{}_\mu M^\beta{}_{\nu\lambda}$. Also (subjecting the $\mu\nu\lambda$ to cyclic permutation)

$$\Gamma_{\nu\lambda\mu} + \Gamma_{\lambda\mu\nu} = 0$$
$$\Gamma_{\lambda\mu\nu} + \Gamma_{\mu\nu\lambda} = 0$$

so

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \Gamma_{\lambda\mu\nu} \\ \Gamma_{\mu\nu\lambda} \\ \Gamma_{\nu\lambda\mu} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

---

[109] See D. van Dantzig, "The fundamental equations of electromagnetism, independent of metric geometry," Proc. Camb. Phil. Soc. **30**, 421 (1935).

[110] For an excellent discussion see the section "Codification of geometry by invariance" (pages 442–453) in E. T. Bell's *The Development of Mathematics* (1945). The Erlangen Program is discussed in scholarly detail in T. Hawkins, *Emergence of the theory of Lie Groups* (2000): see the index. For a short history of tensor analysis, see Bell's Chapter 9.

The 3×3 matrix is non-singular, so we must have

$$\Gamma_{\lambda\mu\nu} = M^{\alpha}{}_{\lambda} g_{\alpha\beta} \partial_\mu M^{\beta}{}_{\nu} = 0 \quad : \quad \text{ditto cyclic permutations}$$

which in matrix notation reads

$$\mathbb{M}^{\mathsf{T}} g (\partial_\mu \mathbb{M}) = \mathbb{O}$$

The matrices $\mathbb{M}$ and g are non-singular, so we can multiply by $(\mathbb{M}^{\mathsf{T}} g)^{-1}$ to obtain

$$\partial_\mu \mathbb{M} = \mathbb{O} \quad : \quad \text{the elements of } \mathbb{M} \text{ must be } \underline{\text{constants}}$$

The functions $x^\mu(x)$ that describe the transformation $\mathfrak{X} \leftarrow \mathfrak{X}$ must possess therefore the <u>inhomogeneous linear</u> structure[111]

$$x^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu \quad : \quad \text{the } \Lambda^\mu{}_\nu \text{ and } a^\mu \text{ are constants}$$

The transformation matrix $\mathbb{M}$, with elements given now by constants $\Lambda^\mu{}_\nu$, will henceforth be denoted $\mathbb{\Lambda}$ to emphasize that it is no longer generic but has been specialized (and also to suggest "Lorentz"). We shall (when the risk of confusion is slight) write

$$x = \mathbb{\Lambda} x + \underbrace{a}_{\text{describes a \underline{translation} in spacetime}} \tag{189.1}$$

to describe an ("inhomogeneous Lorentz" or) **Poincaré transformation**, and

$$x = \mathbb{\Lambda} x \tag{189.2}$$

to describe a (simple homogeneous) **Lorentz transformation**, the assumption in both cases being that

$$\mathbb{\Lambda}^{\mathsf{T}} g \, \mathbb{\Lambda} = g \tag{190}$$

> IMPORTANT REMARK: Linearity of a transformation—constancy of the transformation matrix—is sufficient in itself to kill all "extraneous terms," without the assistance of weight restrictions.

It was emphasized on page 119 that "not every indexed object transforms tensorially," and that, in particular, the x^μ themselves do not transform tensorially *except in the linear case*. We have now in hand just such a case, and for that reason relativity becomes—not just locally but globally—an exercise in linear algebra. Spacetime has become a 4-dimensional vector space; indeed, it has become an *inner product space*, with

$$\left.\begin{aligned} (x,y) &\equiv g_{\mu\nu} x^\mu y^\nu \\ &= (y,x) \quad \text{by} \quad g_{\mu\nu} = g_{\nu\mu} \\ &= x^{\mathsf{T}} g \, y \\ &= x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3 = x^0 y^0 - \boldsymbol{x}\cdot\boldsymbol{y} \end{aligned}\right\} \tag{191.1}$$

[111] Einstein (1905)—on the grounds that what he sought was a *minimal modification of the Galilean transformations* (which are themselves linear)—was content simply to *assume* linearity.

The *Lorentz inner product* (interchangeably: the "Minkowski inner product") described above is, however, "pathological" in the sense that it gives rise to an "indefinite norm;" *i.e.*, to a norm

$$\left.\begin{aligned}(x,x) &= g_{\mu\nu}x^\mu x^\nu\\ &= x^\mathsf{T}\mathbb{g}\,x\\ &= (x^0)^2-(x^1)^2-(x^2)^2-(x^3)^2=(x^0)^2-\boldsymbol{x}\cdot\boldsymbol{x}\end{aligned}\right\}\tag{191.2}$$

which (instead of being positive unless $x=0$) can assume either sign, and can vanish even if $x\neq 0$. From this primitive fact radiates much—arguably all—that is most distinctive about the geometry of spacetime . . . which, as Minkowski was the first to appreciate (and as will emerge) lies at the heart of the theory of relativity.

If A^μ, B^μ and $g_{\mu\nu}$ transform as weightless tensors, then basic tensor algebra informs us that $g_{\mu\nu}A^\mu B^\nu$ transforms by invariance:

$$g_{\mu\nu}A^\mu B^\nu \longrightarrow g_{\mu\nu}A^\mu B^\nu = g_{\mu\nu}A^\mu B^\nu \quad \text{unrestrictedly}$$

What distinguishes Lorentz transformations from transformations-in-general is that

$$g_{\mu\nu} = g_{\mu\nu}$$

To phrase the issue as it relates not to things (like A^μ and B^μ) "written on" spacetime but to the structure of spacetime itself, we can state that the linear transformation

$$x \longrightarrow x = \Lambda x$$

describes a Lorentz transformation if and only if

$$x^\mathsf{T}\mathbb{g}\,y = x^\mathsf{T}\Lambda^\mathsf{T}\mathbb{g}\,\Lambda\,y = x^\mathsf{T}\mathbb{g}\,y \quad \text{for all } x \text{ and } y \quad : \quad \text{entails} \quad \Lambda^\mathsf{T}\mathbb{g}\,\Lambda = \mathbb{g}$$

where, to be precise, we require that $\mathbb{g}$ has the specific design

$$\mathbb{g} \equiv \begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}$$

that at (163) was impressed upon us by our interest in the transformation properties of Maxwell's equations (*i.e.*, by some narrowly prescribed specific physics).

We come away with the realization that *Lorentz transformations have in fact only incidentally to do with electrodynamics*: they are the transformations that <u>preserve Lorentzian inner products</u>, which is to say: that <u>preserve the metric properties of spacetime</u> . . . just as "rotations" $\boldsymbol{x} \longrightarrow \boldsymbol{x} = \mathbb{R}\boldsymbol{x}$ are the linear transformations that preserve Euclidean inner products

$$\boldsymbol{x}^\mathsf{T}\mathbb{I}\,\boldsymbol{y} = \boldsymbol{x}^\mathsf{T}\mathbb{R}^\mathsf{T}\mathbb{I}\,\mathbb{R}\,\boldsymbol{y} = \boldsymbol{x}^\mathsf{T}\mathbb{I}\,\boldsymbol{y} \quad \text{for all } \boldsymbol{x} \text{ and } \boldsymbol{y} \quad : \quad \text{entails} \quad \mathbb{R}^\mathsf{T}\mathbb{R} = \mathbb{I}$$

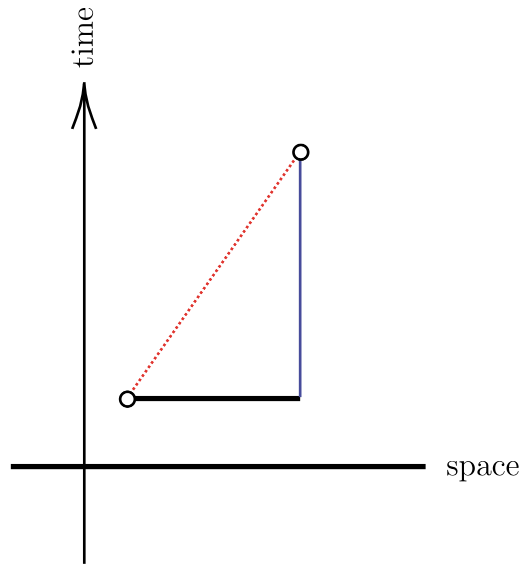

FIGURE 50: *Two "events" identify a triangle in the spacetime. Relativity asks each inertial observer to use metersticks and clocks to assign traditional meanings to the "Euclidean length" of the black side (here thickened to suggest that space is several-dimensional) and to the "duration" of the blue side—meanings which (as will emerge) turn out, however, to yield observer-dependent numbers—but assigns (Lorentz-invariant!) meaning also to the squared length of the hypotenuse.*

and in so doing preserve the lengths/angles/areas/volumes . . . that endow Euclidean 3-space with its distinctive metric properties.

That spacetime can be said to possess metric structure is the great surprise, the great discovery. In pre-relativistic physics one could speak of the duration (quantified by a clock) of the temporal interval $\Delta t = t_a - t_b$ separating a pair of events, and one could speak of the length

$$\Delta\ell = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2 + (z_a - z_b)^2}$$

(quantified by a meter stick) of the spatial interval separating a pair of points; one spoke of "space" and "time," but "spacetime" remained an abstraction of the design space $\otimes$ time. Only with the introduction g did it become possible (see Figure 50) to speak of the (squared) length

$$(\Delta s)^2 = c^2(t_a - t_b)^2 - (x_a - x_b)^2 - (y_a - y_b)^2 - (z_a - z_b)^2$$

of the interval separating $(t_a, \boldsymbol{x}_a)$ from $(t_b, \boldsymbol{x}_b)$:

"space $\otimes$ time" had become "spacetime"

The first person to recognize the profoundly revolutionary nature of what had been accomplished was (not Einstein but) Minkowski, who began an address to the Assembly of German Natural Scientists & Physicians (21 September 1908) with these words:

> *"The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality."*

Electrodynamics had led to the first clear perception of the geometrical design of the spacetime manifold upon which all physics is written. The symmetries inherent in that geometry were by this time know to be reflected in the design of Maxwell's equations. Einstein's *Principle of Relativity* holds that they must, in fact, be reflected in the design of *all* physical theories—irrespective of the specific phenomenology to which any individual theory may refer.

Returning now to the technical mainstream of this discussion ... let the Lorentz condition (190) be written

$$\Lambda^{-1} = g^{-1}\Lambda^{\mathsf{T}} g \tag{192}$$

Generally inversion of a 4×4 matrix is difficult, but (192) shows that *inversion of a Lorentz matrix Λ can be accomplished very easily*.[112].

Equations (190/192) impose a multiplicative condition upon Λ. It was to reduce multiplicative conditions to additive conditions (which are easier) that logarithms were invented. Assume, therefore, that Λ can be written

$$\Lambda = e^{\mathbb{A}} = \mathbb{I} + \mathbb{A} + \tfrac{1}{2!}\mathbb{A}^2 + \cdots$$

It now follows that

$$\Lambda^{-1} = e^{-\mathbb{A}} \quad \text{while} \quad g^{-1}\Lambda^{\mathsf{T}} g = g^{-1} e^{\mathbb{A}^{\mathsf{T}}} g = e^{g^{-1}\mathbb{A}^{\mathsf{T}} g}$$

Evidently Λ will be a Lorentz matrix if

$$-\mathbb{A} = g^{-1}\mathbb{A}^{\mathsf{T}} g$$

which (by $g^{\mathsf{T}} = g$) can be expressed

$$(g\mathbb{A})^{\mathsf{T}} = -(g\mathbb{A})$$

This is an *additive* condition (involves negation instead of inversion) and amounts simply to the statement that $g\mathbb{A} \equiv \|A_{\mu\nu}\|$ is *antisymmetric*. Adopt this notation

$$g\mathbb{A} = \begin{pmatrix} 0 & A_1 & A_2 & A_3 \\ -A_1 & 0 & -a_3 & a_2 \\ -A_2 & a_3 & 0 & -a_1 \\ -A_3 & -a_2 & a_1 & 0 \end{pmatrix}$$

[112] PROBLEM 40

where $\{A_1, A_2, A_3, a_1, a_2, a_3\}$ comprise a *sextet of adjustable real constants*. Multiplication on the left by $\mathbb{g}^{-1}$ gives a matrix of (what I idiosyncratically call) the "$\mathbb{g}$-antisymmetric" design[113]

$$\mathbb{A} \equiv \|A^\mu{}_\nu\| = \begin{pmatrix} 0 & A_1 & A_2 & A_3 \\ A_1 & 0 & a_3 & -a_2 \\ A_2 & -a_3 & 0 & a_1 \\ A_3 & a_2 & -a_1 & 0 \end{pmatrix}$$

We come thus to the conclusion that matrices of the form

$$\mathbb{\Lambda} = \exp \begin{pmatrix} 0 & A_1 & A_2 & A_3 \\ A_1 & 0 & a_3 & -a_2 \\ A_2 & -a_3 & 0 & a_1 \\ A_3 & a_2 & -a_1 & 0 \end{pmatrix} \tag{193}$$

are Lorentz matrices; *i.e.*, they satisfy (190/192), and when inserted into (189) they describe Poincaré/Lorentz transformations.

Does every Lorentz matrix $\mathbb{\Lambda}$ admit of such representation? Not quite. It follows immediately from (190) that $(\det \mathbb{\Lambda})^2 = 1$; *i.e.*, that

$$\Lambda \equiv \det \mathbb{\Lambda} = \pm 1, \text{ according as } \mathbb{\Lambda} \text{ is} \begin{cases} \text{"proper"} \\ \text{"improper"} \end{cases}$$

while the theory of matrices supplies the lovely identity[114]

$$\det(e^{\mathbb{M}}) = e^{\operatorname{tr}\mathbb{M}} \quad : \quad \mathbb{M} \text{ is } any \text{ square matrix} \tag{194}$$

We therefore have $\Lambda = \det(e^{\mathbb{A}}) = 1$ by $\operatorname{tr}\mathbb{A} = 0$:

> Every Lorentz matrix $\mathbb{\Lambda}$ of the form (193) is necessarily *proper*; moreover (as will emerge), every proper $\mathbb{\Lambda}$ admits of such an "exponential representation." (195)

It will emerge also that when one has developed the structure of the matrices $\mathbb{\Lambda} = e^{\mathbb{A}}$ one has "cracked the nut," in the sense that it becomes easy to describe their improper companions.[115]

What it means to "develop the structure of $\mathbb{\Lambda} = e^{\mathbb{A}}$" is exposed most simply in the (physically artificial) case $N = 2$. Taking

$$\mathbb{g} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad : \quad \text{Lorentz metric in 2-dimensional spacetime}$$

---

[113] Notice that $\mathbb{g}$-antisymmetry becomes literal antisymmetry when the metric $\mathbb{g}$ is Euclidean. Notice also that while it makes tensor-algebraic good sense to write $\mathbb{A}^2 = \|A^\mu{}_\alpha A^\alpha{}_\nu\|$ it would be hazardous to write $(\mathbb{g}\mathbb{A})^2 = \|A_{\mu\alpha}A_{\alpha\nu}\|$.

[114] PROBLEM 41.

[115] PROBLEM 42.

as our point of departure, the argument that gave (193) gives

$$\mathbb{\Lambda} = \exp\begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix} = e^{A\mathbb{J}} \tag{196.1}$$

where evidently

$$\mathbb{J} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

By quick calculation (or, more elegantly, by appeal to the *Cayley-Hamilton theorem*, according to which every matrix satisfies its own characteristic equation) we find $\mathbb{J}^2 = \mathbb{I}$, from which it follows that

$$\mathbb{J}^n = \begin{cases} \mathbb{I} & \text{if } n \text{ is even} \\ \mathbb{J} & \text{if } n \text{ is odd} \end{cases}$$

So

$$\begin{aligned} \mathbb{\Lambda} &= \underbrace{\left\{1 + \tfrac{1}{2!}A^2 + \tfrac{1}{4!}A^4 + \cdots\right\}}_{\cosh A}\mathbb{I} + \underbrace{\left\{A + \tfrac{1}{3!}A^3 + \tfrac{1}{5!}A^5 + \cdots\right\}}_{\sinh A}\mathbb{J} \\ &= \begin{pmatrix} \cosh A & \sinh A \\ \sinh A & \cosh A \end{pmatrix} \qquad (196.2) \\ &\equiv \mathbb{\Lambda}(A) \quad : \quad \text{Lorentzian for all real values of } A \end{aligned}$$

It is evident—whether one argues from (196.2) of (more efficiently) from (196.1) —that

$$\begin{aligned} \mathbb{I} &= \mathbb{\Lambda}(0) && : \text{ existence of identity} && (197.1) \\ \mathbb{\Lambda}(A_2)\mathbb{\Lambda}(A_1) &= \mathbb{\Lambda}(A_1 + A_2) && : \text{ compositional closure} && (197.2) \\ \mathbb{\Lambda}^{-1}(A) &= \mathbb{\Lambda}(-A) && : \text{ existence of inverse} && (197.3) \end{aligned}$$

and that all such $\mathbb{\Lambda}$-matrices *commute*.

We are now—but *only* now—in position to consider the kinematic meaning of A, and of the action of $\mathbb{\Lambda}(A)$. We are, let us pretend, a "point PhD" who—having passed the physical tests required to establish our inertiality—use our "good clock and Cartesian frame" to assign coordinates $x \equiv \{x^0, x^1, x^2, x^3\}$ to events. O—a second observer, similarly endowed, who we see to be gliding by with velocity $\boldsymbol{v}$—assigns coordinates $x \equiv \{x^0, x^1, x^2, x^3\}$ to those same events. O shares our confidence in the validity of Maxwellian electrodynamics: we can therefore write $x = \mathbb{\Lambda}x + a$. In the interests merely of simplicity we will *assume that O's origin and our origin coincide*: the translational terms a^μ then drop away and we have $x = \mathbb{\Lambda}x$ ... which in the 2-dimensional case reads

$$\begin{pmatrix} x^0 \\ x^1 \end{pmatrix} = \begin{pmatrix} \cosh A & \sinh A \\ \sinh A & \cosh A \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \end{pmatrix} \tag{198}$$

To describe the successive "ticks of the clock at his origin" O writes

$$\begin{pmatrix} ct \\ 0 \end{pmatrix}$$

while—to describe those same events—*we* write

$$\begin{pmatrix} ct \\ vt \end{pmatrix}$$

Immediately $vt = ct \cdot \sinh A$ and $ct = ct \cdot \cosh A$ which, when we divide the former by the latter, give

$$\tanh A = \beta \tag{199}$$

with

$$\beta \equiv v/c \tag{200}$$

These equations serve to *assign kinematic meaning to* A*, and therefore to* $\mathbb{\Lambda}(A)$. Drawing now upon the elementary identities

$$\cosh A = \frac{1}{\sqrt{1 - \tanh^2 A}} \qquad \text{and} \qquad \sinh A = \frac{\tanh A}{\sqrt{1 - \tanh^2 A}}$$

we find that (198) can be written

$$\begin{pmatrix} x^0 \\ x^1 \end{pmatrix} = \gamma \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \end{pmatrix} \tag{201}$$

with

$$\gamma \equiv \frac{1}{\sqrt{1 - \beta^2}} = 1 + \tfrac{1}{2}\beta^2 + \tfrac{3}{8}\beta^4 + \cdots \tag{202}$$

Evidently γ becomes singular (see Figure 51) at $\beta^2 = 1$; *i.e.*, at $v = \pm c$ ... with diverse consequences which we will soon have occasion to consider. The *non-relativistic limit* arises physically from $\beta^2 \ll 1$; *i.e.*, from $v^2 \ll c^2$, but can be considered formally to arise from $c \uparrow \infty$. One must, however, take careful account of the c that lurks in the definitions of x^0 and x^0: when that is done, one finds that (201) assumes the (less memorably symmetric) form

$$\begin{pmatrix} t \\ x \end{pmatrix} = \gamma \begin{pmatrix} 1 & v/c^2 \\ v & 1 \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix}$$

giving

$$\downarrow$$

$$= \begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} \quad \text{as} \quad c \uparrow \infty \tag{203}$$

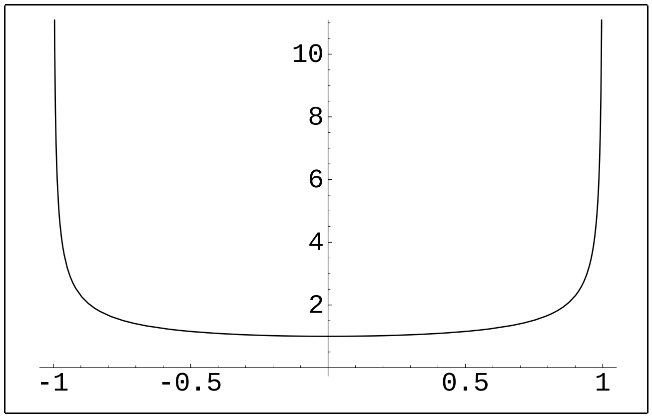

FIGURE 51: *Graph of the β-dependence of $\gamma \equiv 1/\sqrt{1-\beta^2}$, as $\beta \equiv v/c$ ranges on the physical interval $-1 < \beta < +1$. Outside that interval γ becomes imaginary.*

Heretofore we have been content to share our profession with a zippy population of "superluminal inertial observers" who glide past us with speeds $v > c$. But

$$\mathbb{\Lambda}(\beta) \text{ becomes } \underline{\text{imaginary}} \text{ when } \beta^2 > 1$$

We cannot enter into meaningful dialog with such observers; we therefore strip them of their clocks, frames and PhD's and send them into retirement, denied any further collaboration in the development of our relativistic theory of the world[114]—indispensable though they were to our former Galilean activity. Surprisingly, we can get along very well without them, for

$$\begin{aligned}
&\mathbb{\Lambda}(\beta_2)\mathbb{\Lambda}(\beta_1) = \mathbb{\Lambda}(\beta)\\
&\qquad \beta = \beta(\beta_1,\beta_2) = \tanh(A_1 + A_2)\\
&\qquad\qquad = \frac{\tanh A_1 + \tanh A_2}{1 + \tanh A_1 \tanh A_2}\\
&\qquad\qquad = \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2}
\end{aligned} \tag{204}$$

entails (this is immediately evident in Figure 52) that

if $v_1 < c$ and $v_2 < c$ then so also is $v(v_1, v_2) < c$:
one cannot leapfrog into the superluminal domain

The function $\beta(\beta_1, \beta_2)$ plays in (2-dimensional) relativity a role precisely analogous to a "group table" in the theory of finite groups: it describes <u>how Lorentz transformations compose</u>, and possess many wonderful properties, of

[114] This, however, does *not*, of itself, deny any conceivable role to superluminal *signals* or *particles* in a relativistic physics!

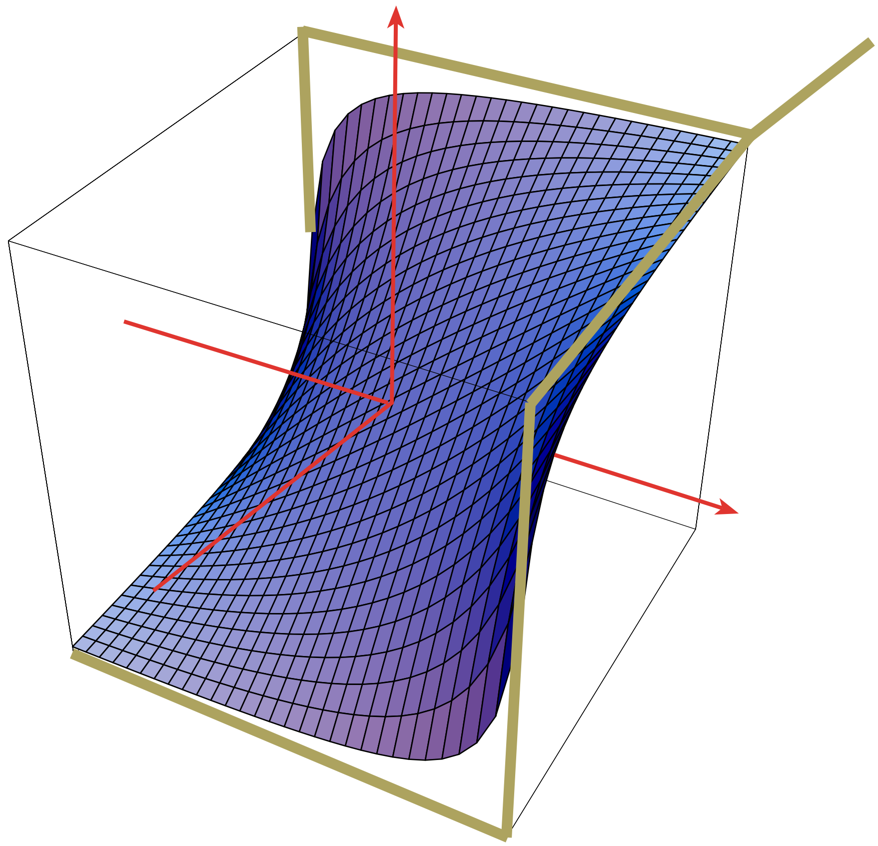

FIGURE 52: *Graph of the function $\beta(\beta_1, \beta_2)$. The vertices of the frame stand at the points $\{\pm 1, \pm 1, \pm 1\}$ in 3-dimensional β-space. If we write $\beta_3 = -\beta(\beta_1, \beta_2)$ then (204) assumes the high symmetry*

$$\beta_1 + \beta_2 + \beta_3 + \beta_1\beta_2\beta_3 = 0$$

clearly evident in the figure. The "β-surface" looks rather like a soap film spanning the 6-sided frame that results when the six untouched edges of the cube are discarded.

which I list here only a few:

$$\begin{aligned}\beta(\beta_1, \beta_2) &= \beta(\beta_2, \beta_1)\\ \beta(\beta_1, \beta_2) &= 0 \quad \text{if} \quad \beta_2 = -\beta_1\\ \beta(1, 1) &= 1\end{aligned}$$

To this list our forcibly retired superluminal friends might add the following:

$$\beta(\beta_1, \beta_2) = \beta(\tfrac{1}{\beta_1}, \tfrac{1}{\beta_2})$$

If β is subluminal then $\frac{1}{\beta}$ is superluminal. So we have here the statement that the compose of two superluminal Lorentz transformations is subluminal (the i's have combined to become real). Moreover, every subluminal Lorentz transformation can be displayed as such a compose (in many ways). Curious!

