RELATIVISTIC ADDITION AND GROUP THEORY

KEITH CONRAD

1. INTRODUCTION

For three particles P, Q, R traveling on a straight line, let v_{PQ} be the (relative) velocity of P as measured by Q, and define v_{QR}, v_{PR} similarly.

According to classical mechanics, the velocity v of a particle moving on a line can be an arbitrary real number, and relative velocities add by the simple formula

$$v_{PR} = v_{PQ} + v_{QR}.$$

On the other hand, the special theory of relativity says velocities are restricted to a bounded range, $-c < v < c$, where c is the speed of light (whose value of course depends on the choice of units, and it is convenient to choose them so $c = 1$, but we won't do that.) The relativistic addition formula for relative velocities along a line is

$$(1.1) \qquad v_{PR} = \frac{v_{PQ} + v_{QR}}{1 + (v_{PQ}v_{QR}/c^2)}.$$

Example 1.1. Suppose Q moves away from R at speed $(1/2)c$ and P moves away from Q in the same direction at speed $(1/2)c$. Then $v_{QR} = (1/2)c$ and $v_{PQ} = (1/2)c$, so (1.1) says

$$v_{PR} = \frac{v_{PQ} + v_{QR}}{1 + (v_{PQ}v_{QR}/c^2)} = \frac{c}{1 + (1/4)} = \frac{4}{5}c < c.$$

Example 1.2. If P moves away from Q at speed $(3/4)c$ and R moves away from Q in the opposite direction at speed $(1/2)c$, so $v_{PQ} = (3/4)c$ and $v_{RQ} = -(1/2)c$, then $v_{QR} = (1/2)c$ and (1.1) says

$$v_{PR} = \frac{v_{PQ} + v_{QR}}{1 + (v_{PQ}v_{QR}/c^2)} = \frac{(5/4)c}{1 + (3/8)} = \frac{10}{11}c < c.$$

There is an interesting algebraic similarity between the classical and relativistic velocity addition formulas. The classical model for velocity addition is the set of real numbers, combined under addition. Special relativity involves velocities in an interval $(-c, c)$ for some $c > 0$, combining them by the formula

$$(1.2) \qquad v \oplus w = \frac{v + w}{1 + vw/c^2}.$$

The operation $\oplus$ on $(-c, c)$ has properties similar to addition on $\mathbf{R}$ (check these!):

- Closure, *i.e.*, if $v_1, v_2 \in (-c, c)$ then $v_1 \oplus v_2 \in (-c, c)$.
- Identity for $\oplus$: $0 \oplus v = v \oplus 0 = v$ for $v \in (-c, c)$.
- Inverse of each v under $\oplus$ is $-v$: $v \oplus -v = -v \oplus v = 0$.
- Associativity: $(v_1 \oplus v_2) \oplus v_3 = v_1 \oplus (v_2 \oplus v_3)$ for all $v_1, v_2, v_3 \in (-c, c)$.

These say (1.2) is a group law on $(-c, c)$. Einstein, in his first paper on special relativity, wrote [1, §5] that "such parallel transformations – necessarily – form a group".[1]

The formula for $v \oplus w$ in (1.2) resembles the addition formula for the tangent function:

$$\tan(x + y) = \frac{\tan x + \tan y}{1 - (\tan x)(\tan y)}.$$

However, there is a minus sign here where there is a plus sign in (1.2).

The *hyperbolic* tangent is better than the tangent in this regard. The hyperbolic tangent function is given by the formula

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}} : \mathbf{R} \to (-1, 1)$$

and a graph is in Figure 1. This function is a bijection from $\mathbf{R}$ to $(-1, 1)$, with inverse

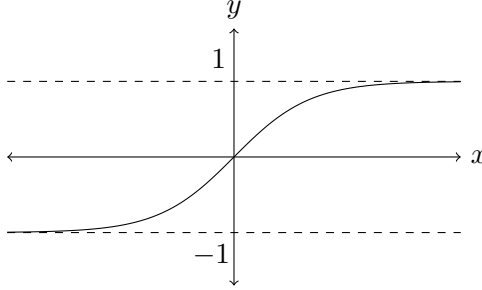$$\tanh^{-1}(x) = \frac{1}{2} \log\left(\frac{1 + x}{1 - x}\right).$$

FIGURE 1. Plot of $y = \tanh x$.

It is a matter of algebra to check that

$$\tanh(x + y) = \frac{\tanh x + \tanh y}{1 + (\tanh x)(\tanh y)}.$$

This is *exactly* like (1.2), up to some factors of c. Taking those into account, we find that the function $\varphi \colon \mathbf{R} \to (-c, c)$ where $\varphi(x) = c \tanh(x)$ satisfies

$$\varphi(x + y) = \varphi(x) \oplus \varphi(y).$$

Going in the other direction, let $\psi \colon (-c, c) \to \mathbf{R}$ by

$$\psi(v) = \frac{1}{2} \log\left(\frac{1 + v/c}{1 - v/c}\right).$$

This "rescaled" velocity turns $\oplus$ into addition:

(1.3) $$\psi(v \oplus w) = \psi(v) + \psi(w).$$

Thus, by a suitably clever transformation, velocities $v \in (-c, c)$ turn into rescaled velocities $\psi(v) \in \mathbf{R}$ and the nonintuitive operation $\oplus$ on $(-c, c)$ turns into *ordinary addition in* $\mathbf{R}$.