Equation (204) is often presented as "relativistic velocity addition formula"

$$\begin{aligned} v &= \frac{v_1 + v_2}{1 + v_1 v_2/c^2} \\ &= (v_1 + v_2) \cdot \Big[1 - \Big(\frac{v_1 v_2}{c^2}\Big) + \Big(\frac{v_1 v_2}{c^2}\Big)^2 - \Big(\frac{v_1 v_2}{c^2}\Big)^3 + \cdots\Big] \\ &= (\text{Galilean formula}) \cdot \Big[\text{relativistic correction factor}\Big] \end{aligned}$$

but that portrayal of the situation—though sometimes useful—seems to me to miss (or to entail risk of missing) the simple origin and essential significance of (204): the tradition that has, for now nearly a century, presented relativity as a source of endless paradox (and which has, during all that time, contributed little or nothing to understanding—paradox being, as it is, a symptom of imperfect understanding) should be allowed to wither.

In applications we will have need also of $\gamma(\beta_1, \beta_2) \equiv [1 - \beta^2(\beta_1, \beta_2)]^{-\frac{1}{2}}$, the structure of which is developed most easily as follows:

$$\begin{aligned} \gamma &= \cosh(A_1 + A_2) \\ &= \cosh A_1 \cosh A_2 \big[1 + \tanh A_1 \tanh A_2\big] \\ &= \gamma_1 \gamma_2 \big[1 + \beta_1 \beta_2\big] \end{aligned} \tag{205}$$

This "γ-composition law"—in which we might (though it is seldom useful) use

$$\beta = \sqrt{1 - \gamma^{-2}} = \frac{\sqrt{(\gamma + 1)(\gamma - 1)}}{\gamma}$$

to eliminate the surviving β's—will acquire importance when we come to the theory of radiation.

5. Geometric considerations. Our recent work has been algebraic. The following remarks emphasize the geometrical aspects of the situation, and are intended to provide a more vivid sense of what Lorentz transformations are all about. By way of preparation: In Euclidean 3-space the equation $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{x} = r^2$ defines a sphere (concentric about the origin, of radius r) which—consisting as it does of points all of which lie at the *same (Euclidean) distance from the origin*—we may reasonably call an "isometric surface." Rotations ($\boldsymbol{x} \to \boldsymbol{x} = \mathbb{R}\,\boldsymbol{x}$ with $\mathbb{R}^{\mathsf{T}}\mathbb{R} = \mathbb{I}$) cause the points of 3-space to shift about, but by a linear rule (straight lines remain straight) that <u>maps isometric spheres onto themselves</u>: such surfaces are, in short, "$\mathbb{R}$-invariant." Similarly …

In spacetime the σ-parameterized equations

$$\boldsymbol{x}^{\mathsf{T}} g\, \boldsymbol{x} = \sigma$$

define a population of <u>Lorentz-invariant isometric surfaces</u> Σ_σ. The surfaces that in 3-dimensional spacetime arise from

$$(x^0)^2 - (x^1)^2 - (x^2)^2 = \sigma$$

which describes a

- hyperboloid of two sheets in the case $\sigma > 0$
- cone in the case $\sigma = 0$
- hyperboloid of one sheet in the case $\sigma < 0$

are shown in Figure 53. The analogous construction in 2-dimensional spacetime (Figure 54) is easier to sketch, and serves most purposes well enough, but is misleading in one important respect: it fails to indicate the profound distinction between one-sheeted and two-sheeted hyperboloids. On the former one can move continuously from any point to any other (one can, in particular, get from one to the other by Lorentz transformation), but passage from one sheet to the other is necessarily *dis*continuous (requires "time reflection," can might be symbolized

$$\text{future} \rightleftarrows \text{past}$$

and cannot be executed "a little bit at a time").

How—within the geometric framework just described—is one to represent the action $x \longrightarrow x = \mathbb{\Lambda} x$ of $\mathbb{\Lambda}(\beta)$? I find it advantageous to approach the question somewhat obliquely: Suppose O to be thinking about the points (events)

$$\begin{pmatrix} +1 \\ +1 \end{pmatrix}, \begin{pmatrix} +1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ +1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

that mark the vertices of a "unit square" on her spacetime diagram. By quick calculation

$$\left.\begin{array}{l} \begin{pmatrix} +1 \\ +1 \end{pmatrix} \longrightarrow K^+(\beta)\begin{pmatrix} +1 \\ +1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ -1 \end{pmatrix} \longrightarrow K^+(\beta)\begin{pmatrix} -1 \\ -1 \end{pmatrix} \\ \begin{pmatrix} +1 \\ -1 \end{pmatrix} \longrightarrow K^-(\beta)\begin{pmatrix} +1 \\ -1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ +1 \end{pmatrix} \longrightarrow K^-(\beta)\begin{pmatrix} -1 \\ +1 \end{pmatrix} \end{array}\right\} \tag{206}$$

where

$$K^+(\beta) \equiv \sqrt{\tfrac{1+\beta}{1-\beta}} \quad \text{and} \quad K^-(\beta) \equiv \sqrt{\tfrac{1-\beta}{1+\beta}} \tag{207}$$

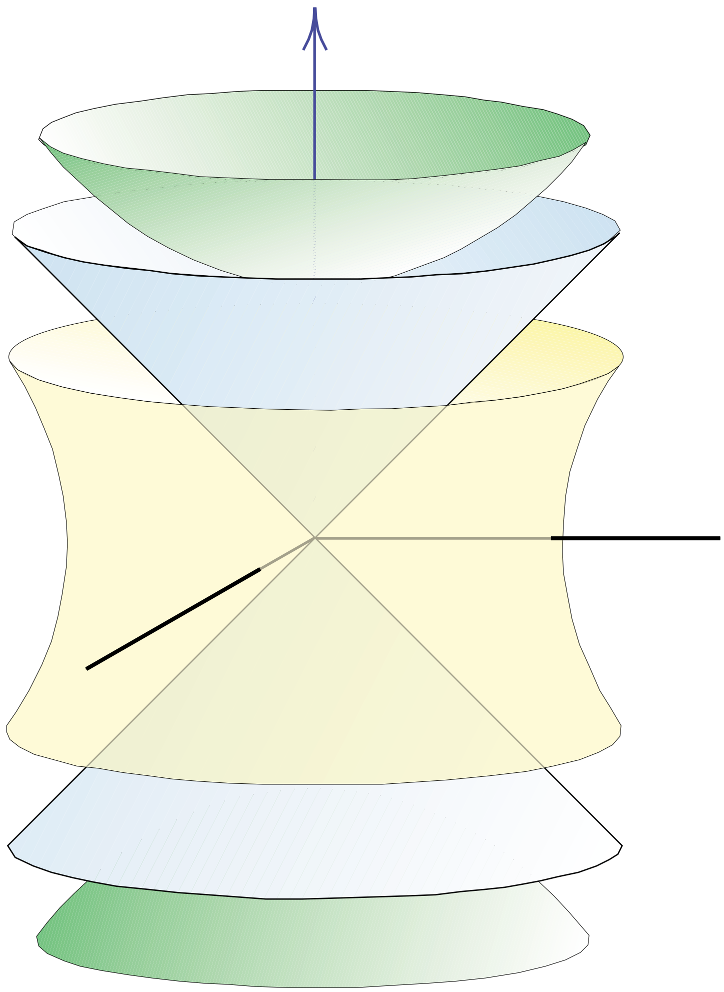

Figure 53: *Isometric surfaces in 3-dimensional spacetime. The arrow is "the arrow of time." Points on the blue "null cone" (or "light cone") are defined by the condition* $\sigma = 0$*: the interval separating such points from the origin has zero squared length (in the Lorentzian sense). Points on the green cup (which is interior to the forward cone) lie in the "future" of the origin, while points on the green cap (interior to the backward cone) lie in the "past:" in both cases* $\sigma > 0$*. Points on the yellow girdle (exterior to the cone) arise from* $\sigma < 0$*: they are separated from the origin by intervals of negative squared length, and are said to lie "elsewhere." In physical (4-dimensional) spacetime the circular cross sections (cut by "time-slices") become spherical. Special relativity acquires many of its most distinctive features from the circumstance that the isometric surfaces* Σ_σ *are hyperboloidal.*

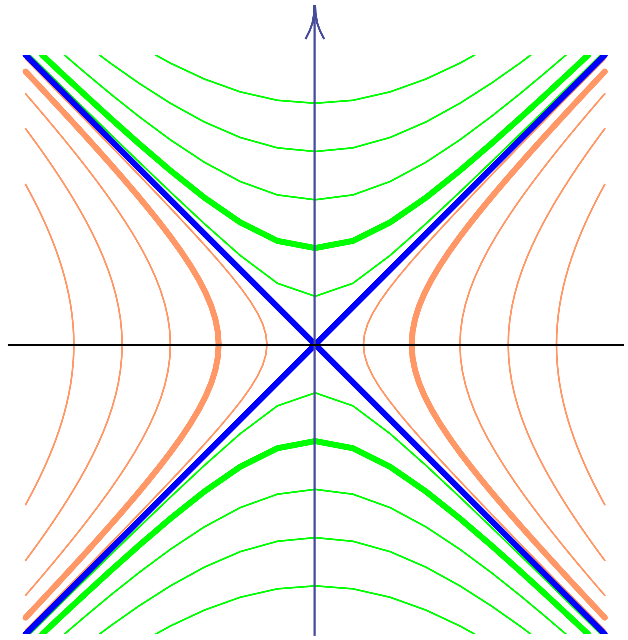

FIGURE 54: *The isometric surfaces shown in the preceding figure become isometric* curves *in 2-dimensional spacetime, where all hyperbolas have two branches. We see that*

$$\begin{pmatrix}1\\0\end{pmatrix} \textit{ gives } \sigma = 1^2 - 0^2 = +1, \textit{ typical of points with timelike}$$

$$\begin{pmatrix}1\\1\end{pmatrix} \textit{ gives } \sigma = 1^2 - 1^2 = 0, \textit{ typical of points with lightlike}$$

$$\begin{pmatrix}0\\1\end{pmatrix} \textit{ gives } \sigma = 0^2 - 1^2 = -1, \textit{ typical of points with spacelike}$$

separation from the origin. And that—since the figure maps to itself under the Lorentz transformations that

- *describe the symmetry structure of spacetime*
- *describe the relationships among inertial observers*

—these classifications are Lorentz-invariant, shared by all *inertial observers.*

Calculation would establish what is in fact made obvious already at (206): the $K^{\pm}(\beta)$ are precisely the *eigenvalues* of $\mathbb{\Lambda}(\beta)$.[115] Nor are we surprised that the associated eigenvectors are *null* vectors, since

$$(x,x) \rightarrow (Kx,Kx) = (x,x) \quad \text{entails} \quad (x,x) = 0$$

[115] We note in passing that $K^-(\beta) = [K^+(\beta)]^{-1} = K^+(-\beta)$.

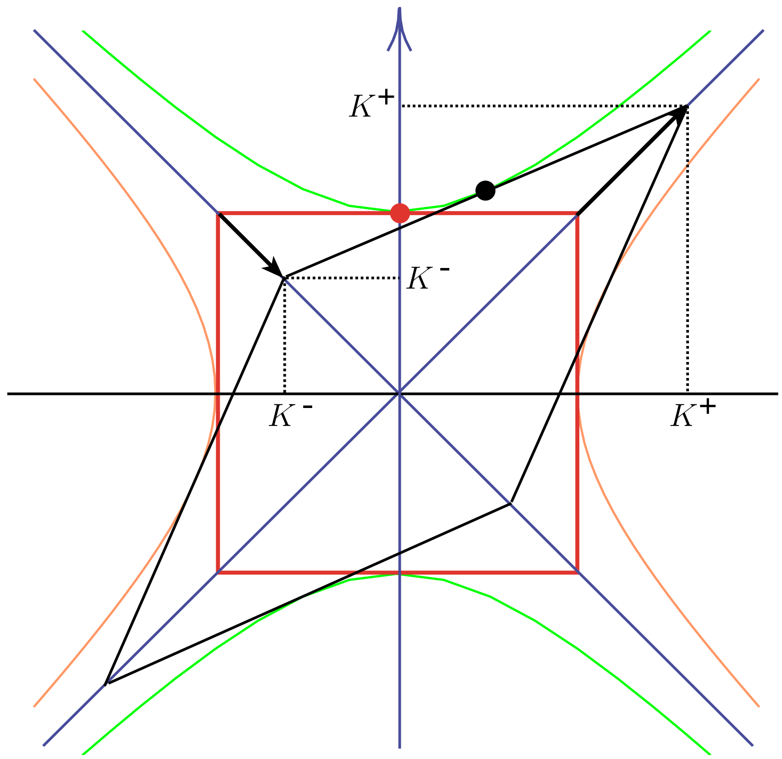

FIGURE 55: *Inertial observer O inscribes a "unit square" $\square$, with lightlike vertices, on her spacetime diagram. $\mathbb{\Lambda}(\beta)$ stretches one diagonal by the factor K^+, and shrinks the other by the factor K^-. That individual points "slide along isometric curves" is illustrated here by the motion $\bullet \to \bullet$ of a point of tangency. Corresponding sides of $\square$ and its transform have different Euclidean lengths, but identical Lorentzian lengths. Curiously, it follows from $K^+K^- = 1$ that $\square$ and its transform have identical Euclidean* areas.[116,117]

The upshot of preceding remarks is illustrated above, and elaborated in the figure on the next page, where I have stated in the caption but here emphasize once again that *such figures, though drawn on the Euclidean page, are to be read as inscriptions on 2-dimensional spacetime*. The distinction becomes especially clear when one examines Figure 57.

---

[116] PROBLEM 43.

[117] Some authors stress the utility in special relativity of what they call the "k-calculus:" see, for example, Hermann Bondi, *Relativity and Common Sense: A New Approach to Einstein* (1962), pages 88–121 and occasional papers in the American Journal of Physics. My K-notation is intended to establish contact with that obscure tradition.

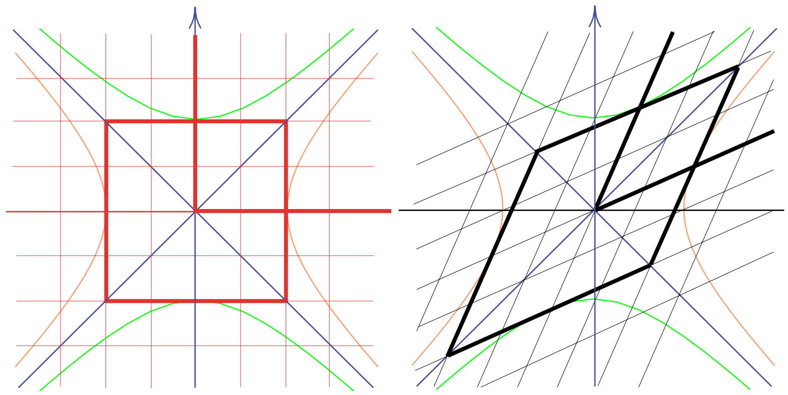

FIGURE 56: *Elaboration of the preceding figure.* O *has inscribed a Cartesian gridwork on spacetime. On the right is shown the Lorentz transform of that coordinate grid. Misner, Thorne & Wheeler* (Gravitation (1973), page 11) *have referred in this connection to the "collapse of the egg crate," though that picturesque terminology is somewhat misleading: egg crates preserve side-length when they collapse, while the present mode of collapse preserves Euclidean area. Orthogonality, though obviously violated in the Euclidean sense, is preserved in the Lorentzian sense . . . which is, in fact, the only relevant sense, since the figure is inscribed not on the Euclidean plane but on 2-dimensional spacetime. Notice that tangents to isometric curves remain in each case tangent to the same such curve. The entire population of isometric curves (see again Figure 54) can be recovered as the population of envelopes of the grid lines, as generated by allowing* β *to range over all allowed values* $(-1 < \beta < +1)$.

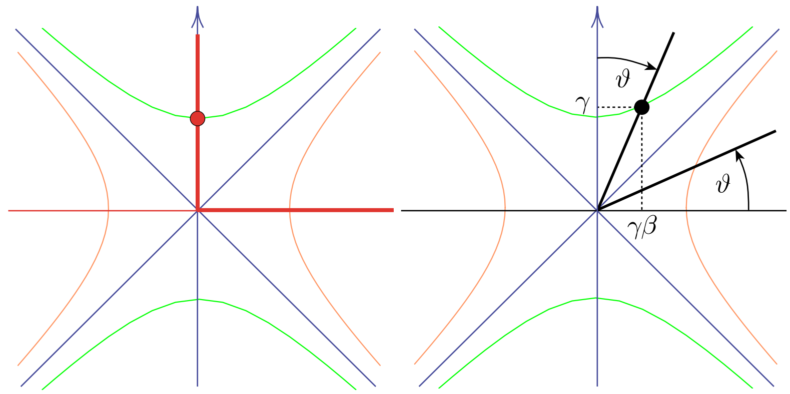

FIGURE 57: *O writes $(ct, 0)$ to describe the "t^{th} tick of her clock." Working from (201) we find that O assigns coordinates $(\gamma t, \gamma\beta t)$ to that same event. The implication is that the (Euclidean) angle ϑ subtended by*

- *O's time axis and*
- *O's representation of O's time axis*

can be described

$$\tan\vartheta = \beta$$

The same angle, by a similar argument, arises when one looks to O's representation of O's space axis. One could, with this information, construct the instance of Figure 56 which is appropriate to any prescribed β-value. Again I emphasize that—their Euclidean appearance notwithstanding—O and O are in agreement that O's coordinate axes are normal in the Lorentzian sense.[118]

We are in position now to **four points of fundamental physical significance**, of which three are temporal, and one spatial. The points I have in mind will be presented in a series of figures, and developed in the captions:

[118] PROBLEM 44.

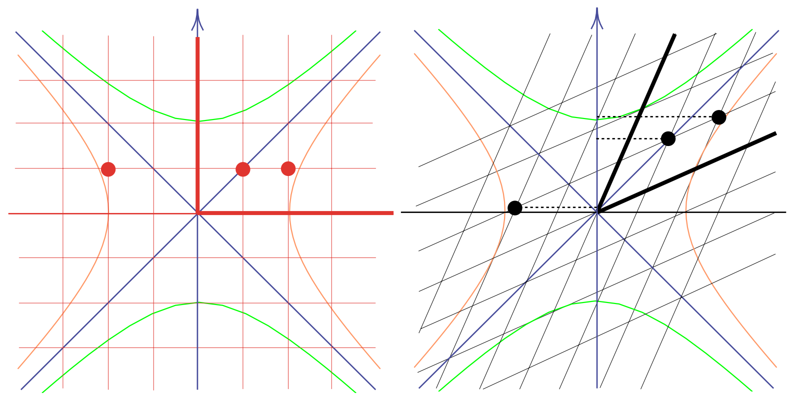

FIGURE 58: **Breakdown of non-local simultaneity.** *O sees three spatially-separated events to be simultaneous. O, on the other hand, assigns distinct x^0-coordinates to those same events (see the figure on the right), which he considers to be non-simultaneous/sequential. It makes relativistic good sense to use the word "simultaneous" only in reference to events which (like the birth of twins) occur at the same moment and at the same spatial point. The Newtonian concept of* "instantaneous action at a distance"*—central to his "Universal Law of Gravitation" but which, on philosophical grounds, bothered not only Newton's contemporaries but also Newton himself —has been rendered relativistically untenable: interactions, in any relativistically coherent physics, have become necessarily local, dominated by what philosophers call the "Principle of Contiguity." They have, in short, become collision-like events, the effects of which propagate like a contagion: neighbor infects neighbor. If "particles" are to participate in collisions they must necessarily be held to be pointlike in the mathematical sense (a hard idealization to swallow), lest one acquire an obligation to develop a physics of processes* interior *to the particle. The language most natural to physics has become field theory—a theory in which all interactions are local field-field interactions, described by partial differential equations.*

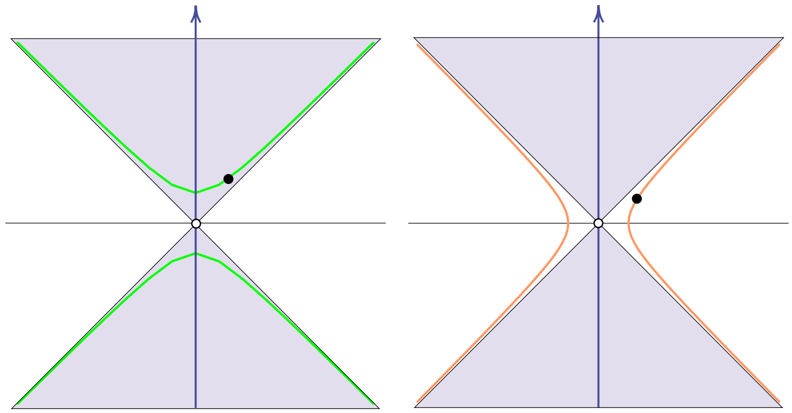

FIGURE 59: **Conditional covariance of causal sequence.** *At left: diverse inertial observers all place the event* • *on a sheet of the isometric hyperboloid that is confined to the interior of the forward lightcone, and all agree that* • *lies "in the future" of the origin* ∘. *But if (as at the right)* • *is separated from* ∘ *by a spacelike interval; i.e., if* • *lies outside the lightcone at* ∘, *then some observers see* • *to lie in the future of* ∘, *while other observers see* • *to lie in its past. In the latter circumstance it is impossible to develop an agreed-upon sense of causal sequence. Generally: physical events at a point* **p** *can be said to have been "caused" only by events that lie in/on the lightcone that extends backward from* **p**, *and can themselves influence only events that lie in/on the lightcone that extends forward from* **p**. *In electrodynamics it will emerge that (owing to the absence of "photon mass terms") effects propagate* <u>*on*</u> *the lightcone. Recent quantum mechanical experiments (motivated by the "EPR paradox") are of great interest because they have yielded results that appear to be "acausal" in the sense implied by preceding remarks: the outcome of a quantum coin-flip at* **p** *predetermines the result of a similar measuremennt at* **q** even though the interval separating **q** from **p** is spacelike.

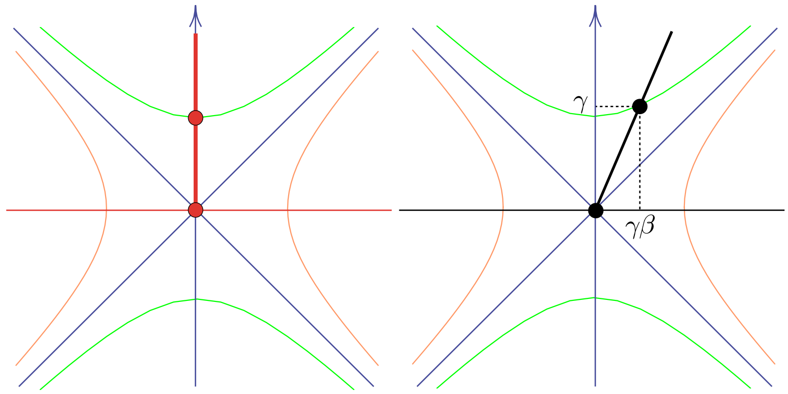

FIGURE 60: **Time dilation.** *Inertial observer O assigns duration x^0 to the interval separating "successive ticks $\bullet\ldots\bullet$ of her clock." A second observer O, in motion relative to O, assigns to those same events (see again Figure 57) the coordinates*

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \textit{and} \quad \begin{pmatrix} x^0 \\ x^1 \end{pmatrix} = \begin{pmatrix} \gamma x^0 \\ \gamma\beta x^0 \end{pmatrix}$$

He assigns the same Lorentzian value to the squared length of the spacetime interval $\bullet\ldots\bullet$ that O assigned to $\bullet\ldots\bullet$

$$(\gamma x^0)^2 - (\gamma\beta x^0)^2 = (x^0)^2 - (0)^2$$

but reports that the 2^{nd} tick occurred at time

$$x^0 = \gamma x^0 > x^0$$

In an example discussed in every text (see, e.g., Taylor & Wheeler, Spacetime Physics (1966), §42) *the "ticking" is associated with the lifetime of an unstable particle—typically a muon—which (relative to the tabulated rest-frame value) seems dilated to observers who see the particle to be in motion.*

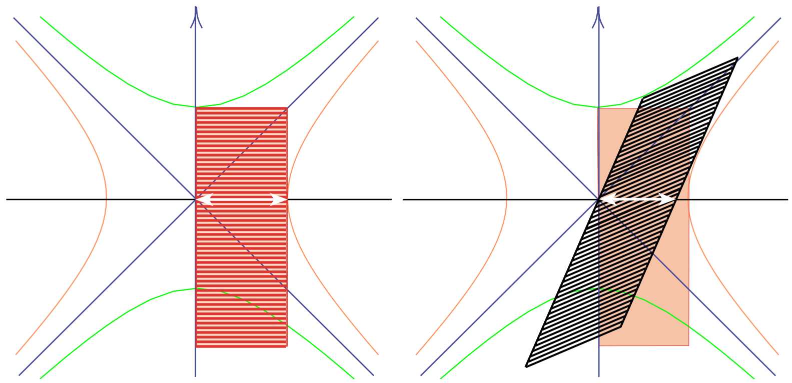

FIGURE 61: **Lorentz contraction.** *This is often looked upon as the flip side of time dilation, but the situation as it pertains to* spatial *intervals is—owing to the fact that <u>metersticks persist, and are therefore not precise analogs of clockticks</u>—a bit more subtle. At left is* O*'s representation of a meterstick sitting there, sitting there, sitting there ... and at right is* O*'s representation of that same construction. The white arrows indicate that while* O *and* O *have the same thought in mind when they talk about the "length of the meterstick" (length of the spatial interval that separates one end from the other <u>at an instant</u>) they are—because they assign distinct meanings to "at an instant"—actually talking about different things. Detailed implications are developed in the following figure.*

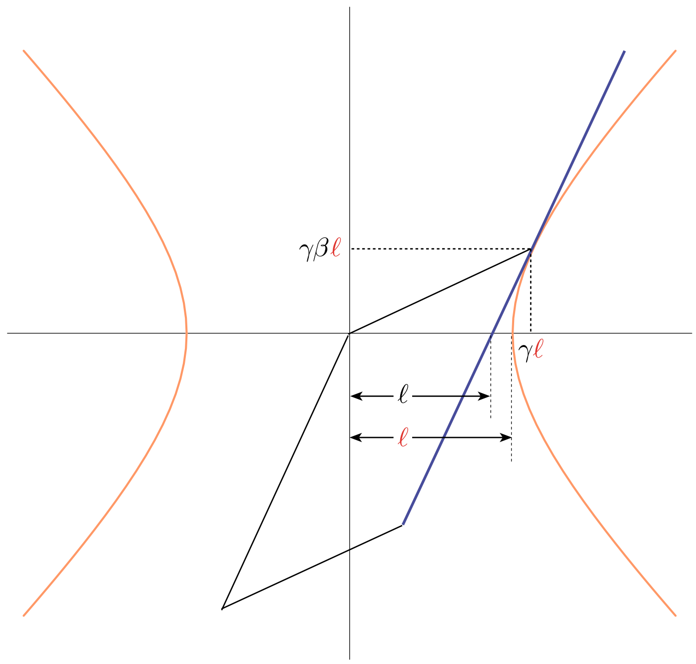

FIGURE 62: **Lorentz contraction (continued).** *When observers speak of the "length of a meterstick" they are really talking about what they perceive to be the width of the "ribbon" which such an extended object inscribes on spacetime. This expanded detail from the preceding figure shows how it comes about that the meterstick which O sees to be at rest, and to which she assigns length ℓ, is assigned length*

$$\ell = \gamma^{-1}\ell < \ell$$

by O, who sees the meterstick to be in uniform motion. This familiar result poses, by the way, a problem which did not escape Einstein's attention, and which contributed to the development of general relativity: The circumference of a rigidly rotating disk has become too short to go all the way around![119]

Prior to Einstein's appearance on the scene (1905) it was universally held that time dilation and "Lorentz-FitzGerald contraction" were *physical* effects, postulated to account for the null result of the Michelson-Morley experiment, and attributed to the interaction of physical clocks and physical metersticks with the physical "æther" through which they were being transported. Einstein

[119] See J. Stachel, "Einstein and the rigidly rotating disk" in A. Held (editor), *General Relativity & Gravitation* (1980), Volume 1, page 1. H. Arzeliès, in *Relativistic Kinematics* (1966), devotes an entire chapter to the disk problem and its relatives.

(with his trains and lanterns) argued that such effects are not "physical," in the sense that they have to do with the properties of "stuff"... but "metaphysical" (or should one say: *pre*-physical?)—artifacts of the *operational procedures by which one assigns meaning* to lengths and times. In preceding pages I have, in the tradition established by Minkowski, espoused a third view: I have represented all such effects are reflections of the circumstance (brought first to our attention by electrodynamics) that *the hyperbolic geometry of spacetime is a primitive fact of the world, embraced by all inertial observers ... and written into the design of all possible physics*.

> REMARK: It would be nice if things were so simple (which in leading approximation they are), but when we dismissed Newton's Law of Universal Gravitation as "relativistically untenable" we acquired a question ("How did the Newtonian theory manage to serve so well for so long?") and an obligation—the development of a "field theory of gravitation." The latter assignment, as discharged by Einstein himself, culminated in the invention of "general relativity" and the realization that it is—except in the approximation that gravitational effects can be disregarded—*incorrect* to speak with global intent about the "hyperbolic geometry of spacetime." The "geometry of spacetime" is "hyperbolic" only in the same approximate/tangential sense that vanishingly small regions inscribed on (say) the unit sphere become "Euclidean."

6. Lorentz transformations in 4-dimensional spacetime. The transition from toy 2-dimensional spacetime to physical 4-dimensional spacetime poses an enriched algebraic problem

$$\mathbb{\Lambda} = \exp\begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix} \tag{196.1}$$

$$\downarrow$$

$$\mathbb{\Lambda} = \exp\begin{pmatrix} 0 & A_1 & A_2 & A_3 \\ A_1 & 0 & a_3 & -a_2 \\ A_2 & -a_3 & 0 & a_1 \\ A_3 & a_2 & -a_1 & 0 \end{pmatrix} \tag{193}$$

and brings to light a physically-important point or two which were overlooked by Einstein himself. The algebraic details are, if addressed with a measure of elegance, of some intrinsic interest[120]... but I must here be content merely to outline the most basic facts, and to indicate their most characteristic kinematic/physical consequences. Consider first the

[120] See ELEMENTS OF RELATIVITY (1966).

$\boxed{\text{CASE } A_1 = A_2 = A_3 = 0}$ in which $\mathbb{\Lambda}$ possesses only space/space generators.[121] Then

$$\mathbb{\Lambda} = \exp\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathbb{A} & \\ 0 & & & \end{pmatrix}$$

where

$$\mathbb{A} \equiv \begin{pmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix} \text{ is real and antisymmetric}$$

It follows quite easily that

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathbb{R} & \\ 0 & & & \end{pmatrix} \tag{208}$$

where $\mathbb{R} \equiv e^{\mathbb{A}}$ is a 3×3 *rotation matrix*. The action of such a $\mathbb{\Lambda}$ can be described

$$\begin{pmatrix} x^0 \\ \boldsymbol{x} \end{pmatrix} \longrightarrow \begin{pmatrix} x^0 \\ \boldsymbol{x} \end{pmatrix} = \begin{pmatrix} x^0 \\ \mathbb{R}\boldsymbol{x} \end{pmatrix}$$

as a spatial rotation that leaves time coordinates unchanged. Look to the case $a_1 = a_2 = 0,\ a_3 = \phi$ and use the *Mathematica* command `MatrixExp[Λ]` to obtain

$$\mathbb{\Lambda} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & \sin\phi & 0 \\ 0 & -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with the evident implication that in the general case such a Lorentz matrix describes a lefthanded rotation through angle $\phi = \sqrt{\boldsymbol{a\cdot a}}$ about the unit vector $\boldsymbol{\lambda} \equiv \hat{\boldsymbol{a}}$.[122] Such Lorentz transformations contain no allusion to $\boldsymbol{v}$ and have *no properly kinematic significance*: O simply stands beside us, using her clock (indistinguishable from ours) and her rotated Cartesian frame to "do physics." What we have learned is that

Spatial rotations are Lorentz transformations

of a special type (a type for which the 2-dimensional theory is too impoverished to make provision). The associated Lorentz matrices will be notated $\mathbb{R}(\phi, \boldsymbol{\lambda})$.

Look next to the complementary . . .

---

[121] "Time/time" means ^{0} appears twice, "time/space" and "space/time" mean that ^{0} appears once, "space/space" means that ^{0} is absent.

[122] See CLASSICAL DYNAMICS (1964/65), Chapter 1, pages 83–89 for a simple account of the detailed argument.

$\boxed{\text{CASE } a_1 = a_2 = a_3 = 0}$ in which $\mathbb{\Lambda}$ possesses only time/space generators. Here (as it turns out) $\mathbb{\Lambda}$ *does* possess kinematic significance. The argument which (on page 139) gave

$$A = \tanh^{-1}\beta \quad \text{with} \quad \beta = v/c$$

now gives

$$\boldsymbol{A} = \tanh^{-1}\beta \cdot \hat{\boldsymbol{v}}$$

while the argument which (on pages 138–139) gave

$$\mathbb{\Lambda} = \exp\left\{\tanh^{-1}\beta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right\} = \begin{pmatrix} \gamma & v\gamma/c \\ v\gamma/c & \gamma \end{pmatrix}$$

now gives

$$\mathbb{\Lambda} = \exp\left\{\tanh^{-1}\beta \begin{pmatrix} 0 & \hat{v}_1 & \hat{v}_2 & \hat{v}_3 \\ \hat{v}_1 & 0 & 0 & 0 \\ \hat{v}_2 & 0 & 0 & 0 \\ \hat{v}_3 & 0 & 0 & 0 \end{pmatrix}\right\}$$
$$= \begin{pmatrix} \gamma & v_1\gamma/c & v_2\gamma/c & v_3\gamma/c \\ v_1\gamma/c & 1+(\gamma-1)v_1v_1/v^2 & (\gamma-1)v_1v_2/v^2 & (\gamma-1)v_1v_3/v^2 \\ v_2\gamma/c & (\gamma-1)v_2v_1/v^2 & 1+(\gamma-1)v_2v_2/v^2 & (\gamma-1)v_2v_3/v^2 \\ v_3\gamma/c & (\gamma-1)v_3v_1/v^2 & (\gamma-1)v_3v_2/v^2 & 1+(\gamma-1)v_3v_3/v^2 \end{pmatrix}$$

Such Lorentz matrices will be notated

$$= \mathbb{\Lambda}(\boldsymbol{\beta}) \tag{209}$$
$$\boldsymbol{\beta} \equiv \boldsymbol{v}/c$$

They give rise to Lorentz transformations $x \longrightarrow x = \mathbb{\Lambda}(\boldsymbol{\beta})x$ which are "pure" (in the sense "rotation-free") and are called "boosts." The construction (208) *looks* complicated, but in fact it possesses *precisely the structure that one might (with a little thought) have anticipated*. For (209) supplies[123]

$$\left.\begin{aligned} t &= \gamma t + (\gamma/c^2)\,\boldsymbol{v}\cdot\boldsymbol{x} \\ \boldsymbol{x} &= \boldsymbol{x} + \left\{\gamma t + (\gamma-1)\,(\boldsymbol{v}\cdot\boldsymbol{x})/v^2\right\}\boldsymbol{v} \end{aligned}\right\} \tag{210.1}$$

and if we resolve $\boldsymbol{x}$ and $\boldsymbol{x}$ into components which are parallel/perpendicular to the velocity $\boldsymbol{v}$ with which O sees O to be gliding by

$$\boldsymbol{x} = \boldsymbol{x}_\perp + \boldsymbol{x}_\parallel \quad \text{with} \quad \begin{cases} \boldsymbol{x}_\parallel \equiv (\boldsymbol{x}\cdot\hat{\boldsymbol{v}})\,\hat{\boldsymbol{v}} \equiv x_\parallel \hat{\boldsymbol{v}} \\ \boldsymbol{x}_\perp \equiv \boldsymbol{x} - \boldsymbol{x}_\parallel \end{cases}$$
$$\boldsymbol{x} = \boldsymbol{x}_\perp + \boldsymbol{x}_\parallel \quad \text{with} \quad \begin{cases} \boldsymbol{x}_\parallel \equiv (\boldsymbol{x}\cdot\hat{\boldsymbol{v}})\,\hat{\boldsymbol{v}} \equiv x_\parallel \hat{\boldsymbol{v}} \\ \boldsymbol{x}_\perp \equiv \boldsymbol{x} - \boldsymbol{x}_\parallel \end{cases}$$

[123] PROBLEM 45, 46.

then (210.1) can be written (compare (203))

$$\left.\begin{aligned}\begin{pmatrix} t \\ x_\parallel \end{pmatrix} &= \gamma \begin{pmatrix} 1 & v/c^2 \\ v & 1 \end{pmatrix} \begin{pmatrix} t \\ x_\parallel \end{pmatrix} \\ \boldsymbol{x}_\perp &= \boldsymbol{x}_\perp \end{aligned}\right\} \tag{210.2}$$

And in the Galilean limit we recover

$$\begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ v_1 & 1 & 0 & 0 \\ v_2 & 0 & 1 & 0 \\ v_3 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{210.3}$$

GENERAL CASE Having discussed the 3-parameter family of rotations $\mathbb{R}(\phi, \boldsymbol{\lambda})$ and the 3-parameter family of boosts $\Lambda(\boldsymbol{\beta})$ the questions arises: *What can one say in the general 6-parameter case*

$$\Lambda = e^{\mathbb{A}}$$

It is—given the context in which the question was posed—natural to write

$$\mathbb{A} = \mathbb{J} + \mathbb{K}$$

with

$$\mathbb{J} \equiv \begin{pmatrix} 0 & A_1 & A_2 & A_3 \\ A_1 & 0 & 0 & 0 \\ A_2 & 0 & 0 & 0 \\ A_3 & 0 & 0 & 0 \end{pmatrix} \equiv \sum_{i=1}^{3} A_i \, \mathbb{J}_i$$

$$\mathbb{K} \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & a_3 & -a_2 \\ 0 & -a_3 & 0 & a_1 \\ 0 & a_2 & -a_1 & 0 \end{pmatrix} \equiv \sum_{i=1}^{3} a_i \, \mathbb{K}_i$$

and one might on this basis be tempted to write $\Lambda = e^{\mathbb{K}} \cdot e^{\mathbb{J}}$, giving

$$\Lambda_{\text{general}} = (\text{rotation}) \cdot (\text{boost}) \tag{211}$$

Actually, a representation theorem of the form (211) is available, but the argument which here led us to (211) is incorrect: one can write

$$e^{\mathbb{J}+\mathbb{K}} = e^{\mathbb{K}} \cdot e^{\mathbb{J}} \quad \text{if and only if } \mathbb{J} \text{ and } \mathbb{K} \ \underline{\text{commute}}$$

and in the present instance we (by computation) have

$$\begin{aligned} \big[\mathbb{J}, \mathbb{K}\big] &= -\sum_{i=1}^{3} (\boldsymbol{A} \times \boldsymbol{a})_i \, \mathbb{J}_i \\ &= \mathbb{O} \quad \text{if and only if } \boldsymbol{A} \text{ and } \boldsymbol{a} \text{ are } \underline{\text{parallel}} \end{aligned} \tag{212}$$

More careful analysis (which requires some fairly sophisticated algebraic machinery[124]) leads back again to (211), but shows the boost and rotational factors of $\mathbb{\Lambda}$ to be different from those initially contemplated. I resist the temptation to inquire more closely into the correct factorization of $\mathbb{\Lambda}$, partly because I have other fish to fry ... but mainly because I have already in hand the facts needed to make my major point, which concerns the *composition of boosts in 4-dimensional spacetime*. It follows immediately from (208) that

$$(\text{rotation})\cdot(\text{rotation}) = \underbrace{(\text{rotation})}_{\text{specific description poses a non-trivial but merely technical (algebraic) problem}} \tag{213.1}$$

It might—on analogical grounds—appear plausible therefore that

$$(\text{boost})\cdot(\text{boost}) = (\text{boost})$$

but (remarkably!) <u>this is not the case</u>: actually

$$= (\underline{\text{rotation}})\cdot(\text{boost}) \tag{213.2}$$

Detailed calculation shows more specifically that

$$\mathbb{\Lambda}(\boldsymbol{\beta}_2)\cdot\mathbb{\Lambda}(\boldsymbol{\beta}_1) = \mathbb{R}(\phi,\boldsymbol{\lambda})\,\mathbb{\Lambda}(\boldsymbol{\beta}) \tag{214.0}$$

where

$$\boldsymbol{\beta} = \frac{\big[1+(\beta_2/\beta_1)(1-\frac{1}{\gamma_1})\cos\omega\big]\boldsymbol{\beta}_1 + \frac{1}{\gamma_1}\boldsymbol{\beta}_2}{1+\beta_1\beta_2\cos\omega} \tag{214.1}$$

$$\boldsymbol{\lambda} = \text{unit vector parallel to } \boldsymbol{\beta}_2\times\boldsymbol{\beta}_1 \tag{214.2}$$

$$\omega = \text{angle between } \boldsymbol{\beta}_1 \text{ and } \boldsymbol{\beta}_2 \tag{214.3}$$

$$\phi = \tan^{-1}\left\{\frac{\epsilon\sin\omega}{1+\epsilon\cos\omega}\right\} \tag{214.4}$$

$$\epsilon = \sqrt{(\gamma_1-1)(\gamma_2-1)/(\gamma_1+1)(\gamma_2+1)} \tag{214.5}$$

and where β_1, β_2, γ_1 and γ_2 have the obvious meanings. One is quite unprepared by 2-dimensional experience for results which are superficially so ugly, and which are undeniably so complex. The following points should be noted:

1. Equation (214.1) is the 4-dimensional *velocity addition formula*. Looking with its aid to $\boldsymbol{\beta}\cdot\boldsymbol{\beta}$ we obtain the *speed addition formula*

$$\beta = \frac{\sqrt{\beta_1^2+\beta_2^2+2\beta_1\beta_2\cos\omega-(\beta_1\beta_2\sin\omega)^2}}{1+\beta_1\beta_2\cos\omega} \tag{215}$$

$$\Downarrow$$

$$\beta \leqslant 1 \quad \text{if } \beta_1 \leqslant 1 \text{ and } \beta_2 \leqslant 1$$

according to which (see the following figure) *one cannot, by composing velocities, escape from the c-ball*. Note also that

$$\downarrow$$

$$\beta = \frac{\beta_1+\beta_2}{1+\beta_1\beta_2} \quad \text{in the collinear case: } \omega = 0$$

---

[124] The requisite machinery is developed in elaborate detail in ELEMENTS OF SPECIAL RELATIVITY (1966).

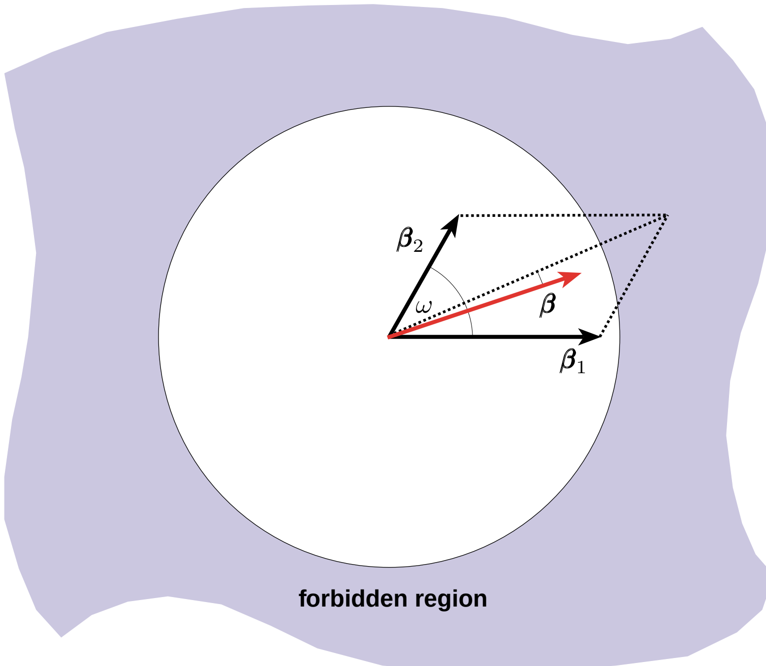

FIGURE 63: *$\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, if not collinear, span a plane in 3-dimensional $\boldsymbol{\beta}$-space. The figure shows the intersection of that plane with what I call the "c-ball," defined by the condition $\beta^2 = 1$. The placement of $\boldsymbol{\beta}$ is given by (214.1). Notice that, while $\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2$ falls into the forbidden exterior of the c-ball, $\boldsymbol{\beta}$ does not. Notice also that $\boldsymbol{\beta}$ lies on the $\boldsymbol{\beta}_1$-side of $\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2$, from which it deviates by an angle that turns out to be precisely the ϕ that enters into the design of the rotational factor $\mathbb{R}(\phi, \boldsymbol{\lambda})$.*

which is in precise conformity with the familiar 2-dimensional formula (204).

2. It is evident in (214.1) that $\boldsymbol{\beta}$ depends *asymmetrically* upon $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$. Not only is $\boldsymbol{\beta} \neq \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2$, is its *not even parallel* to $\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2$, from which it deviates by an angle that turns out to be precisely the ϕ encountered already—in quite another connection—at (214.4). The asymmetry if the situation might be summed up in the phrase "$\boldsymbol{\beta}_1$ predominates." From this circumstance one acquires interest in the angle Ω between $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_1$: we find

$$\begin{aligned} \Omega\ &= \tan^{-1}\left\{\frac{\beta_2 \sin\omega}{\gamma_1(\beta_1 + \beta_2\cos\omega)}\right\} \qquad (216)\\ &\downarrow \\ \Omega_0 &= \tan^{-1}\left\{\frac{\beta_2 \sin\omega}{\beta_1 + \beta_2\cos\omega}\right\} \quad \text{in the non-relativistic limit} \end{aligned}$$

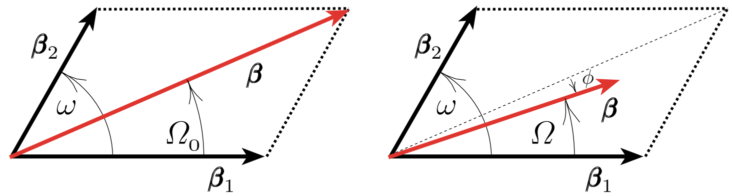

FIGURE 64: *At left: Galilean composition of non-collinear velocities. At right: its Lorentzian counterpart, showing the sense in which "$\boldsymbol{\beta}_1$ predominates." Evidently*

$$\Omega_{\text{relativistic}} = \Omega_0 + \phi \leqslant \Omega_0$$

calculations which are elementary in the Galilean case (see the figure) but become a little tedious in the relativistic case.[125] Asymmetry effects become most pronounced in the ultra-relativistic limit. Suppose, for example, that $\beta_1 = 1$: then $\Omega \downarrow 0$ and

$$\boldsymbol{\beta} \to \boldsymbol{\beta}_1, \text{irrespective of the value assigned to } \boldsymbol{\beta}_2!$$

More physically,[126] suppose $\beta_1 < 1$ but $\beta_2 = 1$: then

$$\Omega = \tan^{-1}\left\{\sqrt{1-\beta_1^2}\,\frac{\sin\omega}{\beta_1+\cos\omega}\right\}$$

The first occurrence of this formula is in §7 of Einstein's first relativity paper (1905), where it is found to provide the relativistic correction to the classic "law of aberration."[127]

3. It is a corollary of (215) that

$$\gamma = \gamma_1\gamma_2\big[1+\beta_1\beta_2\cos\omega\big]$$

which gives back (205) in the collinear case.

---

[125] See page 87 in the notes just cited.

[126] I say "more physically" because $\beta = 1$ cannot pertain to an "observer" (though it can pertain to the flight of a massless particle): while it does make sense to ask what an observer in motion (with respect to us) has to say about the lightbeam to which we assign a certain direction of propagation, it makes no sense to ask what the lightbeam has to say about the observer!

[127] "Aberration" is the name given by astronomers to the fact that "fixed stars" are seen to trace small ellipses in the sky, owing to the earth's annual progress along its orbit. See page 17 in W. Pauli's classic *Theory of Relativity* (first published in 1921, when Pauli was only twenty-one years old; reissued with a few additional notes in 1958) or P. G. Bergmann, *Introduction to the Theory of Relativity* (1942), pages 36–38.

4. In the small-velocity approximation (213.1) and (213.4) give

$$\begin{aligned}\boldsymbol{v} &= \boldsymbol{v}_1 + \boldsymbol{v}_2 - \Big[\tfrac{1}{2}\beta_1\beta_2\cos\omega\cdot\boldsymbol{v}_1 + \big(\tfrac{1}{2}\beta_1^2 + \beta_1\beta_2\cos\omega\big)\boldsymbol{v}_2\Big] + \cdots\\ \phi &= \tfrac{1}{4}\beta_1\beta_2\sin\omega + \cdots\end{aligned}$$

according to which all "relativistic correction terms" are of 2^{nd} order.

The presence of the $\mathbb{R}$-factor on the right side of (213)—*i.e.*, the fact that *rotations arise when one composes non-collinear boosts*—can be traced to the following algebraic circumstance:

$$\big[\,\mathbb{J}_1, \mathbb{K}_2\big] = -\mathbb{J}_3 = \big[\,\mathbb{J}_2, \mathbb{K}_1\big] \tag{217.1}$$

$$\big[\mathbb{K}_1, \mathbb{K}_2\big] = -\mathbb{K}_3 \tag{217.2}$$

$$\big[\,\mathbb{J}_1,\,\mathbb{J}_2\big] = +\mathbb{K}_3 \tag{217.3}$$

—each of which remains valid under cyclic index permutation. Equations (217.1) are but a rewrite of (212). The compositional closure (213.1) to the rotations can be attributed to the fact that it is a $\mathbb{K}$ that stands on the right side of (217.2). The fact (213.2) that *the set of boosts is not compositionally closed* arises from the circumstance that it is again a $\mathbb{K}$—not, as one might have expected, a $\mathbb{J}$—that stands on right side of (217.3).