Example 1.3. If $v = (1/2)c$ and $w = (1/2)c$ then $\psi(v) = (1/2) \log 3$ and $\psi(w) = (1/2) \log 3$, while $\psi(v \oplus w) = \psi((4/5)c) = (1/2) \log 9 = \log 3 = \psi(v) + \psi(w)$.

---

[1]Einstein wrote the speed of light in [1] as V, a notation first used in Maxwell's work on electromagnetism.

Example 1.4. If $v = (3/4)c$ and $w = (1/2)c$ then $\psi(v) = (1/2)\log 7$ and $\psi(w) = (1/2)\log 3$, while $\psi(v \oplus w) = \psi((10/11)c) = (1/2)\log 21 = \psi(v) + \psi(w)$.

Remark 1.5. Writing a velocity $v \in (-c, c)$ as $c \tanh(x)$, the parameter x is called "rapidity" in physics. Combining two relativistic velocities (along the same direction) corresponds to adding their respective rapidities.

Could we have found ψ if we did not know the hyperbolic tangent? Yes! Using the ideas in [2, Chap. 6], we will prove that *every* (continuously differentiable) group law on an open interval in **R** can be rescaled to look like ordinary addition on **R**.

Remark 1.6. Instead of converting relativistic addition on $(-c, c)$ into ordinary addition on **R**, we can convert it into ordinary multiplication on $(0, \infty)$: the function $f \colon (-c, c) \to (0, \infty)$ where $f(v) = (1 + v/c)/(1 - v/c) = (c + v)/(c - v)$ is a bijection and $f(v \oplus w) = f(v)f(w)$ for all v and w in $(-c, c)$.

2. GROUP LAWS ON AN INTERVAL

Let $I \subset \mathbf{R}$ be an open interval with a group law $*$, so $*$ has the following properties:
- Closure. $x, y \in I \Longrightarrow x * y \in I$.
- Identity. There is $u \in I$ such that for all $x \in I$, $x * u = u * x = x$.
- Inverses. For all $x \in I$ there is some $i(x) \in I$ such that $x * i(x) = i(x) * x = u$.
- Associativity. For $x, y, z \in I$, $(x * y) * z = x * (y * z)$.

(We write the identity for $*$ as u rather than, say, e, since we're working with real numbers and don't want confusion with the real number 2.71828...)

It will be useful to write the operation $x * y$ in the notation of a function of two variables: $F(x, y) = x * y$. For example, the classical and relativistic velocity addition formulas are

$$F(v, w) = v + w, \ I = \mathbf{R}; \quad F(v, w) = \frac{v + w}{1 + vw/c^2}, \ I = (-c, c).$$

The above properties of $*$ take the following form in terms of F:
- $x, y \in I \Rightarrow F(x, y) \in I$.
- There is $u \in I$ such that for all $x \in I$, $F(x, u) = F(u, x) = x$.
- For all $x \in I$ there is some $i(x) \in I$ such that $F(x, i(x)) = F(i(x), x) = u$.
- For $x, y, z \in I$, $F(F(x, y), z) = F(x, F(y, z))$.

Our goal is to prove the following theorem.

Theorem 2.1. *If the group law $F(x, y) = x * y$ on an interval I has continuous partial derivatives, then there is a differentiable bijection $\ell \colon I \to \mathbf{R}$ that converts $*$ on I to ordinary addition on $\mathbf{R}$. That is, ℓ is a differentiable bijection with $\ell(x * y) = \ell(x) + \ell(y)$.*

We write the rescaling function as ℓ because we think about it as a 'logarithm' for $*$, just as the usual logarithm turns multiplication on $(0, \infty)$ into addition on **R**.

Theorem 2.1 will be proved by giving an explicit recipe for ℓ. To *discover* ℓ, let's assume it exists: $\ell(x * y) = \ell(x) + \ell(y)$ for all x and y in I. Rewriting this in the functional F-notation instead of the operator $*$-notation,

$$\ell(F(x, y)) = \ell(x) + \ell(y).$$

Assuming ℓ is differentiable, let's differentiate both sides of this equation with respect to x:

$$\ell'(F(x, y))F_1(x, y) = \ell'(x),$$

where we write $F_1(x, y)$ for $\partial F/\partial x$ (and $F_2(x, y) = \partial F/\partial y$). Setting $x = u$,

$$\ell'(y)F_1(u, y) = \ell'(u),$$

so we solve for $\ell'(y)$ and integrate:

$$(2.1) \qquad \ell(y) = \int_u^y \frac{\ell'(u)}{F_1(u, t)} \, \mathrm{d}t.$$

This is a possible formula for the rescaling function ℓ. The constant $\ell'(u)$, where u is the $*$-identity, is just an undetermined scaling factor that we will simply set equal to 1 once we return to rigorous definitions. Incidentally, when we integrated in (2.1), we didn't introduce an additive constant since we want $\ell(u) = 0$ (the $*$-identity should go the additive identity) and the integral formula (2.1) already takes care of that. Of more pressing interest is the validity of dividing by $F_1(u, t)$ in (2.1). Why is it never zero?