The essential presence of the rotational $\mathbb{R}$-factor on the right side of (214) was discovered by L. H. Thomas (1926: relativity was then already 21 years old), whose motivation was not mathematical/kinematic, but intensely physical: Uhlenbeck & Goudsmit had sought (1925) to derive fine details of the hydrogen spectrum from the assumption that the electron in the Bohr atom possesses intrinsic "spin"... but had obtained results which were invariably off by a factor of 2. Thomas—then a post-doctoral student at the Bohr Institute, and for reasons to which I will return in a moment—speculated that a "relativistic correction" would resolve that problem. Challenged by Bohr to develop the idea (for which neither Bohr nor his associate Kramers held much hope), Thomas "that weekend" argued as follows: (i) A proton •, pinned to the origin of an inertial frame, sees an electron • to be revolving with angular velocity Ω_{orbital} on a circular Bohr orbit of radius R. (ii) Go to the frame of the non-inertial observer who is "riding on the electron" (and therefore sees • to be in circular motion): do this by

> going to the frame of the inertial observer who is instantaneously comoving with • at time $t_0 = 0$, then...
>
> **boosting** to the frame of the inertial observer who is instantaneously comoving with • at time $t_1 = \tau$, then...
>
> **boosting** to the frame of the inertial observer who is instantaneously comoving with • at time $t_2 = 2\tau$, then...
>
> ⋮
>
> **boosting** to the frame of the inertial observer who is instantaneously comoving with • at time $t = N\tau$

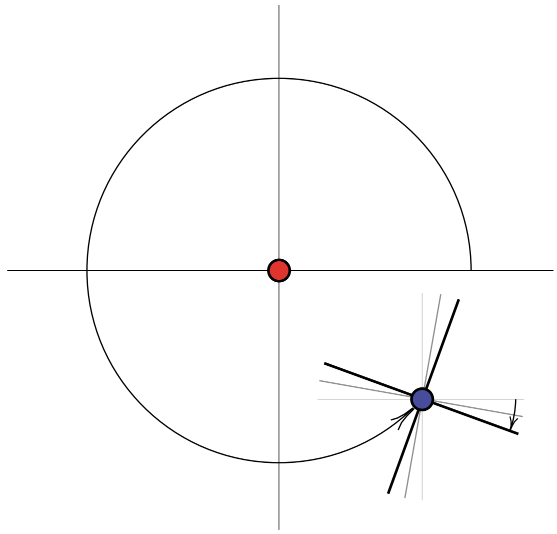

FIGURE 65: *Thomas precession of the non-inertial frame of an observer • in circular orbit about an inertial observer •. In celestial mechanical applications the effect is typically so small (on the order of seconds of arc per century) as to be obscured by dynamical effects. But in the application to (pre-quantum mechanical) atomic physics that was of interest to Thomas the precession becomes quite brisk (on the order of* $\sim 10^{12}$ *Hz.).*

and by taking that procedure to the limit $\tau \downarrow 0$, $N = t/\tau \uparrow \infty$. One arrives thus at method for *Lorentz transforming to the frame of an accelerated observer*. The curvature of the orbit means, however, that successive boosts are not collinear; rotational factors intrude at each step, and have a cumulative effect which (as detailed analysis[128] shows) can be described

$$\begin{aligned}\tfrac{d\phi}{dt} \equiv \Omega_{\text{Thomas}} &= (\gamma - 1)\Omega_{\text{orbital}}\\ &= \tfrac{1}{2}\beta^2\Omega_{\text{orbital}}\big\{1 + \tfrac{3}{4}\beta^2 + \tfrac{15}{24}\beta^4 + \cdots\big\}\end{aligned}$$

in the counterrotational sense (see the figure). It is important to notice that this *Thomas precessional effect* is of relativistic kinematic origin: it does not

[128] See §103 in E. F. Taylor & J. A. Wheeler, *Spacetime Physics* (1963) or pages 95–116 in the notes previously cited.[122] Thomas' own writing—"The motion of the spinning electron," Nature **117**, 514 (1926); "The kinematics of an electron with an axis," Phil. Mag. **3**, 1 (1927); "Recollections of the discovery of the Thomas precessional frequency" in G. M. Bunce (editor), *High Energy Spin Physics*–1982,AIP Conference Proceedings No. 95 (1983)—have never seemed to me to be particularly clear. See also J. Frenkel, "Die Elektrodynamic des rotierenden Elektrons," Z. für Physik **37**, 243 (1926).

arise from impressed forces. (*iii*) Look now beyond the kinematics to the dynamics: from •'s viewpoint the revolving • is, in effect, a current loop, the generator of a magnetic field $\boldsymbol{B}$. Uhlenbeck & Goudsmit had assumed that the electron possesses a magnetic moment proportional to its postulated spin: such an electron senses the $\boldsymbol{B}$-field, to which it responds by precessing, acquiring precessional energy $\mathcal{E}_{\text{Uhlenbeck \& Goudsmit}}$. Uhlenbeck & Goudsmit worked, however, from a mistaken conception of "•'s viewpoint." The point recognized by Thomas is that when relativistic frame-precession is taken into account[129] one obtains

$$\mathcal{E}_{\text{Thomas}} = \tfrac{1}{2}\mathcal{E}_{\text{Uhlenbeck \& Goudsmit}}$$

—in good agreement with the spectroscopic data. This was a discovery of historic importance, for it silenced those (led by Pauli) who had dismissed as "too classical" the spin idea when it had been put forward by Krönig and again, one year later, by Uhlenbeck & Goudsmit: *"spin" became an accepted/fundamental attribute of elementary particles*.[130]

So much for the structure and properties of the Lorentz transformations . . . to which (following more closely in Minkowski's footsteps than Lorentz') we were led by analysis of the condition

$$\mathbb{\Lambda}^{\mathsf{T}}\mathbb{g}\,\mathbb{\Lambda} = \mathbb{g} \quad \text{everywhere} \tag{182}$$

which arose from <u>one natural interpretation</u> of the requirement that $\mathcal{X} \rightarrow \mathcal{X}$ preserve the form of Maxwell's equations . . . but to which Einstein himself was led by quite other considerations: Einstein—recall his trains/clocks/rods and lanterns—proceeded by operational/epistemological analysis of how inertial observers O and O, consistently with the most primitive principles of an idealized macroscopic physics, would *establish* the relationship between their coordinate systems. Einstein's argument was wonderfully original, and lent an air of "inescapability" to his conclusions . . . but (in my view) *must today be dismissed as irrelevant*, for <u>special relativity appears to remain effective in the</u>

---

[129] See pages 116–122 in ELEMENTS OF RELATIVITY (1966).

[130] Thomas precession is a relativistic effect which 2-dimensional theory is too impoverished to expose. Einstein himself missed it, and—so far as I am aware—never commented in print upon Thomas' discovery. Nor is it mentioned in Pauli/s otherwise wonderfully complete *Theory of Relativity*.[125] In 1969 I had an opportunity to ask Thomas himself how he had come upon his essential insight. He responded "Nothing is ever really new. I learned about the subject from Eddington's discussion [Eddington was in fact one of Thomas' teachers] of the relativistic dynamics of the moon—somewhere in his relativity book, which was then new. I'm sure the whole business—except for the application to Bohr's atom—was known to Eddington by 1922. Eddington was a smart man." Arthur Stanley Eddington's *The Mathematical Theory of Relativity* (1922) provided the first English-language account of general relativity. The passage to which Thomas evidently referred occurs in the middle of page 99 in the 2nd edition (1954), and apparently was based upon then-recent work by W. De Sitter.

deep microscopic realm where Einstein's operational devices/procedures (his "trains and lanterns") are—for quantum mechanical reasons—meaningless. Einstein built better than he knew—or could know ... but I'm ahead of my story. The Lorentz transformations enter into the statement of—but do not in and of themselves comprise—special relativity. The "meaning of relativity" is a topic to which I will return in §8.

7. Conformal transformations in N-dimensional spacetime.* We have seen that a second—and hardly less natural—interpretation of "Lorentz' question" gives rise not to (182) but to a condition of the form

$$\mathbb{W}^{\mathsf{T}}\mathbb{g}\,\mathbb{W} = \Omega\,\mathbb{g} \quad \text{everywhere} \tag{185.2}$$

where (as before)

$$\mathbb{g} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

My objective here is to describe the specific structure of the transformations $\mathfrak{X} \to \mathfrak{X}$ which arise from (185.2).

We begin as we began on page 132 (though the argument will not *not* lead to a proof of enforced linearity). If (185.2) is written

$$g_{\alpha\beta}W^\alpha{}_\mu W^\beta{}_\nu = g_{\mu\nu} \tag{218}$$

then (since the elements of $\mathbb{g}$ are constants) application of ∂_λ gives

$$g_{\alpha\beta}W^\alpha{}_{\lambda\mu}W^\beta{}_\nu + g_{\alpha\beta}W^\alpha{}_\mu W^\beta{}_{\nu\lambda} = g_{\mu\nu}\Omega_\lambda \tag{219}$$

where $W^\alpha{}_{\lambda\mu} \equiv \partial_\lambda W^\alpha{}_\mu = \partial^2 x^\alpha/\partial x^\lambda \partial x^\mu$ and $\Omega_\lambda \equiv \partial_\lambda\Omega$. Let functions $\Gamma_{\mu\nu\lambda}$ and $\varphi_\lambda \equiv \partial_\lambda\varphi$ be defined—deviously—as follows:

$$\Omega_\lambda \equiv 2\Omega\varphi_\lambda \tag{220}$$

$$g_{\alpha\beta}W^\alpha{}_\mu W^\beta{}_{\nu\lambda} \equiv \Omega\Gamma_{\mu\nu\lambda} \quad : \quad \nu\lambda\text{-symmetric} \tag{221}$$

Then (since the stipulated invertibility of $\mathfrak{X} \to \mathfrak{X}$ entails $\Omega = \sqrt{W} \neq 0$) equation (219) becomes

$$\Gamma_{\mu\nu\lambda} + \Gamma_{\nu\lambda\mu} = 2g_{\mu\nu}\varphi_\lambda$$

which by the "cyclic permutation argument" encountered on page 132 gives

$$\Gamma_{\lambda\mu\nu} = g_{\lambda\mu}\varphi_\nu + g_{\lambda\nu}\varphi_\mu - g_{\mu\nu}\varphi_\lambda \tag{222}$$

---

* It is the logic of the overall argument—certainly not pedagogical good sense!—that has motivated me to introduce this material (which will *not* be treated in lecture). First-time readers should skip directly to §7.

Now

$$W^\alpha{}_{\mu\nu} = \Gamma_{\lambda\mu\nu}\cdot\underbrace{\Omega M^\lambda{}_\beta g^{\beta\alpha}}_{=\,g^{\lambda\kappa}\,W^\alpha{}_\kappa \quad \text{by (218)}} \qquad \text{by (221)}$$

so by (222)

$$= \varphi_\mu W^\alpha{}_\nu + \varphi_\nu W^\alpha{}_\mu - g_{\mu\nu}\cdot g^{\lambda\kappa}\varphi_\lambda W^\alpha{}_\kappa \tag{223}$$

where the $\mu\nu$-symmetry is manifest. More compactly

$$= \Gamma^\kappa{}_{\mu\nu} W^\alpha{}_\kappa \tag{224}$$

where

$$\Gamma^\kappa{}_{\mu\nu} \equiv g^{\kappa\lambda}\Gamma_{\lambda\mu\nu}$$

Application of ∂_λ to (224) gives $W^\alpha{}_{\lambda\mu\nu} = \dfrac{\partial \Gamma^\kappa{}_{\mu\nu}}{\partial x^\lambda} W^\alpha{}_\kappa + \Gamma^\kappa{}_{\mu\nu} W^\alpha{}_{\lambda\kappa}$ which (since $W^{\bullet}{}_{\bullet\bullet}$, $W^{\bullet}{}_{\bullet\bullet\bullet}$ and $\Gamma^{\bullet}{}_{\bullet\bullet}$ are symmetric in their subscripts, and after relabling some indices) can be written

$$\begin{aligned} W^\alpha{}_{\lambda\mu\nu} &= \frac{\partial \Gamma^\beta{}_{\lambda\nu}}{\partial x^\mu} W^\alpha{}_\beta + \Gamma^\kappa{}_{\nu\lambda}\underbrace{W^\alpha{}_{\kappa\mu}}_{=\,\Gamma^\beta{}_{\kappa\mu} W^\alpha{}_\beta \quad \text{by (224)}} \\ &= \left\{\frac{\partial \Gamma^\beta{}_{\lambda\nu}}{\partial x^\mu} + \Gamma^\beta{}_{\kappa\mu}\Gamma^\kappa{}_{\nu\lambda}\right\} W^\alpha{}_\beta \end{aligned}$$

from which it follows in particular that

$$\begin{aligned} W^\alpha{}_{\lambda\mu\nu} - W^\alpha{}_{\lambda\nu\mu} &= \left\{\frac{\partial \Gamma^\beta{}_{\lambda\nu}}{\partial x^\mu} - \frac{\partial \Gamma^\beta{}_{\lambda\mu}}{\partial x^\nu} + \Gamma^\beta{}_{\kappa\mu}\Gamma^\kappa{}_{\nu\lambda} - \Gamma^\beta{}_{\kappa\nu}\Gamma^\kappa{}_{\mu\lambda}\right\} W^\alpha{}_\beta \\ &\equiv R^\beta{}_{\lambda\mu\nu} W^\alpha{}_\beta \end{aligned} \tag{225}$$

The preceding sequence of manipulations will, I fear, strike naive readers as an unmotivated jumble. But those with some familiarity with patterns of argument standard to differential geometry will have recognized that

- the quantities $W^\alpha{}_\mu$ transform as components of an α-parameterized set of covariant vectors;
- the quantities $\Gamma^\kappa{}_{\mu\nu}$ are components of[131] an *affine connection* to which (222) assigns a specialized structure;
- the α-parameterized equations (224) can be notated

$$D_\nu W^\alpha{}_\mu \equiv \partial_\nu W^\alpha{}_\mu - W^\alpha{}_\kappa \Gamma^\kappa{}_{\mu\nu} = 0$$

  according to which each of the vectors $W^\alpha{}_\mu$ has the property that its *covariant derivative*[129] vanishes;
- the 4th rank tensor $R^\beta{}_{\lambda\mu\nu}$ defined at (225) is just the *Riemann-Christoffel curvature tensor*,[129] to which a specialized structure has in this instance been assigned by (222).

[131] See again page 123.

But of differential geometry I will make explicit use only in the following—independently verifiable—facts: let

$$R_{\kappa\lambda\mu\nu} \equiv g_{\kappa\beta} R^\beta{}_{\lambda\mu\nu}$$

Then—owing entirely to (i) the definition of $R^\beta{}_{\lambda\mu\nu}$ and (ii) the $\mu\nu$-symmetry of $\Gamma^\beta{}_{\mu\nu}$—the tensor $R_{\kappa\lambda\mu\nu}$ possess the following *symmetry properties*:

$$\begin{aligned} R_{\kappa\lambda\mu\nu} &= -R_{\kappa\lambda\nu\mu} &&: \quad \text{antisymmetry on the last pair of indices}\\ &= -R_{\lambda\kappa\mu\nu} &&: \quad \text{antisymmetry on the first pair of indices}\\ &= +R_{\mu\nu\kappa\lambda} &&: \quad \text{supersymmetry}\\ R_{\kappa\lambda\mu\nu} + R_{\kappa\mu\nu\lambda} + R_{\kappa\nu\lambda\mu} &= 0 &&: \quad \text{windmill symmetry} \end{aligned}$$

These serve to reduce the number of independent components from N^4 to $\frac{1}{12}N^2(N^2-1)$:

N	N^4	$\frac{1}{12}N^2(N^2-1)$
1	1	0
2	16	1
3	81	6
4	256	20
5	625	50
6	1296	105
$\vdots$	$\vdots$	$\vdots$

We will, in particular, need to know that in the 2-dimensional case the only non-vanishing components of $R_{\kappa\lambda\mu\nu}$ are

$$R_{0101} = -R_{0110} = -R_{1001} = +R_{1010}$$

Returning now to the analytical mainstream...

The left side of (225) vanishes automatically, and from the invertibility of $\mathbb{W}$ we infer that

$$R_{\kappa\lambda\mu\nu} = 0 \tag{226}$$

Introducing (222) into (225) we find (after some calculation marked by a great deal of cancellation) that $R_{\kappa\lambda\mu\nu}$ has the correspondingly specialized structure

$$R_{\kappa\lambda\mu\nu} = g_{\kappa\nu}\Phi_{\lambda\mu} - g_{\kappa\mu}\Phi_{\lambda\nu} - g_{\lambda\nu}\Phi_{\kappa\mu} + g_{\lambda\mu}\Phi_{\kappa\nu} \tag{227}$$

where

$$\begin{aligned} \Phi_{\lambda\mu} &\equiv \varphi_{\lambda\mu} - \varphi_\lambda\varphi_\mu + \tfrac{1}{2}g_{\lambda\mu}\cdot(g^{\alpha\beta}\varphi_\alpha\varphi_\beta) \\ \varphi_{\lambda\mu} &\equiv \partial\varphi_\lambda/\partial x^\mu = \partial^2\varphi/\partial x^\lambda\partial x^\mu = \varphi_{\mu\lambda} \end{aligned} \tag{228}$$

entail $\Phi_{\lambda\mu} = \Phi_{\mu\lambda}$. It follows now from (227) that

$$R_{\lambda\mu} \equiv R^{\alpha}{}_{\lambda\mu\alpha} = \ (N-2)\,\Phi_{\lambda\mu} + g_{\lambda\mu}\cdot(g^{\alpha\beta}\Phi_{\alpha\beta}) \tag{229.1}$$
$$R \equiv R^{\beta}{}_{\beta} = 2(N-1)\cdot g^{\alpha\beta}\Phi_{\alpha\beta} \tag{229.2}$$

must—in consequence of (226)—both vanish:

$$R_{\lambda\mu} = 0 \tag{230.1}$$
$$R = 0 \tag{230.2}$$

In the $\boxed{\text{CASE } N=2}$ the equations (230) are seen to reduce to a solitary condition

$$g^{\alpha\beta}\Phi_{\alpha\beta} = 0 \tag{231}$$

which in $\boxed{\text{CASES } N>2}$ becomes a corollary of the stronger condition

$$\Phi_{\alpha\beta} = 0 \tag{232}$$

This is the **conformality condition** from which we will work. When introduced into (227) it renders (226) automatic.[132]

Note that (220) can be written $\partial_\lambda\varphi \equiv \varphi_\lambda = \partial_\lambda \log\sqrt{\Omega}$ and entails

$$\varphi = \log\sqrt{\Omega} + \text{constant}$$

Returning with this information to (228), the conformality condition (232) becomes

$$\frac{\partial^2 \log\sqrt{\Omega}}{\partial x^\mu \partial x^\nu} - \frac{\partial \log\sqrt{\Omega}}{\partial x^\mu}\frac{\partial \log\sqrt{\Omega}}{\partial x^\nu} + \tfrac{1}{2}g_{\mu\nu}\cdot g^{\alpha\beta}\frac{\partial \log\sqrt{\Omega}}{\partial x^\alpha}\frac{\partial \log\sqrt{\Omega}}{\partial x^\beta} = 0$$

which—if we introduce

$$F \equiv \frac{1}{\sqrt{\Omega}} \tag{233}$$

—can be written

$$\frac{\partial^2 \log F}{\partial x^\mu \partial x^\nu} = \tfrac{1}{2}g_{\mu\nu}\cdot g^{\alpha\beta}\frac{\partial \log F}{\partial x^\alpha}\frac{\partial \log F}{\partial x^\beta} - \frac{\partial \log F}{\partial x^\mu}\frac{\partial \log F}{\partial x^\nu}$$

[132] When $N=2$ one must, on the other hand, proceed from (231). It is therefore of interest that (231) and (226) are—uniquely in the $\boxed{\text{CASE } N=2}$ —notational variants of the *same* statement . . . for

$$\begin{aligned} R_{0101} &= \text{only independent element} \\ &= g_{01}\Phi_{10} - g_{00}\Phi_{11} - g_{11}\Phi_{00} + g_{10}\Phi_{01} \quad \text{by (227)} \\ &= -g\cdot g^{\alpha\beta}\Phi_{\alpha\beta} \quad \text{by} \quad \begin{pmatrix} g^{00} & g^{01} \\ g^{10} & g^{11}\end{pmatrix} = g^{-1}\cdot\begin{pmatrix} g_{11} & -g_{01} \\ -g_{10} & g_{00}\end{pmatrix}\end{aligned}$$

We write out the derivatives and obtain these simpler-looking statements

$$F_{\mu\nu} = g_{\mu\nu} \cdot \frac{g^{\alpha\beta}F_\alpha F_\beta}{2F} \tag{234}$$

where $F_\mu \equiv \partial_\mu F$ and $F_{\mu\nu} \equiv \partial_\mu \partial_\nu F$. The implication is that

$$\begin{aligned} \partial_\nu(g^{\lambda\mu}F_\mu) &= g^{\lambda\mu}F_{\mu\nu} \\ &= \Big\{\tfrac{1}{2}\frac{g^{\alpha\beta}F_\alpha F_\beta}{F}\Big\}\delta^\lambda{}_\nu \quad : \quad \text{vanishes unless } \nu = \lambda \end{aligned}$$

which is to say: $g^{\lambda\mu}F_\mu$ is a function only of x^λ. But $\mathbb{g}$ is, by initial assumption, a *constant diagonal* matrix, so we have

$$F_\mu \text{ is a function only of } x^\mu\text{, and so are all of its derivatives } F_{\mu\nu}$$

Returning with this information to (233), we are brought to the conclusion that the expression $\{\text{etc.}\}$ is a function only of x^0, only of x^1, ...; that it is, in short, a *constant* (call it $2C$), and that (233) can be written

$$F_{\mu\nu} = 2Cg_{\mu\nu}$$

giving

$$\begin{aligned} F &= Cg_{\alpha\beta}x^\alpha x^\beta - 2b_\alpha x^\beta + A \\ &= C\cdot(x,x) - 2(b,x) + A \end{aligned} \tag{235}$$

where b_α and A are constants of integration. Returning with this information to (234) we obtain

$$\begin{aligned} 4CF = g^{\alpha\beta}F_\alpha F_\beta &= g^{\alpha\beta}(2Cx_\alpha - 2b_\alpha)(2Cx_\beta - 2b_\beta) \\ &= 4C\Big[C\cdot(x,x) - 2(b,x) + \frac{(b,b)}{C}\Big] \end{aligned}$$

the effect of which, upon comparison with (235), is to constrain the constants $\{A, b_\alpha, C\}$ to satisfy

$$AC = (b,b)$$

This we accomplish by setting $C = (b,b)/A$, giving

$$F = A - 2(b,x) + \frac{(b,b)(x,x)}{A} \quad : \quad A \text{ and } b_\alpha \text{ now } \textit{un}\text{constrained}$$

Finally we introduce $a_\alpha \equiv b_\alpha/A$ to obtain the pretty result

$$F = A\big[1 - 2(a,x) + (a,a)(x,x)\big] \tag{236}$$

The conformal transformations $\mathfrak{X} \leftarrow \mathfrak{X}$ have yet to be described, but we now know this about W, the *Jacobian* of such a transformation:

$$\Omega = W^{\frac{2}{N}} = \frac{1}{F^2} = \frac{1}{A^2[1-2(b,x)+(a,a)(x,x)]^2} \tag{237}$$

Clearly, *tensor weight distinctions do not become moot in the context provided by the conformal group*, as they did (to within signs) in connection with the Lorentz group.

To get a handle on the functions $x^\alpha(x)$ that describe specific conformal transformations $\mathfrak{X} \leftarrow \mathfrak{X}$ we introduce

$$\partial_\mu \varphi \equiv \varphi_\mu = \partial_\mu \log \sqrt{\Omega} = -\partial_\mu \log F = -\frac{1}{F} F_\mu$$

into (223) to obtain

$$FW^\alpha{}_{\mu\nu} + F_\mu W^\alpha{}_\nu + F_\nu W^\alpha{}_\mu = g_{\mu\nu} \cdot g^{\lambda\kappa} F_\lambda W^\alpha{}_\kappa$$

or again (use $W^\alpha{}_\mu = \partial x^\alpha / \partial x^\mu$)

$$(Fx^\alpha)_{\mu\nu} = F_{\mu\nu} x^\alpha + g_{\mu\nu} \cdot g^{\lambda\kappa} F_\lambda W^\alpha{}_\kappa \tag{238}$$

To eliminate some subsequent clutter we agree translate from x-coordinates to y-coordinates whose origin coincides with that of the x-coordinate system: we write

$$x^\alpha(x) = y^\alpha(x) + Kt^\alpha \quad \text{with} \quad K \equiv A^{-1}$$

and achieve $y^\alpha(0) = 0$ by setting $Kt^\alpha \equiv x^\alpha(0)$. Clearly, if the functions $x^\alpha(x)$ satisfy (238) then so also do the functions $y^\alpha(x)$, and conversely. We change dependent variables now once *again*, writing

$$Fy^\alpha \equiv z^\alpha$$

Then $y^\alpha{}_\mu = -\frac{1}{F^2} F_\mu x^\alpha + \frac{1}{F} z^\alpha{}_\mu$ and (238) assumes the form

$$z^\alpha{}_{\mu\nu} = \frac{1}{F}\Big\{ \underbrace{\Big(F_{\mu\nu} - g_{\mu\nu} \cdot \frac{g^{\lambda\kappa} F_\lambda F_\kappa}{F}\Big)} z^\alpha + g_{\mu\nu} \cdot g^{\lambda\kappa} F_\lambda z^\alpha{}_\kappa \Big\}$$

It follows, however, from the previously established structure of F that

$$= -F_{\mu\nu} = -2Cg_{\mu\nu}$$

so

$$= g_{\mu\nu} \cdot \frac{1}{F}\Big\{ -2Cz^\alpha + g^{\lambda\kappa} F_\lambda z^\alpha{}_\kappa \Big\} \tag{239}$$

Each of these α-parameterized equations is structurally analogous to (234), and the argument that gave (235) no gives

$$z^\alpha(x) = P^\alpha \cdot (x,x) + \Lambda^\alpha{}_\beta x^\beta + \begin{bmatrix} \text{now no } x\text{-independent term} \\ \text{because } y(0){=}0 \;\Rightarrow\; z(0){=}0 \end{bmatrix}$$

Returning with this population of results to (239) we obtain

$$
\begin{aligned}
2P^\alpha\big[C(x,x)-2(b,x)+A\big] = -2C\big[P^\alpha(x,x)+\Lambda^\alpha{}_\beta x^\beta\big]\\
+\big[2Cx^\beta-2b^\beta\big]\big[2P^\alpha x_\beta+\Lambda^\alpha{}_\beta\big]
\end{aligned}
$$

—the effect of which (after much cancellation) is to constrain the constants P^α and $\Lambda^\alpha{}_\beta$ to satisfy $P^\alpha = -\frac{1}{A}\Lambda^\alpha{}_\beta b^\beta = -\Lambda^\alpha{}_\beta a^\beta$. Therefore

$$
z^\alpha(x) = \Lambda^\alpha{}_\beta\big\{x^\beta-(x,x)a^\beta\big\}
$$

Reverting to y-variables this becomes

$$
y^\alpha(x) = K\frac{\Lambda^\alpha{}_\beta\big\{x^\beta-(x,x)a^\beta\big\}}{1-2(a,x)+(a,a)(x,x)}
$$

so in x-variables—the variables of primary interest—we have

$$
x^\alpha(x) = K\left[t^\alpha+\frac{\Lambda^\alpha{}_\beta\big\{x^\beta-(x,x)a^\beta\big\}}{1-2(a,x)+(a,a)(x,x)}\right] \tag{240}
$$

Finally we set $K=1$ and $a^\alpha=0$ (all α) which by (237) serve to establish $\Omega = 1$. But in that circumstance (240) assumes the simple form

$$
\begin{aligned}
&\downarrow\\
&= \Lambda^\alpha{}_\beta x^\beta
\end{aligned}
$$

and the equation (185.2) that served as our point of departure becomes $\mathbb{\Lambda}^{\mathsf{T}}\mathbb{g}\,\mathbb{\Lambda} = \mathbb{g}$, from which we learn that the $\Lambda^\alpha{}_\beta$ must be elements of a Lorentz matrix.

Transformations of the form (240) have been of interest to mathematicians since the latter part of the 19^{th} Century. Details relating to the derivation of (240) by iteration of <u>*infinitesimal*</u> *conformal transformations* were worked out by S. Lie, and are outlined on pages 28–32 of J. E. Campbell's *Theory of Continuous Groups* (1903). The finitistic argument given above—though in a technical sense "elementary"—shows the toolmarks of a master's hand, and is in fact due (in essential outline) to H. Weyl (1923). I have borrowed most directly from V. Fock, *The Theory of Space, Time & Gravitation* (1959), Appendix A: "On the derivation of the Lorentz transformations."

Equation (240) describes—for $N \neq 2$—the most general <u>N-dimensional conformal transformation</u>, and can evidently be considered to arise by composition from the following:

Lorentz transformation : $x \to x = \mathbb{\Lambda}x$ (241.1)

Translation : $x \to x = x + t$ (241.2)

Dilation : $x \to x = Kx$ (241.3)

Möbius transformation : $x \to x = \dfrac{x-(x,x)a}{1-2(a,x)+(a,a)(x,x)}$ (241.4)

To specify such a transformation one must assign values to

$$\tfrac{1}{2}N(N-1)+N+1+N=\tfrac{1}{2}(N+2)(N+2)$$

adjustable parameters $\{t^\alpha, K, a^\alpha \text{ and the elements of } \log \mathbb{\Lambda}\}$, the physical dimensionalities of which are diverse but obvious. The associated numerology is summarized below:

N	$\tfrac{1}{2}(N+2)(N+1)$
1	3
2	$6+\infty$
3	10
4	15
5	21
6	28
$\vdots$	$\vdots$

Concerning the entry at $N=2$: equation (240) makes perfect sense in the $\boxed{\text{CASE } N=2}$, and that case provides a diagramatically convenient context within which to study the meaning of (240) in the general case. But (240) was derived from (232), which was seen on page 167 to be stronger that the condition (231) appropriate to the 2-dimensional case. The weakened condition requires alternative analysis,[133] and admits of more possibilities—actually *infinitely many more*, corresponding roughly to the infinitely many ways of selecting $f(z)$ in the theory of conformal transformations as it is encountered in complex function theory.[134] I do not pursue the topic because the physics of interest to us is inscribed (as are we) on 4-dimensional spacetime.

Some of the mystery which surrounds the Möbius transformations—which are remarkable for their *nonlinearity*—is removed by the remark that they can be assembled from translations and "inversions," where the latter are defined as follows:

$$\textbf{Inversion} \quad : \quad x \to x = \mu^2 \frac{x}{(x,x)} \tag{241.5}$$

Here μ^2 is a constant of arbitrary value, introduced mainly for dimensional reasons. The proof is by construction:

$$\left.\begin{aligned} x &\xrightarrow[\text{inversion}]{\hspace{3cm}} x = \mu^2 x/(x,x)\\ &\xrightarrow[\text{translation with } t=-\mu^2 a]{\hspace{3cm}} x = x-\mu^2 a\\ &\xrightarrow[\text{inversion}]{\hspace{3cm}} x = \mu^2 x/(x,x)\\ &\qquad\qquad\qquad\quad = \frac{x-(x,x)a}{1-2(a,x)+(a,a)(x,x)} \end{aligned}\right\} \tag{242}$$

---

[133] The problem is discussed in my TRANSFORMATIONAL PHYSICS OF WAVES (1979–1981).
[134] See again page 129.

Inversion—which

- admits readily of geometrical interpretation (as a kind of "radial reflection" in the isometric surface $(x,x)=\mu^2$)
- can be looked upon as the ultimate *source of the nonlinearity* which is perhaps the most striking feature of the conformal transformations (240)

—is one of the sharpest tools available to the conformal theorist, so I digress to examine some of its properties:

We have, in effect, already shown (at (242): set $a=0$) that *inversion is—like every kind of "reflection"—self-reciprocal*:

$$(\text{inversion})\cdot(\text{inversion}) = \text{identity} \tag{243}$$

That inversion is conformal in the sense "angle-preserving" can be established as follows: let x and y be the inversive images of x and y. Then

$$(x,y)=\mu^4\frac{(x,y)}{(x,x)(y,y)}$$

shows that inversion does not preserve inner products. But immediately

$$\frac{(x,y)}{\sqrt{(x,x)(y,y)}}=\frac{(x,y)}{\sqrt{(x,x)(y,y)}} \tag{244}$$

which is to say:

$$\text{angle} = \text{angle}$$

Inversion, since conformal, must be describable in terms of the primitive transformations listed at (241). How is that to be accomplished? We notice that each of those transformations—with the sole exception of the improper Lorentz transformations—is *continuous with the identity* (which arises at $\Lambda = \mathbb{I}$, at $t=0$, at $K=1$, at $a=0$). Evidently improper Lorentz transformations—in a word: reflections—must enter critically into the fabrication of inversion, and it is this observation that motivates the following short digression: For arbitrary non-null a^μ we can always write

$$x=\Big[x-\frac{(x,a)}{(a,a)}\,a\Big]+\frac{(x,a)}{(a,a)}\,a \equiv x_\parallel + x_\perp$$

which serves to resolve x^μ into components parallel/normal to a^μ. It becomes in this light natural to define

$$\textbf{a-reflection} \quad : \quad \begin{aligned} x &= x_\perp + x_\parallel \\ &\downarrow \\ \hat{x} &= x_\perp - x_\parallel = x - 2\frac{(x,a)}{(a,a)}\,a \end{aligned} \tag{245}$$

and to notice that (by quick calculation)

$$(\hat{x},\hat{y})=(x,y) \quad : \quad a\text{-reflection is inner-product preserving}$$

This simple fact leads us to notice that (245) can be written

$$\hat{x} = \mathbb{\Lambda}\, x \quad \text{with} \quad \mathbb{\Lambda} \equiv \|\Lambda^\mu{}_\nu\| = \|\delta^\mu{}_\nu - 2(a,a)^{-1}a^\mu a_\nu\|$$

where a brief calculation (examine $\Lambda^\alpha{}_\mu g_{\alpha\beta}\Lambda^\beta{}_\nu$) establishes that $\mathbb{\Lambda}$ is a Lorentz matrix with (according to *Mathematica*) $\det \mathbb{\Lambda} = -1$. In short:

$$\text{a-reflections are improper Lorentz transformations} \tag{246}$$

Thus prepared, we are led after a little exploratory tinkering to the following sequence of transformations:

$$\begin{aligned}
x \xrightarrow[\text{translation}]{\hspace{3cm}} x &= x - \frac{1}{(a,a)}\,a\\
\xrightarrow[\text{reflection}]{\hspace{3cm}} x &= x - 2\frac{(x,a)}{(a,a)}\,a\\
\xrightarrow[\text{Möbius}]{\hspace{3cm}} x &= \frac{x-(x,x)a}{1-2(a,x)+(a,a)(x,x)}\\
&\ \vdots\ \text{algebraic simplification}\\
&= \frac{1}{(a,a)}\left\{\frac{x}{(x,x)} - a\right\}\\
\xrightarrow[\text{reverse translation}]{\hspace{3cm}} x &= x + \frac{1}{(a,a)}a\\
&= \mu^2\frac{x}{(x,x)} \quad \text{with} \quad \mu^2 \equiv (a,a)^{-1}
\end{aligned}$$

The preceding equations make precise the sense in which

$$\text{inversion} = (\text{translation})^{-1}\cdot(\text{Möbius})\cdot(\text{reflection})\cdot(\text{translation}) \tag{247}$$

and confirm the conclusion reached already at (244): inversion is conformal. Finally, if one were to attempt direct evaluation of the Jacobian W of the general conformal transformation (240)—thus to confirm the upshot

$$W = \pm K^N\Big[\frac{1}{1-2(a,x)+(a,a)(x,x)}\Big]^N$$

of (237)—one would discover soon enough that one had a job on one's hands! But the result in question can be obtained as an easy consequence of the following readily-established statements:

$$
\begin{aligned}
W_{\text{inversion}} &= -\mu^{2N}\frac{1}{(x,x)^N} &&(248.1)\\
W_{\text{Lorentz}} &= \pm 1 &&(248.2)\\
W_{\text{translation}} &= \ \ 1 &&(248.3)\\
W_{\text{dilation}} &= \ \ K^N &&(248.4)
\end{aligned}
$$

It follows in particular from (242) that

$$
\begin{aligned}
W_{\text{Möbius}} &= (-)^2\mu^{2N}\frac{1}{(x,x)^N}\cdot 1\cdot \mu^{2N}\frac{1}{(x,x)^N} \quad \text{with} \quad x = \mu^2\Big[\frac{x}{(x,x)} - a\Big]\\
&= \Big[\frac{1}{1-2(a,x)+(a,a)(x,x)}\Big]^N &&(248.5)
\end{aligned}
$$

We are familiar with the fact that specialized Lorentz transformations serve to boost one to the frame of an observer O in uniform motion. I discuss now a related fact with curious electrodynamic implications: *specialized Möbius transformations serve to boost one to the frame of a uniformly accelerated observer*. From (241.4) we infer that a_μ has the dimensionality of reciprocal length, so

$$\tfrac{1}{2}g_\mu \equiv c^2 a_\mu \quad \text{is dimensionally an “acceleration”}$$

and in this notation (241.4) reads

$$x^\mu \longrightarrow x^\mu = \frac{x^\mu - \frac{1}{2c^2}(x,x)g^\mu}{1-\frac{1}{c^2}(g,x)+\frac{1}{4c^4}(g,g)(x,x)} \tag{249}$$

We concentrate now on implications of the assumption that g_μ possesses the specialized structure

$$\begin{pmatrix} g_0\\ g_1\\ g_2\\ g_3\end{pmatrix} = \begin{pmatrix} 0\\ \boldsymbol{g}\end{pmatrix}$$

that results from setting $g_0 = 0$. To describe (compare page 139) the “successive ticks of the clock at his origin” O writes

$$\begin{pmatrix} ct\\ \boldsymbol{0}\end{pmatrix}$$

which to describe those same events we write

$$\begin{pmatrix} ct\\ \boldsymbol{x}\end{pmatrix} = \frac{1}{1-(gt/2c)^2}\begin{pmatrix} ct\\ \boldsymbol{0}+\frac{1}{2}\boldsymbol{g}t^2\end{pmatrix}$$

where $g \equiv \sqrt{\boldsymbol{g}\cdot\boldsymbol{g}}$ and the + intruded because we are talking here about g^μ; *i.e.*, because we raised the index. In the *non-relativistic limit* this gives

$$\begin{pmatrix} t\\ \boldsymbol{x}\end{pmatrix} = \begin{pmatrix} t\\ \frac{1}{2}\boldsymbol{g}t^2\end{pmatrix} \tag{250}$$

which shows clearly the sense in which <u>we see O to be in a state of uniform acceleration</u>. To simplify more detailed analysis of the situation we (without loss of generality) sharpen our former assumption, writing

$$\boldsymbol{g} = \begin{pmatrix} g \\ 0 \\ 0 \end{pmatrix}$$

Then

$$1 - \tfrac{1}{c^2}(g, x) + \tfrac{1}{4c^4}(g,g)(x,x) = \frac{\big[(x-\lambda)^2 + y^2 + x^2 - c^2 t^2\big]}{\lambda^2}$$

$$\lambda \equiv \frac{2c^2}{g} \text{ is a "length"}$$

and (249) becomes

$$\left.\begin{aligned} t &= \frac{\lambda^2}{[\text{etc.}]}\cdot t \\ x &= \frac{\lambda^2}{[\text{etc.}]}\cdot\big\{x + \lambda^{-1}(c^2t^2 - x^2)\big\} \\ &= \frac{\lambda}{[\text{etc.}]}\cdot\Big\{c^2t^2 - \big(x - \tfrac{1}{2}\lambda\big)^2 + \big(\tfrac{1}{2}\lambda\big)^2\Big\} \\ &= \frac{\lambda}{[\text{etc.}]}\cdot\Big\{-\big[\text{etc.}\big] - \lambda(x-\lambda)\Big\} \\ y &= \frac{\lambda^2}{[\text{etc.}]}\cdot y \\ z &= \frac{\lambda^2}{[\text{etc.}]}\cdot z \end{aligned}\right\} \tag{251}$$

It is evident that $\big[\text{etc.}\big]$ vanishes—and the transformation (251) becomes therefore *singular*—on the lightcone $c^2t^2 - (x-\lambda)^2 - y^2 - x^2 = 0$ whose vertex is situated at $\big\{t, x, y, z\big\} = \big\{0, \lambda, 0, 0\big\}$. It is to gain a diagramatic advantage that we now set $y = z = 0$ and study what (251) has to say about how t and x depend upon t and x. We have

$$t = \frac{\lambda^2}{[(x-\lambda)^2 - c^2t^2]}\cdot t \tag{252.1}$$

$$(x+\lambda) = -\frac{\lambda^2}{[(x-\lambda)^2 - c^2t^2]}\cdot(x-\lambda) \tag{252.2}$$

which jointly entail

$$\big[c^2t^2 - (x+\lambda)^2\big]\big[c^2t^2 - (x-\lambda)^2\big] = \lambda^4 \tag{253}$$

But equations (252) can be written

$$\big[c^2t^2 - (x-\lambda)^2\big] = -\lambda^2\,\frac{t}{t} \tag{254.1}$$

$$= \lambda^2\,\frac{x-\lambda}{x+\lambda} \tag{254.2}$$

and when we return with the latter to (253) we find

$$\big[c^2t^2 - (x+\lambda)^2\big] = \lambda^2\,\frac{x+\lambda}{x-\lambda}$$

from which t has been eliminated: complete the square and obtain

$$\Big(x+\lambda\,\frac{x-\frac{1}{2}\lambda}{x-\lambda}\Big)^2 - (ct)^2 = \Big(\frac{\lambda^2}{2(x-\lambda)}\Big)^2 \tag{255.1}$$

which is seen to describe a x-parmeterized family of hyperbolas inscribed on the (t,x)-plane. These are *Möbius transforms of the* <u>*lines of constant* x</u> *inscribed on the* (t,x)*-plane*. Proceeding similarly to the elimination of $(x-\lambda)$ we find

$$\big[c^2t^2 - (x+\lambda)^2\big] = -\lambda^2\,\frac{t}{t}$$

giving

$$\Big(ct+\frac{\lambda^2}{2ct}\Big)^2 - (x+\lambda)^2 = \Big(\frac{\lambda^2}{2ct}\Big)^2 \tag{255.2}$$

which describes a t-parameterized family of hyperbolas—*Möbius transforms of the "time-slices" or* <u>*lines of constant* t</u> *inscribed on the* (t,x)*-plane*. The following remarks proceed from the results now in hand:

- O, by (252), assigns to O's origin the coordinates $t_0 = 0$, $x_0 = 0$; their origins, in short, coincide.
- In (255.1) set $x = 0$ and find that O writes

$$(x+\tfrac{1}{2}\lambda)^2 - (ct)^2 = (\tfrac{1}{2}\lambda)^2$$

  to describe O's worldline, which O sees to be hyperbolic, with x-intercepts at $x = 0$ and $x = -\lambda$ and asymptotes $ct = \pm(x+\frac{1}{2}\lambda)$ that intersect at $t = 0$, $x = -\frac{1}{2}\lambda$.
- If, in (252), we set $x = 0$ we obtain

$$t = \frac{\lambda^2}{[\lambda^2 - c^2t^2]}\cdot t$$
$$x = \frac{\lambda^3}{[\lambda^2 - c^2t^2]} - \lambda$$

  which provide O's t-parameterized description of O's worldline. Notice that <u>t and x both become infinite</u> at $t = \lambda/c$, and that t thereafter becomes <u>negative</u>!
- To describe her lightcone O writes $x = \pm ct$. Insert $x = +ct$ into (252.1), (ask *Mathematica* to) solve for t and obtain $ct = \lambda ct/(2ct+\lambda)$. Insert that result and $x = +ct$ into (252.2) and, after simplifications, obtain $x = +ct$. Repeat the procedure taking $x = -ct$ as your starting point: obtain $ct = -\lambda ct/(2ct-\lambda)$ and finally $x = -ct$. The striking implication is that (252) sends

$$O\text{'s lightcone} \quad\longmapsto\quad O\text{'s lightcone}$$

The conformal group is a wonderfully rich mathematical object, of which I have scarcely scratched the surface.[135] But I have scratched deeply enough to illustrate the point which motivated this long and intricate digression, a point made already on page 126:

The covariance group of a theory depends
in part upon *how the theory is expressed*:

One rendering of Maxwell's equations led us to the Lorentz group, and to special relativity. An almost imperceptibly different rendering committed us, however, to an entirely different line of analysis, and led us to an entirely different place—the conformal group, which contains the Lorentz group as a subgroup, but contains also much else . . . including transformations to the frames of "uniformly *accelerated* observers." Though it was electrodynamics which inspired our interest in the conformal group,[136] if you were to ask an elementary particle theorist about the conformal group you would be told that "the group arises as the covariance group of the wave equation

$$\Box\varphi = 0 \quad : \quad \text{conformally covariant}$$

Conformal covariance is broken (reduced to Lorentz covariance) by the inclusion of a "mass term"

$$(\Box + m^2)\varphi = 0 \quad : \quad \text{conformal covariance is broken}$$

It becomes the dominant symmetry in particle physics because at high energy mass terms can, in good approximation, be neglected

$$\text{rest energy } mc^2 \ll \text{total particle energy}$$

and enters into electrodynamics because the photon has no mass." That the group enters also into the physics of *massy* particles[133] is, in the light of such a remark, somewhat surprising. Surprises are imported also into classical electrodynamics by the occurrence of accelerations within the conformal group, for the question then arises: Does a uniformly accelerated charge radiate?[137]

---

[135] I scratch deeper, and discuss the occurance of the conformal group in connection with a rich variety of physical problems, in APPELL, GALILEAN & CONFORMAL TRANSFORMATIONS IN CLASSICAL/QUANTUM FREE PARTICLE DYNAMICS (1976) and TRANSFORMATIONAL PHYSICS OF WAVES (1979–1981).

[136] In "'Electrodynamics' in 2-dimensional spacetime" (1997) I develop a "toy electrodynamics" that gives full play to the exceptional richness that the conformal group has been seen to acquire in the 2-dimensional case.

[137] This question—first posed by Pauli in §32γ of his *Theory of Relativity*—once was the focus of spirited controversy: see T. Fulton & F. Rohrlich, "Classical radiation from a uniformly accelerated charge," Annals of Physics **9**,

8. Transformation properties of electromagnetic fields. To describe such a field at a spacetime point P we might display the values assumed there by the respective components of the electric and magnetic field vectors $\boldsymbol{E}$ and $\boldsymbol{B}$. Or we might display the values assumed there by the components $F^{\mu\nu}$ of the electromagnetic field tensor. To describe the *same physical facts* a second[138] observer O would display the values assumed by $\boldsymbol{E}$ and $\boldsymbol{B}$, or perhaps by $F^{\mu\nu}$. The *question* is

$$\text{How are } \{\boldsymbol{E}, \boldsymbol{B}\} \text{ and } \{\boldsymbol{E}, \boldsymbol{B}\} \text{ related?}$$

The *answer* has been in our possession ever since (at **A** on page 127, and on the "natural" grounds there stated) we assumed it to be the case that

$$F^{\mu\nu} \text{ transforms as a } \textbf{tensor density of unit weight} \tag{256}$$

But now we know things about the "allowed" coordinate transformations that on page 127 we did not know. Our task, therefore, is <u>to make explicit the detailed mathematical/physical consequences of (256)</u>. We know (see again (186) on page 129) that (256) pertains even when $\mathcal{X} \to \mathcal{X}$ is conformal, but I will restrict my attention to the (clearly less problematic, and apparently more important) case (184) in which

$$\mathcal{X} \to \mathcal{X} \text{ is } \textbf{Lorentzian}$$

The claim, therefore, is that

$$x \to x = \mathbb{\Lambda} x \quad \text{induces} \quad \mathbb{F} \to \mathbb{F} = V \cdot \mathbb{\Lambda} \mathbb{F} \mathbb{\Lambda}^{\mathsf{T}}$$

where $\mathbb{\Lambda}^{\mathsf{T}} g \mathbb{\Lambda} = g$ entails

$$V \equiv \frac{1}{\det \mathbb{\Lambda}} = \pm 1$$

and $\mathbb{F} = V \cdot \mathbb{\Lambda} \mathbb{F} \mathbb{\Lambda}^{\mathsf{T}}$ means $F^{\mu\nu} = V \Lambda^{\mu}{}_{\alpha} F^{\alpha\beta} \Lambda^{\nu}{}_{\beta}$. It is known, moreover, that (see again (211) on page 157) $\mathbb{\Lambda}$ can be considered to have this factored structure:

$$\mathbb{\Lambda} = \mathbb{R} \cdot \mathbb{\Lambda}(\boldsymbol{\beta})$$

(continued from the preceding page) 499 (1960); T. Fulton, F. Rohrlich & L. Witten, "Physical consequences of a coordinate transformation to a uniformly accelerated frame," Nuovo Cimento **26**, 652 (1962) and E. L. Hill, "On accelerated coordinate systems in classical and relativistic mechanics," Phys. Rev. **67**, 358 (1945); "On the kinematics of uniformly accelerated motions & classical electromagnetic theory," Phys. Rev. **72**, 143 (1947). The matter is reviewed by R. Peierls in §8.1 of *Surprises in Theoretical Physics* (1979), and was elegantly laid to rest by D. Boulware, "Radiation from a uniformly accelerated charge," Annals of Physics **124**, 169 (1980). For more general discussion see T. Fulton, F. Rohrlich & L. Witten, "Conformal invariance in physics," Rev. Mod. Phys. **34**, 442 (1962) and L. Page, "A new relativity," Phys. Rev. **49**, 254 (1936). Curiously, Boulware (with whom I was in touch earlier today: 30 October 2001) proceeded without explicit reference to the conformal group, of which he apparently was (and remains) ignorant.

[138] In view of the conformal covariance of electrodynamics I hesitate to insert here the adjective "inertial."

This means that we can study separately the response of $\mathbb{F}$ to spatial rotations $\mathbb{R}$ and its response to boosts $\mathbb{\Lambda}(\boldsymbol{\beta})$.

RESPONSE TO ROTATIONS Write out again (159)

$$\mathbb{F} = \mathbb{A}(\boldsymbol{E},\boldsymbol{B}) \equiv \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_1 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix} \equiv \begin{pmatrix} 0 & -\boldsymbol{E}^{\mathsf{T}} \\ \boldsymbol{E} & \mathbb{B} \end{pmatrix}$$

and (208)

$$\mathbb{R} \equiv \begin{pmatrix} 1 & \mathbf{0}^{\mathsf{T}} \\ \mathbf{0} & \mathbb{R} \end{pmatrix}$$

where

$$\mathbb{R} = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix}$$

is a 3×3 rotation matrix: $\mathbb{R}^{-1} = \mathbb{R}^{\mathsf{T}}$. It will, in a moment, become essential to notice that the latter equation, when spelled out in detail, reads

$$\begin{aligned} \frac{1}{\det\mathbb{R}} &\begin{pmatrix} (R_{22}R_{33} - R_{23}R_{32}) & (R_{13}R_{32} - R_{12}R_{33}) & (R_{12}R_{23} - R_{13}R_{22}) \\ (R_{23}R_{31} - R_{21}R_{33}) & (R_{11}R_{33} - R_{13}R_{31}) & (R_{21}R_{13} - R_{23}R_{11}) \\ (R_{32}R_{21} - R_{31}R_{22}) & (R_{31}R_{12} - R_{32}R_{11}) & (R_{11}R_{22} - R_{12}R_{21}) \end{pmatrix} \\ &= \begin{pmatrix} R_{11} & R_{21} & R_{31} \\ R_{12} & R_{22} & R_{32} \\ R_{13} & R_{23} & R_{33} \end{pmatrix} \end{aligned} \tag{257}$$

where

$$\frac{1}{\det\mathbb{R}} = \pm 1 \quad \text{according as } \mathbb{R} \text{ is proper/improper}$$

Our task now is the essentially elementary one of evaluating

$$\begin{aligned} \mathbb{F} &= \frac{1}{\det\mathbb{R}} \begin{pmatrix} 1 & \mathbf{0}^{\mathsf{T}} \\ \mathbf{0} & \mathbb{R} \end{pmatrix} \begin{pmatrix} 0 & -\boldsymbol{E}^{\mathsf{T}} \\ \boldsymbol{E} & \mathbb{B} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^{\mathsf{T}} \\ \mathbf{0} & \mathbb{R}^{\mathsf{T}} \end{pmatrix} \\ &= \frac{1}{\det\mathbb{R}} \begin{pmatrix} 0 & -(\mathbb{R}\boldsymbol{E})^{\mathsf{T}} \\ \mathbb{R}\boldsymbol{E} & \mathbb{R}\mathbb{B}\mathbb{R}^{\mathsf{T}} \end{pmatrix} \end{aligned}$$

which supplies

$$\boldsymbol{E} = (\det\mathbb{R})^{-1}\cdot\mathbb{R}\boldsymbol{E} \tag{258.1}$$

$$\mathbb{B} = (\det\mathbb{R})^{-1}\cdot\mathbb{R}\mathbb{B}\mathbb{R}^{\mathsf{T}} \tag{258.2}$$

The latter shows clearly how the antisymmetry of $\mathbb{B}$ comes to be inherited by $\mathbb{B}$, but does not much resemble its companion. HOWEVER . . . if we[139] first spell out

---

139 PROBLEM 47.

the meaning of (258.2)

$$\begin{pmatrix} 0 & -B_3 & B_2 \\ B_3 & 0 & -B_1 \\ -B_2 & B_1 & 0 \end{pmatrix} = (\det \mathbb{R})^{-1} \cdot \mathbb{R} \begin{pmatrix} 0 & -B_3 & B_2 \\ B_3 & 0 & -B_1 \\ -B_2 & B_1 & 0 \end{pmatrix} \mathbb{R}^{\mathsf{T}} \tag{259.1}$$

then (on a large sheet of paper) construct a detailed description of the matrix on the right, and finally make simplifications based on the rotational identity (257) ... we find that (258.1) is precisely equivalent to (which is to say: simply a notational variant of) the statement[140]

$$\Updownarrow$$

$$\begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} = \mathbb{R} \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} \tag{259.2}$$

Equations (258) can therefore be expressed

$$\boldsymbol{E} = (\det \mathbb{R})^{-1} \cdot \mathbb{R}\boldsymbol{E} \tag{260.1}$$

$$\boldsymbol{B} = \mathbb{R}\boldsymbol{B} \tag{260.2}$$

REMARK: In the conventional language of 3-dimensional physics, objects $\boldsymbol{A}$ that respond to rotation $\boldsymbol{x} \to \boldsymbol{x} = \mathbb{R}\boldsymbol{x}$ by the rule

$$\boldsymbol{A} \to \boldsymbol{A} = \mathbb{R}\boldsymbol{A}$$

are said to transform as **vectors** (or "polar vectors"), which objects that transform by the rule

$$\boldsymbol{A} \to \boldsymbol{A} = (\det \mathbb{R}) \cdot \mathbb{R}\boldsymbol{A}$$

are said to transform as **pseudovectors** (or "axial vectors"). Vectors and pseudovectors respond identically to proper rotations, but the latter respond to reflections (improper rotations) by acquisition of a minus sign. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are both vectors (or both pseudovectors) then $\boldsymbol{C} \equiv \boldsymbol{A}\times\boldsymbol{B}$ provides the standard example of a pseudovector ... for reasons that become evident when one considers what mirrors do to the "righthand rule."

The assumption[141] that

$F^{\mu\nu}$ transforms as a **tensor density of unit weight**

---

[140] For a more elegant approach to the proof of this important lemma see pages 22–22 in CLASSICAL GYRODYNAMICS (1976).

[141] See again the $\boxed{\text{FIRST POINT OF VIEW}}$, page 126.

was seen at (260) to carry the implication that

$$\left.\begin{array}{l}\boldsymbol{E} \text{ responds to rotation as a } \textbf{pseudovector}\\ \boldsymbol{B} \text{ responds to rotation as a } \textbf{vector}\end{array}\right\} \tag{261.1}$$

If we were, on the other hand, to assume[142] that

$$F^{\mu\nu} \text{ transforms as a } \textbf{weightless} \text{ tensor}$$

then the $(\det \mathbb{R})^{-1}$ factors would disappear from the right side of (258), and we would be led to the opposite conclusion:

$$\left.\begin{array}{l}\boldsymbol{E} \text{ responds to rotation as a } \textbf{vector}\\ \boldsymbol{B} \text{ responds to rotation as a } \textbf{pseudovector}\end{array}\right\} \tag{261.2}$$

The transformation properties of $\boldsymbol{E}$ and $\boldsymbol{B}$ are in either case "opposite,"[143] and it is from $\boldsymbol{E}$ that the transformation properties of ρ and $\boldsymbol{j}$ are inherited. The mirror image of the Coulombic field of a positive charge looks

- like the Coulombic field of a *negative* charge according to (261.1), but
- like the Coulombic field of a *positive* charge according to (261.2).

Perhaps it is for this reason (supported by no compelling physical argument) that (261.2) describes the tacitly-adopted *convention standard to the relativistic electrodynamical literature*. The factors that distinguish tensor densities from weightless tensors are, in special relativity, so nearly trivial ($\det \Lambda = \pm 1$) that many authors successfully contrive to neglect the distinction altogether.

RESPONSE TO BOOSTS All boosts are proper. Our task, therefore, is to evaluate

$$\mathbb{A}(\boldsymbol{E},\boldsymbol{B}) = \Lambda(\boldsymbol{\beta})\,\mathbb{A}(\boldsymbol{E},\boldsymbol{B})\,\Lambda^{\mathsf{T}}(\boldsymbol{\beta}) \tag{262}$$

where $\Lambda(\boldsymbol{\beta})$ has the structure (209) described on page 156. It will serve our exploratory purposes to suppose initially that

$$\boldsymbol{\beta} = \begin{pmatrix}\beta\\0\\0\end{pmatrix}$$

[142] See again the SECOND POINT OF VIEW, page 128.

[143] This fact has been latent ever since—at (67)—we alluded to the "$\boldsymbol{E}$-like character" of $\frac{1}{c}\boldsymbol{v}\times\boldsymbol{B}$, since

$$\text{vector} \times \begin{cases}\text{vector}\\ \text{pseudovector}\end{cases} = \begin{cases}\text{pseudovector}\\ \text{vector}\end{cases}$$

—*i.e.*, that we are boosting along the x-axis: then

$$\mathbb{A}(\boldsymbol{\beta}) = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and it follows from (262) by quick calculation that

$$\mathbb{A}(\boldsymbol{E},\boldsymbol{B}) = \begin{pmatrix} 0 & -E_1 & -\gamma(\boldsymbol{E}-\boldsymbol{\beta}\times\boldsymbol{B})_2 & -\gamma(\boldsymbol{E}-\boldsymbol{\beta}\times\boldsymbol{B})_3 \\ E_1 & 0 & -\gamma(\boldsymbol{B}+\boldsymbol{\beta}\times\boldsymbol{E})_3 & +\gamma(\boldsymbol{B}+\boldsymbol{\beta}\times\boldsymbol{E})_2 \\ \gamma(\boldsymbol{E}-\boldsymbol{\beta}\times\boldsymbol{B})_2 & +\gamma(\boldsymbol{B}+\boldsymbol{\beta}\times\boldsymbol{E})_3 & 0 & -B_1 \\ \gamma(\boldsymbol{E}-\boldsymbol{\beta}\times\boldsymbol{B})_3 & -\gamma(\boldsymbol{B}+\boldsymbol{\beta}\times\boldsymbol{E})_2 & +B_1 & 0 \end{pmatrix}$$

Noting that

$$\begin{aligned} E_1 &= (\boldsymbol{E}-\boldsymbol{\beta}\times\boldsymbol{B})_1 \quad \text{because} \quad (\boldsymbol{\beta}\times\boldsymbol{B}) \perp \boldsymbol{\beta} \\ B_1 &= (\boldsymbol{B}+\boldsymbol{\beta}\times\boldsymbol{E})_1 \quad \text{because} \quad (\boldsymbol{\beta}\times\boldsymbol{E}) \perp \boldsymbol{\beta} \end{aligned}$$

we infer that

$$\left.\begin{aligned} \boldsymbol{E} &= (\boldsymbol{E}-\boldsymbol{\beta}\times\boldsymbol{B})_\parallel + \gamma(\boldsymbol{E}-\boldsymbol{\beta}\times\boldsymbol{B})_\perp \\ \boldsymbol{B} &= (\boldsymbol{B}+\boldsymbol{\beta}\times\boldsymbol{E})_\parallel + \gamma(\boldsymbol{B}+\boldsymbol{\beta}\times\boldsymbol{E})_\perp \end{aligned}\right\} \tag{263}$$

where components $\parallel$ and $\perp$ to $\boldsymbol{\beta}$ are defined in the usual way: generically

$$\begin{aligned} \boldsymbol{A} &= \boldsymbol{A}_\parallel + \boldsymbol{A}_\perp \\ \boldsymbol{A}_\parallel &\equiv (\boldsymbol{A}\cdot\hat{\boldsymbol{\beta}})\hat{\boldsymbol{\beta}} = \underbrace{\frac{1}{\beta^2}\begin{pmatrix} \beta_1\beta_1 & \beta_1\beta_2 & \beta_1\beta_3 \\ \beta_2\beta_1 & \beta_2\beta_2 & \beta_2\beta_3 \\ \beta_3\beta_1 & \beta_3\beta_2 & \beta_3\beta_3 \end{pmatrix}}_{\text{projects onto }\boldsymbol{\beta}}\boldsymbol{A} \end{aligned}$$

Several comments are now in order:

1. We had already on page 46 (when we are arguing from *Galilean* relativity) reason to suspect that "$\boldsymbol{E}$ & $\boldsymbol{B}$ fields transform in a funny, interdependent way." Equations (263) first appear—somewhat disguised—in §4 of Lorentz (1904).[78] They appear also in §6 of Einstein (1905).[78] They were, in particular, unknown to Maxwell.

2. Equations (263) are ugly enough that they invite reformulation, and can in fact be formulated in a great variety of (superficially diverse) ways . . . some obvious—in the 6-vector formalism[86] one writes

$$\begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{B} \end{pmatrix} = \mathbb{M}(\boldsymbol{\beta}) \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{B} \end{pmatrix}$$

where $\mathbb{M}(\boldsymbol{\beta})$ is a 6×6 matrix whose elements can be read off from (263)—and some not so obvious. I would pursue this topic in response to some specific formal need, but none will arise.

3. The following statements are equivalent:

$$
\left.\begin{array}{l}
\text{Maxwell's equations}\\[6pt]
\qquad\begin{aligned}
\boldsymbol{\nabla}\!\cdot\!\boldsymbol{E} &= \rho\\
\boldsymbol{\nabla}\!\times\!\boldsymbol{B} - \tfrac{1}{c}\tfrac{\partial}{\partial t}\boldsymbol{E} &= \tfrac{1}{c}\boldsymbol{j}\\
\boldsymbol{\nabla}\!\cdot\!\boldsymbol{B} &= 0\\
\boldsymbol{\nabla}\!\times\!\boldsymbol{E} + \tfrac{1}{c}\tfrac{\partial}{\partial t}\boldsymbol{B} &= \boldsymbol{0}
\end{aligned}\\[6pt]
\text{simply “turn black” in response to}\\[6pt]
\qquad\begin{aligned}
t &= \gamma t + \tfrac{1}{c^2}\gamma\boldsymbol{v}\!\cdot\!\boldsymbol{x}\\
\boldsymbol{x} &= \boldsymbol{x} + \big\{\gamma t + (\gamma-1)\tfrac{1}{v^2}\boldsymbol{v}\!\cdot\!\boldsymbol{x}\big\}\boldsymbol{v}\\
\rho &= \gamma\rho + \tfrac{1}{c^2}\gamma\boldsymbol{v}\!\cdot\!\boldsymbol{j}\\
\boldsymbol{j} &= \boldsymbol{j} + \big\{\gamma\rho + (\gamma-1)\tfrac{1}{v^2}\boldsymbol{v}\!\cdot\!\boldsymbol{j}\big\}\boldsymbol{v}\\
\boldsymbol{E} &= (\boldsymbol{E}-\boldsymbol{\beta}\times\boldsymbol{B})_{\parallel} + \gamma(\boldsymbol{E}-\boldsymbol{\beta}\times\boldsymbol{B})_{\perp}\\
\boldsymbol{B} &= (\boldsymbol{B}+\boldsymbol{\beta}\times\boldsymbol{E})_{\parallel} + \gamma(\boldsymbol{B}+\boldsymbol{\beta}\times\boldsymbol{E})_{\perp}
\end{aligned}
\end{array}\right\}\tag{264.1}
$$

$$
\left.\begin{array}{l}
\text{Maxwell's equations}\\[6pt]
\qquad\begin{aligned}
\partial_\mu F^{\mu\nu} &= \tfrac{1}{c}j^\nu\\
\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} &= 0
\end{aligned}\\[6pt]
\text{simply “turn black” in response to}\\[6pt]
\qquad\begin{aligned}
x^\mu &= \Lambda^\mu{}_\alpha x^\alpha\\
j^\nu &= \Lambda^\nu{}_\beta j^\beta\\
F^{\mu\nu} &= \Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta F^{\alpha\beta}
\end{aligned}
\end{array}\right\}\tag{264.2}
$$

and provide detailed statements of what one means when one refers to the “Lorentz covariance of Maxwellian electrodynamics.” Note that it is not enough to know how Lorentz transformations act on spacetime coordinates: one must know also how they act on fields and sources. The contrast in the formal appearance of (264.1: Lorentz & Einstein) and (264.2: Minkowski) is striking, and motivates me to remark that

- it is traditional in textbooks to view (264.1) as “working equations,” and to regard (264.2) as “cleaned-up curiosities,” to be written down and admired as a kind of afterthought . . . but
- my own exposition has been designed to emphasize the *practical utility* of (264.2): I view (264.1) as “elaborated commentary” upon (264.2)—too complicated to work with except in some specialized applications.

4. We know now how to translate electrodynamical statements from one inertial frame to another. But we do *not* at present possess answers to questions such as the following:

- How do electromagnetic fields and/or Maxwell's equations look to an observer in a *rotating* frame?
- How—when Thomas precession is taken into account—does the nuclear Coulomb field look to an observer sitting on an electron in Bohr orbit?
- How do electromagnetic fields and the field equations look to an *arbitrarily accelerated* observer?

We are, however, *in position now to attack* such problems, should physical motivation arise.

5. Suppose O sees a pure $\boldsymbol{E}$-field: $\boldsymbol{B}(\boldsymbol{x}) = 0$ (all $\boldsymbol{x}$). It follows from (263) that *we* would see and electro*magnetic* field of the form

$$\begin{aligned}\boldsymbol{E} &= \boldsymbol{E}_{\parallel} + \gamma\boldsymbol{E}_{\perp} &&= \gamma\boldsymbol{E} + (1-\gamma)\tfrac{1}{v^2}(\boldsymbol{v}\cdot\boldsymbol{E})\,\boldsymbol{v} \\ \boldsymbol{B} &= \qquad\gamma(\boldsymbol{\beta}\times\boldsymbol{E}) &&= \tfrac{1}{c}\gamma(\boldsymbol{v}\times\boldsymbol{E})\end{aligned}$$

Our $\boldsymbol{B}$-field is, however, <u>structurally atypical</u>: it has a specialized ancestory, and (go to O's frame) *can be transformed away—globally*. In general it is not possible by Lorentz transformation to kill $\boldsymbol{B}$ (or $\boldsymbol{E}$) even locally, for to do so would be (unless $\boldsymbol{E}\perp\boldsymbol{B}$ at the spacetime point in question) to stand in violation of the second of the following remarkable equations[144]

$$\boldsymbol{E}\cdot\boldsymbol{E} - \boldsymbol{B}\cdot\boldsymbol{B} = \boldsymbol{E}\cdot\boldsymbol{E} - \boldsymbol{B}\cdot\boldsymbol{B} \tag{265.1}$$

$$\boldsymbol{E}\cdot\boldsymbol{B} = \boldsymbol{E}\cdot\boldsymbol{B} \tag{265.2}$$

The preceding remark makes vividly clear, by the way, why it is that attempts to "derive" electrodynamics from "Coulomb's law + special relativity" are doomed to fail: with only that material to work with one cannot escape from the force of the special/atypical condition $\boldsymbol{E}\cdot\boldsymbol{B} = 0$.

6. We do *not* have in hand the statements analogous to (264) that serve to lend detailed meaning to the "<u>conformal</u> covariance of Maxwellian electrodynamics." To gain a sense of the most characteristic features of the enriched theory it would be sufficient to describe how electromagnetic fields and sources respond to dilations and inversions.

7. An uncharged copper rod is transported with velocity $\boldsymbol{v}$ in the presence of a homogeneous magnetic field $\boldsymbol{B}$. We see a *charge separation* to take place (one end of the rod becomes positively charge, the other negatively: see Figure 66), which we attribute the presence $q(\boldsymbol{v}\times\boldsymbol{B})$-forces. But an observer O co-moving with the rod sees no such forces (since $\boldsymbol{v} = \boldsymbol{0}$), and must attribute the charge separation phenomenon to the presence of an *electric* field $\boldsymbol{E}$. <u>It was to account for such seeming "explanatory asymmetry" that Einstein *invented* the theory of relativity</u>. I quote from the beginning of his 1905 paper:

---

[144] PROBLEM 48.

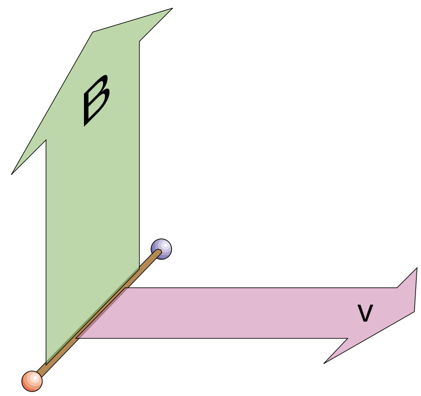

Figure 66: *A copper rod is transported with constant velocity* $\boldsymbol{v}$ *in a homogeneous magnetic field. Charge separation is observed to occur in the rod. Observers in relative motion explain the phenomenon in—unaccountably, prior to the invention of special relativity—quite different ways.*

ON THE ELECTRODYNAMICS OF MOVING BODIES

A. EINSTEIN

It is known that Maxwell's electrodynamics—as usually understood at the present time—when applied to moving bodies, leads to asymmetries which do not appear to be inherent in the phenomena. Take, for example, the reciprocal electrodynamic action of a magnet and a conductor. The observable phenomenon here depends only on the relative motion of the conductor and the magnet, whereas the customary view draws a sharp distinction between the two cases in which either the one or the other of these bodies is in motion. For if the magnet is in motion and the conductor at rest, there arises in the neighborhood of the magnet an electric field with a certain definite energy, producing a current at the places where parts of the conductor are situated. But if the magnet is stationary and the conductor in motion, no electric field arises in the neighborhood of the magnet. In the conductor, however, we find an electrtomotive force, to which in itself there is no corresponding energy, but which gives rise—assuming equality of relative motion in the two cases discussed—to elecric currents of the same path and intensity as those produced by the electric forces in the former case.

Examples of this sort, together with the unsuccessful attempts to discover any motion of the earth relatively to the "light medium," suggest that the phenomena of electrodynamics as well as of mechanics possess no properties corresponding to the idea of absolute rest.

After sixteen pages of inspired argument Einstein arrives at equations (263), from which he concludes that

> . . . electric and magnetic forces do not exist independently of the state of motion of the system of coordinates.
>
> Furthermore it is clear that the asymmetry mentioned in the introduction as arising when we consider the currents produced by the relative motion of a magnet and a conductor now disappears.

He comes to the latter conclusion by arguing that to determine the force $\boldsymbol{F}$ experienced by a moving charge q in an electromagnetic field $\{\boldsymbol{E},\boldsymbol{B}\}$ a typical inertial observer should

i) transform $\{\boldsymbol{E},\boldsymbol{B}\} \rightarrow \{\boldsymbol{E}_0,\boldsymbol{B}_0\}$ to the instantaneous rest frame of the charge;

ii) write $\boldsymbol{F}_0 = q\boldsymbol{E}_0$;

iii) transform back again to his own reference frame: $\boldsymbol{F} \leftarrow \boldsymbol{F}_0$.

We don't, as yet, know how to carry out the last step (because we have yet to study relativistic mechanics). It is already clear, however, that Einstein's program *eliminates asymmetry because it issues identical instructions to every inertial observer*. Note, moreover, that it contains no reference to "the" velocity . . . but refers only to the *relative* velocity (of charge and observer, of observer and observer).

The field-transformation equations (263) lie, therefore, at the motivating heart of Einstein's 1905 paper. All the rest can be read as "technical support"—evidence of the extraordinary surgery Einstein was willing to perform to remove a merely *aesthetic blemish* from a theory (Maxwellean electrodynamics) which—after all—worked perfectly well as it was! Several morals could be drawn. Most are too obvious to state . . . and all are too important for the creative physicist to ignore.

9. Principle of relativity. The arguments which led Einstein to the Lorentz transformations differ profoundly from those which (unbeknownst to Einstein) had led Lorentz to the same result. Lorentz argued (as we have seen . . . and done) from the structure of Maxwell's equations. Einstein, on the other hand (and though he had an electrodynamic problem in mind), extracted the Lorentz transformations from an unprecedented *operational analysis*: his argument assumed very little . . . and he had, therefore, correspondingly greater confidence in the inevitability and generality of his conclusions. His argument was, in particular, entirely free from any reference to Maxwell's equations, so his conclusion—that inertial observers are interrelated by Lorentz transformations —*could not be specific to Maxwellean electrodynamics*. It was this insight—and the firmness[145] with which he adhered to it—which distinguished Einstein's thought from that of his contemporaries (Lorentz, Poincaré). It led him to

[145] I have indicated on page 163 why, in the light of subsequent developments, Einstein's "firmness" can be argued to have been *inappropriately* strong.

propose, at the beginning of his §2, two principles . . . which amount, in effect, to this, the

> **Principle of Relativity:** The concepts, statements and formulæ of physics—whatever the phenomenology to which they specifically pertain—must *preserve their structure* when subjected to Lorentz transformation. (266)

The principle of relativity functions as a "syntactical constraint" on the "statements that physicists may properly utter"—at least when they are doing *local* physics. Concepts/statements/theories which fail to pass the (quite stringent) "Lorentz covariance test" can, according to the principle of relativity, be dismissed out of hand as ill-formed, inconsistent with the grammar of physics . . . and therefore physically untenable. Theories that pass the test are said to be "relativistic," "Lorentz invariant" or (more properly) *Lorentz covariant*. The *physical correctness* of such a theory is, of course, not guaranteed. What is guaranteed is the ultimate physical incorrectness of any theory—whatever may be its utility in circumscribed contexts (think of non-relativistic classical and quantum mechanics!)—that stands in violation of the principle of relativity.[146]

Some theories—such as the version of Maxwellean electrodynamics that was summarized at (264.1)—conform to the principle of relativity, but do so "non-obviously." Other theories—see again (264.2)—conform more obviously. Theories of the latter type are said to be "*manifestly Lorentz covariant*." Manifest is, for obvious reasons, a very useful formal attribute for a physical theory to possess. Much attention has been given therefore to the cultivation of principles and analytical techniques which sharpen one's ability to generate manifestly covariant theories "automatically." Whence the importance which theoretical physicists nowadays attach to *variational principles, tensor analysis, group representation theory,* . . . (Einstein did without them all!).

Clearly, the principle of relativity involves much besides the simple "theory of Lorentz transformations" (it involves, in short, all of physics!) . . . but one must have a good command of the latter subject in order to *implement* the principle. If in (266) one substitutes for the word "Lorentz" the words "Galilean," "conformal," . . . one obtains the "**principle of Galilean relativity**," the "**principle of conformal relativity**," etc. These do have some physically illuminating formal consequences, but appear to pertain only *approximately* to the world-as-we-find-it . . . while the principle announced by Einstein pertains "exactly/universally."

I have several times emphasized the <u>universal applicability</u> of the principle

[146] But *every* physical theory is *ultimately* incorrect! So the question that confronts physicists in individual cases is this: Is Lorentz non-covariance the *principal* defect of the theory in question, the defect worth of my corrective attention? Much more often than not, the answer is clearly "No."

of relativity. It is, therefore, *by way of illustrative application* that in Part II of his paper Einstein turns to the specific physics which had served initially to motivate his research—Maxwellean electrodynamics. It is frequently stated that "electrodynamics was *already* relativistic (while Newtonian dynamics had to be deformed to conform)." But this is not quite correct. The electrodynamics inherited by Einstein contained field equations, but it *contained no allusion to a field transformation law*. Einstein *produced* such a law—namely (263)—by insisting *that Maxwell's field equations conform to the principle of relativity*. Einstein *derived* (from Maxwell's equations + relativity, including prior knowledge of the Lorentz transformations) a result—effectively: that the $F^{\mu\nu}$ transform tensorially—which we were content (on page 127) to *assume*. We, on the other hand, used Maxwell's equations + tensoriality to *deduce the design of the Lorentz transformations*. Our approach—which is effectively Lorentz'—is efficient (also free of allusions to trains & lanterns), but might be criticized on the ground that it is excessively "parochial," too much rooted in specifics of electrodynamics. It is not at all clear that our approach would have inspired anyone to utter a generalization so audacious as Einstein's (266). Historically it didn't: both Lorentz and Poincaré were in possession of the technical rudiments of relativity already in 1904, yet both—for distinct reasons—failed to recognize the revolutionary force of the idea encapsulated at . Einstein was, in this respect, well served by his trains and lanterns. But it was not Einstein but Minkowski who first appreciated that at Einstein had in effect prescribed that

> The physics inscribed on spacetime must mimic
> the symmetry structure of spacetime itself.

10. Relativistic mechanics of a particle. We possess a Lorentz covariant field dynamics. We want a theory of *fields and (charged) particles in interaction*. Self-consistency alone requires that the associated particle dynamics be Lorentz covariant. So also—irrespective of any reference to electromagnetism—does the principle of relativity.

The discussion which follows will illustrate how non-relativistic theories are "deformed to conform" to the principle of relativity. But it is offered to serve a more explicit and pressing need: my primary goal will be *to develop descriptions of the relativistic analogs of the pre-relativistic concepts of energy, momentum, force*, . . . though a number of collateral topics will be treated *en route*.

In Newtonian dynamics the "worldline" of a mass point m is considered to be described by the 3-vector-valued solution $\boldsymbol{x}(t)$ of a differential equation of the form

$$\boldsymbol{F}(t,\boldsymbol{x}) = m\tfrac{d^2}{dt^2}\boldsymbol{x}(t) \tag{267}$$

This equation conforms to the **principle of Galilean covariance** (and it was from this circumstance that historically we acquired our interest in the "population

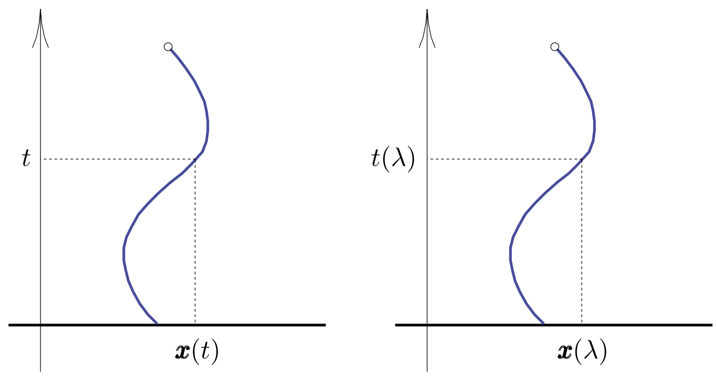

FIGURE 67: *At left: the time-parameterized flight of a particle, standard to Newtonian mechanics, where t is assigned the status of an independent variable and $\boldsymbol{x}$ is a set of dependent variables. At right: arbitrarily parameterization permits t to join the list of dependent variables; i.e., to be treated co-equally with $\boldsymbol{x}$.*

of inertial observers"), but *its Lorentz non-covariance is manifest* . . . for the equations treats t and $\boldsymbol{x}$ with a distinctness which the Lorentz transformations do not allow because they do not preserve. We confront therefore this problem: How to describe a worldline in conformity with the requirement that space and time coordinates be treated *co-equally*? One's first impulse it to give up t-parameterization in favor of an *arbitary* parameterization of the worldline (Figure 67), writing $x^\mu(\lambda)$. This at least treats space and time co-equally . . . but leaves every inertial observer to his own devices: the resulting theory (kinematics) would be too sloppy to support sharp physics. The "slop" would, however, disappear if λ could be assigned a "natural" meaning—a meaning which *stands in the same relationship to all inertial observers*. Einstein's idea—foreshadowed already on page 186—was to assign to λ the meaning/value of "time as measured by a comoving clock." The idea is implemented as follows (see Figure 68): Let O write $x(\lambda)$ to describe a worldline, and let him write

$$dx(\lambda) \equiv x(\lambda + d\lambda) - x(\lambda) = \begin{pmatrix} cdt \\ \boldsymbol{dx} \end{pmatrix}$$

to describe the interval separating a pair of "neighboring points" (points on the tangent at $x(\lambda)$). If and only if $dx(\lambda)$ is timelike will O be able to *boost to the instantaneous restframe* (*i.e.*, to the frame of an observer O who sees the particle to be *momentarily resting at her origin*):

$$\begin{pmatrix} cdt \\ \boldsymbol{dx} \end{pmatrix} = \mathbb{\Lambda}(\boldsymbol{\beta}) \begin{pmatrix} cd\tau \\ \boldsymbol{0} \end{pmatrix}$$

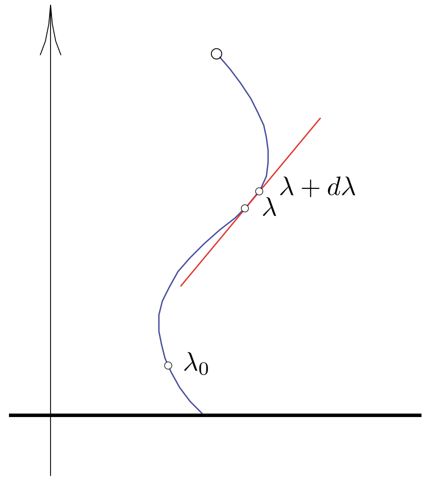

FIGURE 68: *An accelerated observer/particle borrows his/its* proper time *increments* $d\tau$ *from the wristwatches of momentarily comoving inertial observers*.

where from the boost-invariant structure of spacetime it follows that

$$\begin{aligned} d\tau &= \sqrt{(dt)^2 - \tfrac{1}{c^2}\boldsymbol{dx}\cdot\boldsymbol{dx}} = \sqrt{1-\beta^2(t)}\,dt \qquad (268)\\ &\equiv \text{time differential measured by instantaneously comoving clock}\\ &= \tfrac{1}{c}\sqrt{\big(\tfrac{dx^0}{d\lambda}\big)^2 - \big(\tfrac{dx^1}{d\lambda}\big)^2 - \big(\tfrac{dx^2}{d\lambda}\big)^2 - \big(\tfrac{dx^3}{d\lambda}\big)^2}\,d\lambda\\ &= \tfrac{1}{c}ds \end{aligned}$$

The <u>*proper time* τ</u> associated with a *finitely*-separated pair of points is defined

$$\begin{aligned} \tau(\lambda,\lambda_0) &= \tfrac{1}{c}\int_{\lambda_0}^{\lambda}\sqrt{g_{\alpha\beta}\frac{dx^\alpha(\lambda')}{d\lambda'}\frac{dx^\beta(\lambda')}{d\lambda'}}\,d\lambda' = \int d\tau = \frac{\text{arc-length}}{c}\\ &= 0 \text{ at } \lambda=\lambda_0 : \begin{cases} x(\lambda_0) \text{ is the reference point at which}\\ \text{we “start the proper clock”}\end{cases} \end{aligned}$$

Functional inversion gives

$$\lambda = \lambda(\tau,\lambda_0)$$

and in place of $x(\lambda)$ it becomes natural to write

$$x(\tau) \equiv x(\lambda(\tau,\lambda_0)) \;\;:\;\; \tau\text{-parameterized description of the worldline}$$

Evidently τ-parameterization is equivalent (to within a c-factor) to *arc-length parameterization*—long known by differential geometers to be “most natural” in metric spaces. Two points deserve comment:

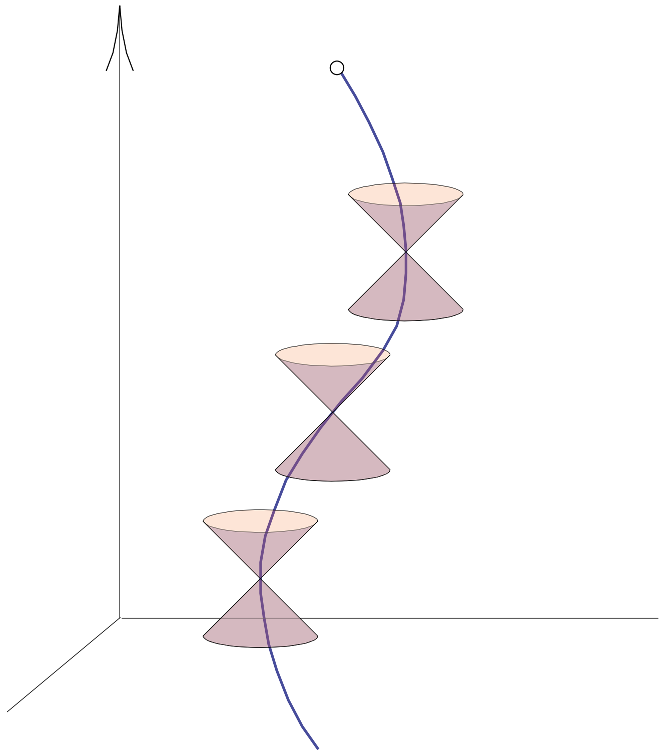

FIGURE 69: *The worldline of a masspoint lies everywhere interior to lightcones with vertices on the worldline. The spacetime interval separating any two points on a worldline is therefore time-like, and the constituent points of the worldline fall into a temporal sequence upon which all inertial observers agree.*

1. Einstein's program works if and only if all tangents to the worldline are *timelike* (Figure 69). One cannot, therefore, τ-parameterize the worldline of a photon. Or of a "tachyon." The reason is that one cannot boost such particles to rest: one cannot Lorentz transform the tangents to such worldlines into local coincidence with the x^0-axis.

2. The $d\tau$'s in $\int d\tau$ refer to a *population of osculating inertial observers*. It is a big step—a step which Einstein (and also L. H. Thomas) considered quite "natural," but a big step nonetheless—to suppose that τ has anything literally to do with "time as measured by a comoving (which in the general case means an accelerating) clock." The relativistic dynamics of particles is, in fact, *independent* of whether attaches literal meaning to the preceding phrase. Close reading of Einstein's paper shows, however, that he *did* intend to be understood literally (even though—patent clerk that he was—he would not have expected his mantle clock to keep good time if jerked about). Experimental evidence supportive of Einstein's view derives from the *decay of accelerated radioactive* *particles* and from recent observations pertaining to the so-called *twin paradox* (see below).

Given a τ-parameterized (whence everywhere timelike) worldline $x(\tau)$, we define by

$$u(\tau) \equiv \tfrac{d}{d\tau}x(\tau) = \tfrac{dt}{d\tau}\tfrac{d}{dt}\begin{pmatrix} ct \\ \boldsymbol{x} \end{pmatrix} = \gamma\begin{pmatrix} c \\ \boldsymbol{v} \end{pmatrix} \tag{269}$$

the *4-velocity* $u^\mu(\tau)$, and by

$$\begin{aligned} a(\tau) &\equiv \tfrac{d^2}{d\tau^2}x(\tau) \\ &= \tfrac{d}{d\tau}u(\tau) = \tfrac{dt}{d\tau}\tfrac{d}{dt}\gamma\begin{pmatrix} c \\ \boldsymbol{v} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{c}\gamma^4(\boldsymbol{a}\cdot\boldsymbol{v}) \\ \gamma^2\boldsymbol{a} + \frac{1}{c^2}\gamma^4(\boldsymbol{a}\cdot\boldsymbol{v})\boldsymbol{v} \end{pmatrix} \end{aligned} \tag{270}$$

the *4-acceleration* $a^\mu(\tau)$. These are equations written by inertial observer O: $\boldsymbol{v}$ refers to O's perception of the particle's instantaneous velocity $\boldsymbol{v}(t)$, and $\gamma \equiv \left[1 - \frac{1}{c^2}\boldsymbol{v}\cdot\boldsymbol{v}\right]^{-\frac{1}{2}}$.[147] Structurally similar equations (but with everything turned red) would be written by a second observer O. In developing this aspect of the subject one must be very careful to distinguish—both notationally and conceptually—the following:

O's perception of the instantaneous particle velocity $\boldsymbol{v}$
O's perception of O's velocity $\boldsymbol{s}$
O's perception of the instantaneous particle velocity $\boldsymbol{v}$

Supposing O and O to be boost-equivalent (no frame rotation)

$$x = \mathbb{A}(\boldsymbol{s}/c)\,x$$

we have

$$u = \mathbb{A}(\boldsymbol{s}/c)\,u \tag{271.1}$$

$$a = \mathbb{A}(\boldsymbol{s}/c)\,a \tag{271.2}$$

These equations look simple enough, but their explcit meaning is—owing to the complexity of $\mathbb{A}(\boldsymbol{s}/c)$, of u^μ and particularly of a^μ—actually quite complex. I will develop the detail only when forced by explicit need.[148]

It follows from (269) that

$$(u,u) = g_{\alpha\beta}u^\alpha u^\beta = \gamma^2(c^2 - v^2) = c^2\cdot\gamma^2(1-\beta^2) = c^2 \tag{272}$$

---

[147] PROBLEM 49.
[148] In the meantime, see my ELECTRODYNAMICS (1972/73), pages 202–205.

according to which all velocity 4-vectors have the same Lorentzian length. All are, in particular (since $(u,u)=c^2>0$), *timelike*. Differentiating (272) with respect to τ we obtain

$$\tfrac{d}{d\tau}(u,u)=2(u,a)=0 \tag{273}$$

according to which it is invariably the case that $u\perp a$ in the Lorentzian sense. It follows now from the timelike character of u that all acceleration 4-vectors are *spacelike*. Direct verification of these statements could be extracted from (269) and (270). The statement $(u,u)=c^2$—of which (273) is an immediate corollary—has no precursor in non-relativistic kinematics,[149] but is, as will emerge, absolutely fundamental to relativistic kinematics/dynamics.

Looking "with relativistic eyes" to Newton's 2$^{\text{nd}}$ law (267) we write

$$K^\mu = m\tfrac{d^2}{d\tau^2}x^\mu(\tau) \tag{274}$$

This equation would be Lorentz covariant—*manifestly* covariant—if

$$K^\mu \equiv \textit{Minkowski force} \text{ transforms like a 4-vector}$$

and m transforms as an invariant. The Minkowski equation (274) can be reformulated

$$K^\mu = m\tfrac{d}{d\tau}u^\mu = ma^\mu$$

or again
$$= \tfrac{d}{d\tau}p^\mu$$

where
$$p^\mu \equiv mu^\mu = \gamma m\begin{pmatrix} c\\ \boldsymbol{v}\end{pmatrix} \equiv \begin{pmatrix} p^0\\ \boldsymbol{p}\end{pmatrix} \tag{275}$$

From the γ-expansion (202) we obtain

$$\begin{aligned} p^0 &= \gamma mc \qquad\qquad (276.1)\\ &= \big(1+\tfrac{1}{2}\beta^2+\tfrac{3}{8}\beta^4+\cdots\big)mc\\ &= \tfrac{1}{c}\big(mc^2+\tfrac{1}{2}mv^2+\cdots\big)\end{aligned}$$

$\quad\uparrow$—familiar from non-relativistic dynamics as kinetic energy

$$\begin{aligned} \boldsymbol{p} &= \gamma m\boldsymbol{v} \qquad\qquad (276.2)\\ &= m\boldsymbol{v}+\cdots\end{aligned}$$

$\uparrow$—familiar from non-relativistic dynamics as linear momentum

It becomes in this light reasonable to call p^μ the *energy-momentum* 4-vector.

---

[149] The constant speed condition

$$\boldsymbol{v}\cdot\boldsymbol{v} = \text{constant}$$

is sometimes encountered, but has no claim to "universality" in non-relativistic physics: *when* encountered (as in uniform circular motion), it entails $\boldsymbol{v}\perp\boldsymbol{a}$.

Looking to the finer details of standard relativistic terminology . . . one writes

$$p^0 = \tfrac{1}{c}E \tag{277}$$

and calls

$$E = \gamma mc^2 = mc^2 + \tfrac{1}{2}mv^2 + \cdots$$

the *relativistic energy*. More particularly

$$\begin{aligned} E_0 &\equiv mc^2 && \text{is the } \textit{rest energy} \\ T &\equiv E - E_0 && \text{is the } \textit{relativistic kinetic energy} \end{aligned} \tag{278}$$

In terms of the $\boldsymbol{v}$-dependent "relativistic mass" defined[150]

$$M \equiv \gamma m = \frac{m}{\sqrt{1 - v^2/c^2}} \tag{279}$$

we have

$$E = Mc^2 \tag{280.1}$$

and

$$T = (M - m)c^2 = \left\{\frac{1}{\sqrt{1 - v^2/c^2}} - 1\right\}mc^2$$

The *relativistic momentum* can in this notation be described

$$\boldsymbol{p} = M\boldsymbol{v} \tag{280.2}$$

It is—so far as I can tell—the "non-relativistic familiarity" of (280.2) that tempts some people[151] to view (283) as the fruit of an astounding "empirical discovery," lying (they would have us believe) near the physical heart of special relativity. But (283) is, I insist, a *definition*—an occasional convenience, nothing more—one incidental detail among many in a coherent theory. It is naive to repeat the tired claim that "in relativity mass becomes velocity dependent:" it is profoundly wrongheaded to attempt to force relativistic dynamics to look less relativistic than it is.

We have

$$p = \begin{pmatrix} \frac{1}{c}E \\ \boldsymbol{p} \end{pmatrix} = mu$$

and from (272) it follows that

$$(p,p) = (E/c)^2 - \boldsymbol{p}\cdot\boldsymbol{p} - m^2c^2 \tag{281}$$

This means that p lies always on a certain m-determined hyperboloid (called the "*mass shell*": see Figure 70) in 4-dimensional energy-momentum space.

[150] It becomes natural in the following context to call m the *rest mass*, though in grown-up relativistic physics there is *really no other kind*. Those who write m when they mean M are obliged to write m_0 to distinguish the rest mass.

[151] See, for example, A. P. French, *Special Relativity: The MIT Introductory Physics Series* (1968), page 23.

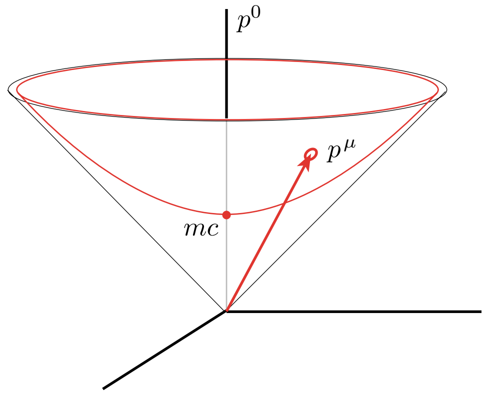

FIGURE 70: *The hyperboloidal mass shell, based upon (281) and drawn in energy-momentum space. The p^0-axis (energy axis) runs up. The mass shell intersects the p^0-axis at a point determined by the value of m:*

$$p^0 = mc \quad i.e., \quad E = mc^2$$

The figure remains meaningful (though the hyperboloid becomes a cone) even in the limit $m \downarrow 0$, which provides first indication that relativistic mechanics supports a theory of massless particles.

From (281) we obtain

$$\begin{aligned} E &= \pm c\sqrt{\boldsymbol{p}\cdot\boldsymbol{p} + (mc)^2} \\ &= \pm\Big\{mc^2 + \tfrac{1}{2m}\boldsymbol{p}\cdot\boldsymbol{p} + \cdots\Big\} \end{aligned} \tag{282}$$

which for a relativistic particle describes the $\boldsymbol{p}$-dependence of the energy E, and should be compared with its non-relativistic free-particle counterpart

$$E = \tfrac{1}{2m}\boldsymbol{p}\cdot\boldsymbol{p}$$

The $\pm$ assumes major importance in relativistic *quantum* mechanics (where it must be explained away lest it provide a rathole that would de-stabilize the world!), but in relativistic classical mechanics one simply *abandons* the minus sign—dismisses it as an algebraic artifact.

Looking next to the structure of K^μ ...ot follows from the Minkowski equation $K = ma$ by $(u, a) = 0$ that

$$(K, u) = 0 \quad : \quad K \perp u \text{ in the Lorentzian sense} \tag{283}$$

We infer that the 4-vectors that describe Minkowski forces are *invariably spacelike*. It follows moreover from (283) that as $p \sim u$ moves around the K-vector must move in concert, contriving always to be $\perp$ to u: in relativistic dynamics all forces are velocity-dependent. What was fairly exceptional in non-relativistic dynamics (where $\boldsymbol{F}_{\text{damping}} = -b\,\boldsymbol{v}$ and $\boldsymbol{F}_{\text{magnetic}} = (q/c)\,\boldsymbol{v}\times\boldsymbol{B}$ are the only vecocity-dependent forces that come readily to mind) is in relativistic dynamics *universal*. Symbolically

$$K = K(u,\ldots)$$

where the dots signify such other variables as may in particular cases enter into the construction of K. The simplest case—which is, as we shall see, the case of electrodynamical interest—arises when K depends *linearly* on u:

$$K_\mu = A_{\mu\nu}u^\nu \tag{284.1}$$

where $(K,u) = A_{\mu\nu}u^\mu u^\nu = 0$ forces the quantities $A_{\mu\nu}(\ldots)$ to satisfy the

$$\text{antisymmetry condition} \quad : \quad A_{\mu\nu} = -A_{\nu\mu} \tag{284.2}$$

K-vectors that depend *quadratically* upon u exist in much greater variety: the following example

$$K_\mu = \phi^\alpha(x)\big[c^2 g_{\alpha\mu} - u_\alpha u_\mu\big]$$

figured prominently in early (unsuccessful) efforts to construct a special relativistic theory of gravitation.[152,153]

If K is notated

$$K = \begin{pmatrix} K^0 \\ \boldsymbol{K} \end{pmatrix} \tag{285}$$

then (283)—written $\gamma(K^0 c - \boldsymbol{K}\cdot\boldsymbol{v}) = 0$—entails

$$K^0 = \tfrac{1}{c}\boldsymbol{K}\cdot\boldsymbol{v} \quad : \quad \text{knowledge of } \boldsymbol{K} \text{ determines } K^0 \tag{286}$$

It follows in particular that

$$K^0 = 0 \quad \text{in the (momentary) } \textit{rest frame} \tag{287}$$

It is, of course, the *non*-zero value of $\boldsymbol{K}$ that causes the particle to take leave of (what a moment ago was) the rest frame. Borrowing notation from (275) and

---

[152] This work (∼1912) is associated mainly with the name of G. Nordström, but for a brief period engaged the enthusiastic attention of Einstein himself: see page 144 in Pauli,[135] and also A. O. Barut, *Electrodynamics and Classical Theory of Fields and Particles* (1965), page 56; A. Pais, *Subtle is the Lord: The Science and Life of Albert Einstein* (1982), page 232.

[153] For further discussion of the "general theory of K-construction" see my RELATIVISTIC DYNAMICS (1967), pages 13–22.

(285), the Minkowski equation (274) becomes

$$\begin{pmatrix} K^0 \\ \boldsymbol{K} \end{pmatrix} = \gamma \tfrac{d}{dt} \begin{pmatrix} \gamma mc \\ \gamma m\boldsymbol{v} \end{pmatrix} \tag{288}$$

where use has been made once again of $\frac{d}{d\tau} = \gamma\frac{d}{dt}$. In the *non-relativistic limit*

$$\begin{pmatrix} 0 \\ \boldsymbol{F} \end{pmatrix} \overset{\downarrow}{=} \begin{pmatrix} 0 \\ \frac{d}{dt}m\boldsymbol{v} \end{pmatrix} \longleftarrow \text{Newtonian!}$$

where we have written

$$\boldsymbol{F} \equiv \lim_{c\uparrow\infty} \boldsymbol{K} \tag{289}$$

to account for such c-factors as may lurk in the construction of $\boldsymbol{K}$. We are used to thinking of the "non-relativistic limit" as an *approximiation* to relativistic physics, but at this point it becomes appropriate to remark that

> In fully relativistic particle dynamics the "non-relativistic limit" becomes literally effective in the momentary rest frame.

The implication is that *if we knew* the force $\boldsymbol{F}$ experienced by a particle *at rest* then we could by Lorentz transformation *obtain the Minkowski force* K *active upon a* <u>*moving*</u> *particle*:

$$\begin{pmatrix} K^0 \\ \boldsymbol{K} \end{pmatrix} = \mathbb{A}(\boldsymbol{\beta}) \begin{pmatrix} 0 \\ \boldsymbol{F} \end{pmatrix} \tag{290}$$

Reading from (210.1) it follows more particularly that

$$\left.\begin{aligned} K^0 &= \gamma\tfrac{1}{c}\boldsymbol{v}\cdot\boldsymbol{F} \\ \boldsymbol{K} &= \boldsymbol{F} + \big\{(\gamma-1)(\boldsymbol{v}\cdot\boldsymbol{F})/v^2\big\}\,\boldsymbol{v} = \boldsymbol{F}_\perp + \gamma\boldsymbol{F}_\parallel \end{aligned}\right\} \tag{291}$$

from which, it is gratifying to observe, one can recover both (289) and (286).

We stand not (at last) in position to trace the details of the program proposed[154] in a specifically electrodynamical setting by Einstein. Suppose that a <u>charged particle</u> experiences a force

$$\boldsymbol{F} = q\boldsymbol{E} \;\;:\;\; \boldsymbol{E} \equiv \text{electrical field in the particle's rest frame}$$

Then

$$\boldsymbol{K} = q(\boldsymbol{E}_\perp + \gamma\boldsymbol{E}_\parallel)$$

But from the field transformation equations (263) it follows that

$$\begin{aligned} \boldsymbol{E}_\perp &= \gamma(\boldsymbol{E} + \boldsymbol{\beta}\times\boldsymbol{B})_\perp \\ \boldsymbol{E}_\parallel &= (\boldsymbol{E} + \boldsymbol{\beta}\times\boldsymbol{B})_\parallel \end{aligned}$$

where $\boldsymbol{E}$ and $\boldsymbol{B}$ refer to *our* perception of the electric and magnetic fields at the particle's location, and $\boldsymbol{\beta}$ to our perception of the particle's velocity. So (because the γ-factors interdigitate so sweetly) we have

$$\boldsymbol{K} = \gamma q(\boldsymbol{E} + \tfrac{1}{c}\boldsymbol{v}\times\boldsymbol{B}) \tag{292}$$

---

[154] See again page 186.

But (288) supplies $\boldsymbol{K} = \gamma\frac{d}{dt}(\gamma m\boldsymbol{v})$, so (dropping the γ-factors on left and right) we have[155]

$$q(\boldsymbol{E} + \tfrac{1}{c}\boldsymbol{v}\times\boldsymbol{B}) = \tfrac{d}{dt}(\gamma m\boldsymbol{v}) \tag{293}$$

This famous equation describes the <u>relativistic motion of a charged particle in an impressed electromagnetic field</u> (no radiation or radiative reaction), and is the upshot of[156] the *Lorentz force law*—obtained here not as an it ad hoc assumption, but as a forced consequence of

- some general features of relativistic particle dynamics
- the transformation properties of electromagnetic fields
- the operational definition of $\boldsymbol{E}$ . . . all fitted into
- Einstein's "go to the frame of the particle" program (pages 186 & 189).

Returning with (292) to (286) we obtain

$$K^0 = \tfrac{1}{c}\gamma q\boldsymbol{E}\cdot\boldsymbol{v} \tag{294}$$

so the *Minkowski 4-force experienced by a charged particle in an impressed elecromagnetic field* becomes

$$\begin{aligned} K = \begin{pmatrix} K^0 \\ \boldsymbol{K} \end{pmatrix} &= \gamma q \begin{pmatrix} \frac{1}{c}\boldsymbol{E}\cdot\boldsymbol{v} \\ \boldsymbol{E} + \frac{1}{c}\boldsymbol{v}\times\boldsymbol{B} \end{pmatrix} \\ &= (q/c)\begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix}\begin{pmatrix} \gamma c \\ \gamma v_1 \\ \gamma v_2 \\ \gamma v_3 \end{pmatrix} \\ &\downarrow \\ K^\mu &= (q/c)F^\mu{}_\nu u^\nu \end{aligned} \tag{295}$$

We are brought thus to the striking conclusion that *the electromagnetic Minkowski force is, in the described at (284), <u>simplest possible</u>*.

The theory in hand descends from $\boldsymbol{F} = m\ddot{\boldsymbol{x}}$, and might plausibly be called "relativistic Newtonian dynamics." Were we to continue this discussion we might expect to busy ourselves with the construction of

- a "relativistic Lagrangian dynamics"
- a "relativistic Hamiltonian dynamics"
- a "relativistic Hamilton-Jacobi formalism"
- "relativistic variational principles," *etc*.

—all in an effort to produce a full-blown "relativistic dynamics of particles." The subject[157] is, however, a minefield, and must be persued with much greater delicacy than the standard texts suggest. Relativistic particle mechanics

[155] PROBLEM 50.
[156] See again equation (67) on page 35.
[157] See the notes[153] already cited.

remains in a relatively primitive state of development because many of the concepts central to non-relativistic mechanics are—for reasons having mainly to do with the breakdown of non-local simultaneity—in conflict with the principle of relativity. But while the relativistic theory of interacting particles presents awkwardnesses at every turn, the relativistic theory of interacting fields unfolds with great ease and naturalness: *it appears to be a lesson of relativity that we should adopt a field-theoretic view of the world*.

We have already in hand a relativistic particle mechanics which, though rudimentary, is sufficient to our electrodynamic needs. Were we to pursue this subject we would want to look to the problem of *solving* Minkowski's equation of motion (274) isn illustrative special cases . . . any short list of which would include

- the relativistic harmonic oscillator
- the relativistic Kepler problem
- motion in a (spatially/temporally) constant electromagnetic field.

This I do on pages 245–275 of ELECTRODYNAMICS (1972/73), where I give also many references. The most significant point to emerge from that discussion is that *distinct relativistic systems can have the same non-relativistic limit; i.e., that constructing the relativistic generalization of a non-relativistic system is an inherently ambiguous process*. For the present I must be content to examine two physical questions that have come already to the periphery of our attention.

HYPERBOLIC MOTION: THE "TWIN PARADOX" We—who call ourselves O—are inertial. A second observer Q sits on a mass point m which we see to be moving with (some dynamically possible but otherwise) arbitrary motion along our x-axis. I am tempted to say that Q rides in a little rocket, but that would entail (on physical grounds extraneous to my main intent) the temporal variability of m: let us suppose therefore that Q moves (accelerates) because m is acted on by impressed forces. In any event, we imagine Q to be equipped with

- a **clock** which—since co-moving—measures proper time τ
- an **accelerometer**, with output g. If Q were merely a passenger then $g(\tau)$ would constitute a king of log. But if Q were a rocket captain then $g(\tau)$ might describe his flight instructions, his prescribed "throttle function."

Finally, let O_τ designate the *inertial* observer who at proper time τ sees O_τ to be instantaneously at rest: spacetime points to which we assign coordinates x are by O_τ assigned coordinates x_τ. Our interest attaches initially to questions such as the following: Given the throttle function $g(\tau)$,

1) What is the boost $\Lambda(\tau)$ associated with $O \leftarrow O_\tau$?
2) What is the functional relationship between t and τ?
3) What are the functions
   $x(t)$ that describes our sense of Q's position at time t
   $\beta(t)$ that describes our sense of Q's velocity at time t
   $a(t)$ that describes our sense of Q's acceleration at time t?

Since O_τ sees Q to be momentarily resting at O_τ's origin we have

$$\begin{aligned} u(\tau) &= \mathbb{A}(\tau)\begin{pmatrix} c \\ 0 \end{pmatrix} && \text{by (269)} \\ a(\tau) &= \mathbb{A}(\tau)\begin{pmatrix} 0 \\ g(\tau) \end{pmatrix} && \text{by (269)} \end{aligned} \tag{296}$$

But

$$= \frac{du(\tau)}{d\tau} = \frac{d\mathbb{A}(\tau)}{d\tau}\begin{pmatrix} c \\ 0 \end{pmatrix}$$

We know, moreover, that[158]

$$\mathbb{A}(\tau) = e^{A(\tau)\mathbb{J}} \qquad \text{with} \quad \mathbb{J} \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \;\; A(\tau) = \tanh^{-1}\beta(\tau)$$

so

$$\begin{aligned} \frac{d\mathbb{A}(\tau)}{d\tau} &= \mathbb{A}(\tau)\cdot\frac{dA(\tau)}{d\tau}\mathbb{J} \\ &\quad \frac{dA(\tau)}{d\tau} = \frac{1}{1-\beta^2}\frac{d\beta}{d\tau} \end{aligned}$$

Returning with this information to (296) we obtain

$$\frac{1}{1-\beta^2}\frac{d\beta}{d\tau} = \tfrac{1}{c}g(\tau)$$

where integration of $dt/d\tau = \gamma$ supplies

$$\tau = \int^t \sqrt{1-\beta^2(t')}\,dt' \tag{297}$$

Given $g(\bullet)$, our assignment therefore is to solve

$$\Big[\frac{1}{1-\beta^2(t)}\Big]^{\frac{3}{2}}\frac{d\beta(t)}{dt} = \tfrac{1}{c}g\Big(\int^t \sqrt{1-\beta^2(t')}\,dt'\Big) \tag{298}$$

for $\beta(t)$: a final integration would then supply the $x(t)$ that describes our perception of Q's worldline. The problem presented by (298) appears in the general case to be hopeless . . . but let us at this point **assume** that the throttle function has the simple structure

$$g(\tau) = g \quad : \quad \textit{constant}$$

[158] See again pages 138 and 139.

The integrodifferential equation (298) then becomes a differential equation which integrates at once: assuming $\beta(0) = 0$ we obtain $\beta/\sqrt{1-\beta^2} = (g/c)t$ giving

$$\beta(t) = \frac{t}{\sqrt{(c/g)^2 + t^2}} \tag{299.1}$$

By integration we therefore have[159]

$$\big[x(t) - x(0) + (c^2/g)\big]^2 - (ct)^2 = (c^2/g)^2 \tag{299.2}$$

and

$$\tau(t) = (c/g)\sinh^{-1} gt/c \tag{299.3}$$

while expansion in powers of gt/c (which presumes $gt \ll c$) gives

$$\begin{aligned} v(t) &= \underbrace{\phantom{x(0)+{}}gt}\big[1 - \tfrac{1}{2}(gt/c)^2 + \cdots\big] \\ x(t) &= x(0) + \tfrac{1}{2}gt^2\big[1 - \tfrac{1}{4}(gt/c)^2 + \cdots\big] \\ \tau(t) &= \underbrace{\phantom{x(0)+\tfrac{1}{2}g}t\big[1}_{\text{conform to non-relativistic experience}} - \tfrac{1}{6}(gt/c)^2 + \cdots\big] \end{aligned} \tag{300}$$

According to (299.2) we see Q to trace out (not a parabolic worldline, as in non-relativistic physics, but) a *hyperbolic worldline*, as shown in Figure 71.

The results now in hand place us in position to construct concrete illustrations of several points that have been discussed thus far only as vague generalities:

1. Equation (299.1) entails

$$\gamma(t) = \sqrt{1 + (gt/c)^2}$$

which places us in position to construct an explicit description

$$\underset{\substack{\uparrow \\ t\,=\,(c/g)\sinh g\tau/c\,,\ \text{by (299.3)}}}{\mathbb{A}(t)} = \gamma(t)\begin{pmatrix} 1 & \beta(t) \\ \beta(t) & 1 \end{pmatrix} \quad : \quad \text{recall (201)}$$

of the Lorentz matrix that achieves $O \leftarrow O_\tau$, and thus to answer a question posed on page 199. We can use that information to (for example) write

$$K(t) = ma(t) = \mathbb{A}(t)\begin{pmatrix} 0 \\ mg \end{pmatrix}$$

to describe the relationship between

$$\begin{aligned} K(t) &\equiv \text{our perception of the Minkowski force impressed upon } m \\ \begin{pmatrix} 0 \\ mg \end{pmatrix} &\equiv O\text{'s perception of that Minkowski force} \end{aligned}$$

[159] PROBLEM 51.

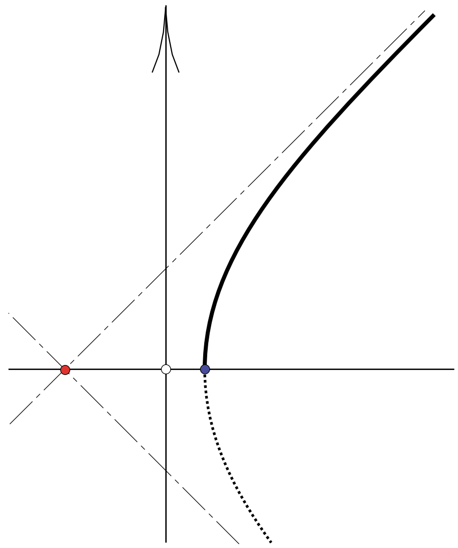

FIGURE 71: *Our (inertial) representation of the hyperbolic worldline of a particle which initially rests at the point* $x(0)$ *but moves off with (in its own estimation) constant acceleration* g*. With characteristic time* c/g *it approaches (and in Galilean physics would actually achieve) the speed of light. If we assign to* g *the comfortable value 9.8 meters/second*^{2} *we find* $c/g = 354.308$ *days.*

2.* In (299.2) set $x(0) = 0$. The resulting spacetime hyperbola is, by notational adjustment $\frac{1}{2}\lambda \mapsto c^2/g$, identical to that encountered at the middle of page 176: our perception of Q's worldline is a *conformal transform* Q's own perception of her (from her point of view trivial) worldline. If Q elected to pass her time doing electrodynamics she would—though non-inertial—use equations that are structurally identical to the (conformally covariant) equations that we might use to describe those same electrodynamical events.

3. O is inertial, content to sit home at $x = 0$. Q—O's twin—is an astronaut, who at time $t = 0$ gives her brother a kiss and sets off on a flight along the x-axis, on which her instruction is to execute the following throttle function:

$$g(\tau) = \begin{cases} +g & : \quad 0 < \tau < \frac{1}{4}\mathfrak{T} \\ -g & : \quad \frac{1}{4}\mathfrak{T} < \tau < \frac{3}{4}\mathfrak{T} \\ +g & : \quad \frac{3}{4}\mathfrak{T} < \tau < \mathfrak{T} \end{cases}$$

Pretty clearly, O's representation of Q's worldline will be assembled from four hyperbolic segments (Figure 72), each of duration $(c/g)\sinh g\mathfrak{T}/4c$. At

* This remark will be intelligible only to those brave readers who ignored my recommendation that they skip §6.

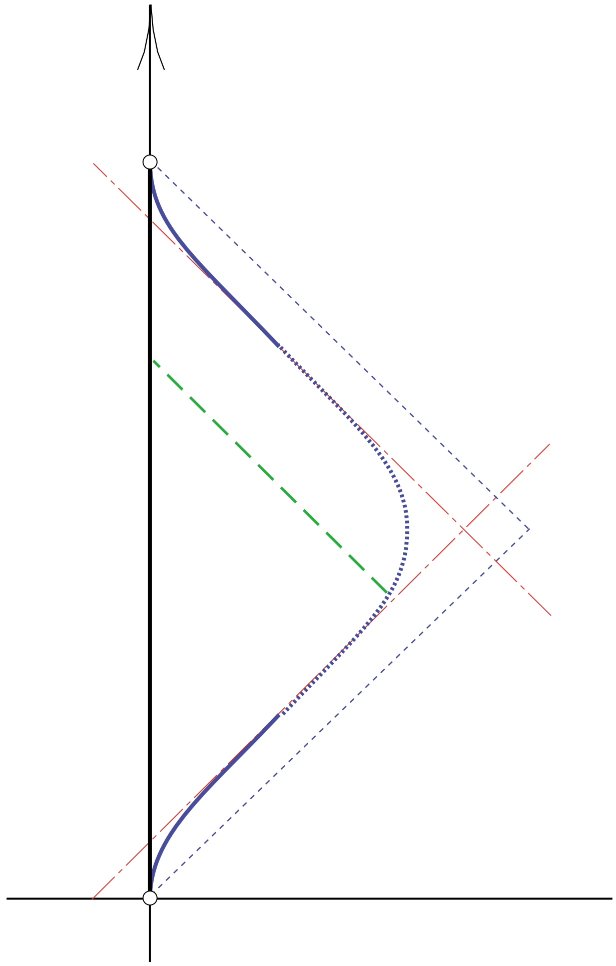

Figure 72: *Inertial observer O's representation of the rocket flight of his twin sister Q. If $\mathcal{T} \gg 4c/g$ then O will see Q to be moving much of the time at nearly the speed of light (hyperbola approaches its asymptote). The dashed curve represents the flight of a lightbeam that departs/returns simultaneously with Q.*

the moment of her return the clock on Q's control panel will read $\mathcal{T}$, but according to O's clock the

$$\text{return time} = \mathcal{T}\cdot(4c/g\mathcal{T})\sinh g\mathcal{T}/4c = \begin{cases} > \mathcal{T} \\ \sim \mathcal{T} \quad \text{only if} \quad \mathcal{T} \ll 4c/g \end{cases} \tag{301.1}$$

and Q's adventure will have taken her to a turn-around point lying[160] a

[160] Work from (299.2).

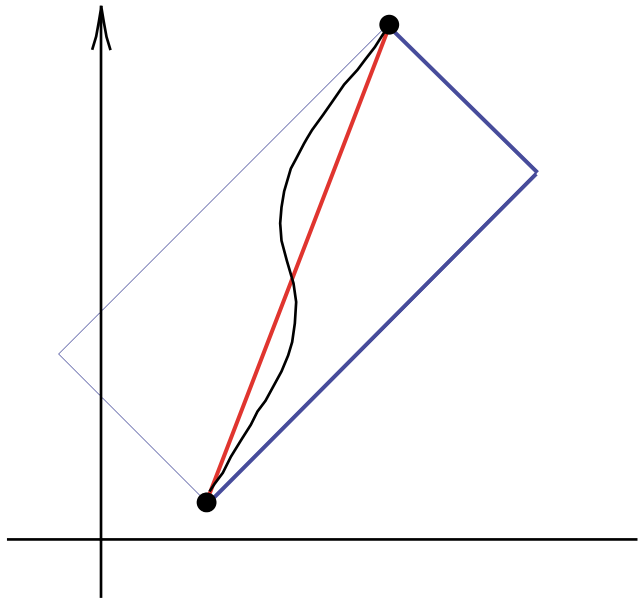

FIGURE 73: *Particle worldlines $\bullet \rightarrow \bullet$ all lie within the confines of the blue box (interior of the spacetime region bounded by the lightcones that extend forward form the lower vertex, and backward from the later vertex). The red trajectory—though shortest-possible in the Euclidean sense—is longest-possible in Minkowski's sense, while the blue trajectory has zero length. The "twin paradox" hinges on the latter fact. The acceleration experienced by the rocket-borne observer Q is, however, not abrupt (as at the kink in the blue trajectory) but evenly distributed.*

$$\text{distance} = 2\Big\{\sqrt{(ct)^2 + (c^2/g)^2} - c^2/g\Big\} \qquad t = \tfrac{1}{4}(\text{return time}) \tag{301.2}$$

away. For brief trips we therefore have

$$\text{distance} = 2(c^2/g)\Big\{\sqrt{1+(gt/c)^2} - 1\Big\} = 2\cdot\Big\{\tfrac{1}{2}gt^2 + \cdots\Big\}$$

while for long trips

$$\text{distance} = 2ct\underbrace{\Big\{\sqrt{1+(c/gt)^2} - (c/gt)\Big\}}_{\text{this factor is always positive, always } <1, \text{ and approaches unity as } t\uparrow\infty}$$

—both of which make good intuitive sense.[161] Notice (as Einstein—at the end of §4 in his first relativity paper—was the first to do) that

$$Q \text{ is younger than } O \text{ upon her return}$$

and that this surprising fact can be attributed to a basic metric property of spacetime (Figure 73).[162] The so-called **twin paradox** arises when one argues that from Q's point of view it is O who has been doing the accelerating, and who should return younger . . . and they can't *both* be younger! But those who pose the "paradox" misconstrue the meaning of the "relativity of motion." Only O remained inertial throughout the preceding exercise, and only Q had to purchase rocket fuel . . . and those facts break the supposed "symmetry" of the situation. The issue becomes more interesting with the observation that *we* have spent our lives in (relative to the inertial frames falling through the floor) "a rocket accelerating upward with acceleration g" (but have managed to do so without an investment in "fuel"). Why does our predicament not more nearly resemble the the predicament of Q than of O?[163]

CURRENT-CHARGE INTERACTION FROM TWO POINTS OF VIEW We possess a command of relativistic electrodynamics/particle dynamics that is now so complete that we can contemplate detailed analysis of the "asymmetries" that served to motivate Einstein's initial relativistic work. The outline of the illustrative discussion which follows was brought to my attention by Richard Crandall.[164] The discussion involves rather more than mere "asymmetry:" on its face it involves a "paradox." The system of interest, and the problem it presents, are described in Figure 74. The observer O who is *at rest with respect to the wire* sees an electromagnetic field which (at points exterior to the wire) can be described

$$\boldsymbol{E} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{B} = \begin{pmatrix} 0 \\ -Bz/R \\ +By/R \end{pmatrix}$$

where $B = I/2\pi cR$ and $R = \sqrt{y^2 + z^2}$. The Minkowski 4-force experienced by q therefore becomes (see again (295))

$$\begin{pmatrix} K^0 \\ K^1 \\ K^2 \\ K^3 \end{pmatrix} = (q/c) \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & By/R & Bz/R \\ 0 & -By/R & 0 & 0 \\ 0 & -Bz/R & 0 & 0 \end{pmatrix} \begin{pmatrix} \gamma c \\ \gamma v \\ 0 \\ 0 \end{pmatrix}$$

---

[161] PROBLEM 52.

[162] PROBLEM 53.

[163] See at this point C. W. Sherwin, "Some recent experimental tests of the clock paradox," Phys. Rev. **120**, 17 (1960).

[164] For parallel remarks see §5.9 in E. M. Purcell's *Electricity & Magnetism: Berkeley Physics Course–Volume 2* (1965) and §13.6 of *The Feynman Lectures on Physics–Volume 2* (1964).

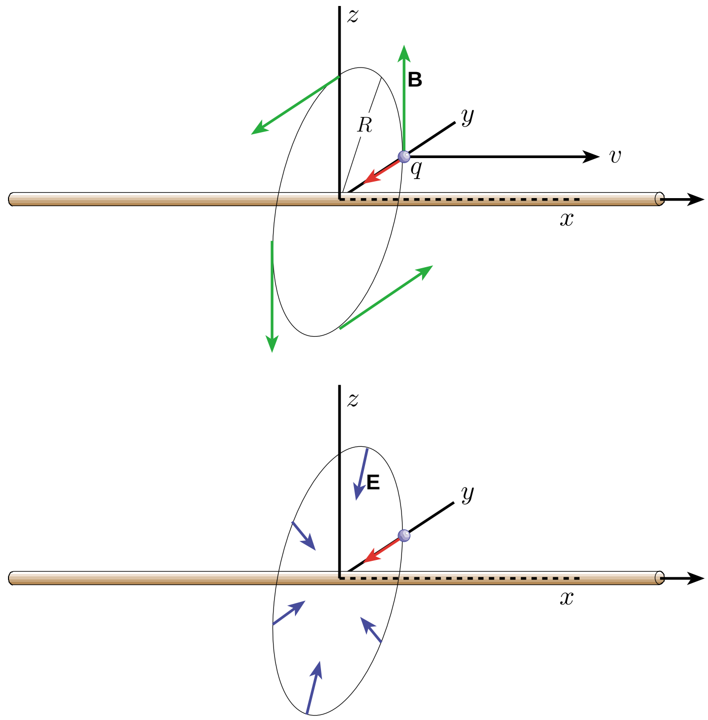

Figure 74: *At top: O's view of the system of interest . . . and at bottom: O's view. O—at rest with respect to a cylindrical conductor carrying current I—sees a charge q whose initial motion is parallel to the wire. He argues that the wire is wrapped round by a solenoidal magnetic field, so the moving charge experiences a* $(\boldsymbol{v}\times\boldsymbol{B})$*-force directed toward the wire, to which the particle responds by veering toward and ultimately impacting the wire. O is (initially) at rest with respect to the particle, so must attribute the impact an electrical force. But electrical forces arise (in the absence of time-dependent magnetic fields) only from charges. The nub of the problem*: How do uncharged current-carrying wires manage to appear charged to moving observers?

So we have

$$\begin{pmatrix} K^0 \\ K^1 \\ K^2 \\ K^3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -(\gamma qBv/c)y/R \\ -(\gamma qBv/c)z/R \end{pmatrix} = \begin{pmatrix} K^0 \\ \boldsymbol{K} \end{pmatrix}$$

according to which $\boldsymbol{K}$ is directed radially toward the wire. To describe this same physics O—who sees O to be moving *to the left* with speed v—writes

$$K = \Lambda K = (q/c)\cdot \underbrace{\Lambda \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & By/R & Bz/R \\ 0 & -By/R & 0 & 0 \\ 0 & -Bz/R & 0 & 0 \end{pmatrix} \Lambda^{-1}}_{\mathbb{F}} \cdot \underbrace{\Lambda \begin{pmatrix} \gamma c \\ \gamma v \\ 0 \\ 0 \end{pmatrix}}_{u}$$

with

$$\Lambda = \gamma \begin{pmatrix} 1 & -\beta & 0 & 0 \\ -\beta & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Straightforward computation supplies

$$= (q/c)\cdot \begin{pmatrix} 0 & 0 & -\beta\gamma By/R & -\beta\gamma By/R \\ 0 & 0 & +\ \gamma By/R & +\ \gamma Bz/R \\ -\beta\gamma By/R & \gamma By/R & 0 & 0 \\ -\beta\gamma Bz/R & \gamma Bz/R & 0 & 0 \end{pmatrix} \begin{pmatrix} c \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 0 \\ -(\gamma qBv/c)y/R \\ -(\gamma qBv/c)z/R \end{pmatrix} = \begin{pmatrix} K^0 \\ \boldsymbol{K} \end{pmatrix}$$

While O saw only a $\boldsymbol{B}$-field, it is clear from the computed structure of $\mathbb{F}$ that O sees both a $\boldsymbol{B}$-field (γ times stronger that O's) and an $\boldsymbol{E}$-field. We have known since (210.2) that

$$\text{(spatial part of any 4-vector)}_\perp \text{ boosts by invariance}$$

so (since $\boldsymbol{K} \perp \boldsymbol{v}$) are not surprised to discover that

$$\boldsymbol{K} = \boldsymbol{K}, \text{ but observe that } \begin{cases} O \text{ considers } \boldsymbol{K} \text{ to a magnetic effect} \\ O \text{ considers } \boldsymbol{K} \text{ to an electric effect} \end{cases}$$

More specifically, O sees (Figure 74) a centrally-directed electric field of just the strength

$$E = \beta\gamma B = \beta\gamma I/2\pi cR$$

that would arise from an infinite line charge linear density

$$\lambda = -\beta\gamma I/c$$

The question now before us: *How* does the current-carrying wire acquire, in O's estimation, a net charge? An answer of sorts can be obtained as follows: Assume (in the interest merely of simplicity) that the current is uniformly distributed on the wire's cross-section:

$$I = ja \quad \text{where } a \equiv \pi r^2 = \text{cross-sectional area}$$

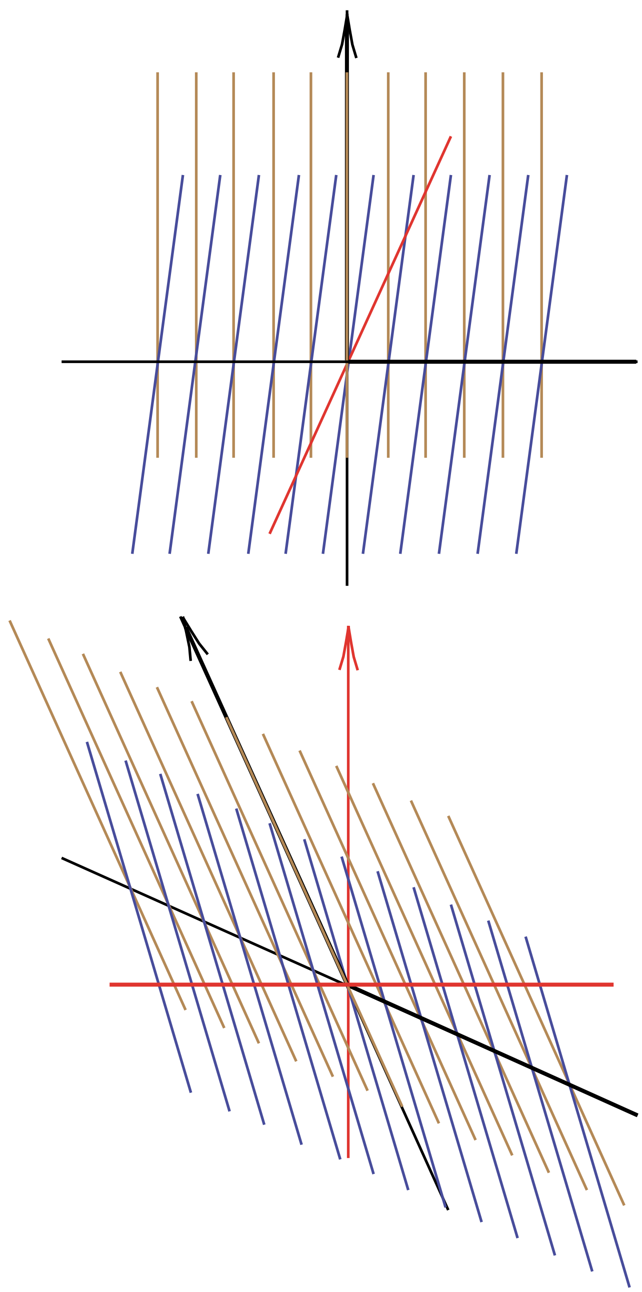

FIGURE 75: *O's representation of current flow in a stationary wire and (below) the result of Lorentz transforming that diagram to the frame of the passing charge q. For interpretive commentary see the text.*

To describe the *current 4-vector interior to the wire* O therefore writes

$$j = \begin{pmatrix} 0 \\ I/a \\ 0 \\ 0 \end{pmatrix}$$

O, on the other hand, writes the Lorentz transform of j:

$$j \equiv \begin{pmatrix} c\rho \\ \boldsymbol{j} \end{pmatrix} = \mathbb{\Lambda}\, j = \begin{pmatrix} -\beta\gamma I/a \\ \gamma I/a \\ 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad \rho = -\beta\gamma I/ac$$

O and O assign identical values to the cross-sectional area

$$a = a \quad \text{because cross-section} \perp \boldsymbol{v}$$

so O obtains

$$\begin{aligned} \lambda \equiv \text{charge per unit length} &= \rho a \\ &= -\beta\gamma I/c \end{aligned}$$

—in precise agreement with the result deduced previously. Sharpened insight into the mechanism that lies at the heart of this counterintuitive result can be gained from a comparison of the spacetime diagrams presented in Figure 75. At top we see O's representation of current in a stationary wire: negatively ionized atoms stand in place, positive charges drift in the direction of current flow.[165] In the lower figure we see how the situation presents itself to an observer O who is moving with speed v in a direction parallel to the current flow. At any instant of time (look, for example, to his $x^0 = 0$ timeslice, drawn in red) O sees ions and charge carriers to have distinct linear densities . . . the reason being that she sees ions and charge carriers to be moving with distinct speeds, and the intervals separating one ion from the next, one charge carrier from the next to be Lorentz contracted by distinct amounts. O's charged wire is, therefore, a *differential Lorentz contraction* effect. That such a small velocity differential

$$\text{drift velocity relative to ions} \sim 10^{-11}c$$

can, from O's perspective, give rise to a measureable net charge is no more surprising than that it can, from O's perspective, give rise to a measureable net current: both can be attributed to the fact that an awful *lot* of charges participate in the drift.

---

[165] O knows perfectly well that in point of physical fact the ionized atoms are positively charged, the current carriers negatively charged, and their drift opposite to the direction of current flow: the problem is that Benjamin Franklin did not know that. But the logic of the argument is unaffected by this detail.

Just about *any* electro-mechanical system would yield similar asymmetries/ “paradoxes” when analysed by alternative inertial observers O and O. The preceding discussion is in all respects typical, and serves to illustrate two points of general methodological significance:

- The formal mechanisms of (manifestly covariant) relativistic physics are so powerful that they tend to lead one *automatically* past conceptual difficulties of the sort that initially so bothered Einstein, and (for that very reason) . . .
- They tend, when routinely applied, to divert one’s attention from certain (potentially quite useful) physical insights: there exist points of physical principle which relativistic physics illuminates only when explicitly interrogated.

When using powerful tools one should always wear goggles.