Lemma 2.2. *For all $t \in I$, $F_1(u, t) > 0$.*

Proof. Differentiate the associative law, $F(F(x, y), z) = F(x, F(y, z))$, with respect to x:

$$F_1(F(x, y), z)F_1(x, y) = F_1(x, F(y, z)).$$

Setting $x = u$, the $*$-identity,

$$(2.2) \qquad F_1(y, z)F_1(u, y) = F_1(u, F(y, z)) = F_1(u, y * z).$$

So if $F_1(u, y) = 0$ for some y, then $F_1(u, y * z) = 0$ for all z. Choose $z = i(y)$ to get $F_1(u, u) = 0$. But this is not true:

$$F(x, u) = x \text{ for all } x \Rightarrow F_1(x, u) = 1 \Rightarrow F_1(u, u) = 1.$$

So $F_1(u, y)$ is nonzero for every y. Since it equals 1 at $y = u$ and is continuous, it must always be positive by the Intermediate Value Theorem. $\qquad\square$

Lemma 2.2 allows us to divide by $F_1(u, t)$ for each $t \in I$, and we will do this often without explicitly appealing to the lemma each time.

Since $1/F_1(u, t)$ is continuous in t, hence integrable, we are justified in making the following *definition*, for each x in the interval I:

$$(2.3) \qquad \ell(x) \stackrel{\text{def}}{=} \int_u^x \frac{\mathrm{d}t}{F_1(u, t)}.$$

By the Fundamental Theorem of Calculus, ℓ is differentiable and

$$(2.4) \qquad \ell'(x) = \frac{1}{F_1(u, x)}.$$

In particular, $\ell'(u) = 1$.

Using (2.3) as our rescaling function, we now prove Theorem 2.1.

Proof. We need to check (i) $\ell(F(x, y)) = \ell(x) + \ell(y)$ and (ii) $\ell \colon I \to \mathbf{R}$ is a bijection.

For item (i), *fix $y \in I$*. We consider the x-derivatives of

$$\ell(F(x, y)), \qquad \ell(x) + \ell(y).$$

By (2.4), the derivative of the first function is

$$\ell'(F(x, y))F_1(x, y) = \frac{F_1(x, y)}{F_1(u, F(x, y))}.$$

Does this equal the x-derivative of the second function, namely $\ell'(x) = 1/F_1(u, x)$? Setting them equal, we want to consider:

$$F_1(x, y)F_1(u, x) \stackrel{?}{=} F_1(u, F(x, y)).$$

This is just (2.2) with x, y, z relabelled as u, x, y. Therefore $\ell(F(x, y))$ and $\ell(x) + \ell(y)$ have equal x-derivatives for all x, which means they differ by an additive constant. Since they are equal at $x = u$, the additive constant is 0. This verifies the first item: $\ell(F(x, y)) = \ell(x) + \ell(y)$.

For item (ii), bijectivity, since $\ell'(y) = 1/F_1(u, y) > 0$ we get ℓ is increasing, hence injective. To show surjectivity, note $\ell(I)$ is an interval by continuity. Choose $x \in I$, $x \neq u$. Since $\ell(x) + \ell(i(x)) = \ell(x * i(x)) = \ell(u) = 0$, $\ell(x)$ and $\ell(i(x))$ have opposite sign. For a positive integer n,

$$\ell(\overbrace{x * \cdots * x}^{n \text{ times}}) = n\ell(x), \quad \ell(\overbrace{i(x) * \cdots * i(x)}^{n \text{ times}}) = n\ell(i(x)),$$

As $n \to \infty$, one tends to ∞, the other to $-\infty$. Since $\ell(I)$ is an interval, we must have $\ell(I) = \mathbf{R}$. $\qquad\square$

Corollary 2.3. *When $x * y = F(x, y)$ in Theorem 2.1 has continuous partial derivatives, it is commutative.*

Proof. Commutativity was never used in the proof of Theorem 2.1, so commutativity of addition on $\mathbf{R}$ implies commutativity of $*$ on I. $\qquad\square$

Remark 2.4. Without using differentiability, it can be shown that every continuous group law $*$ on an open interval I has the form $x * y = f(f^{-1}(x) + f^{-1}(y))$ for some homeomorphism $f: \mathbf{R} \to I$.[2] In particular, every continuous group law on I is commutative.

3. EXAMPLES

Let's look at some examples, to see which functions rescale various group laws on an interval to the additive group of all real numbers.

Example 3.1. If $F(x, y) = x + y$ on $I = \mathbf{R}$, then $u = 0$, $F_1(0, x) = 1$, and

$$\ell(x) = \int_0^x \mathrm{d}t = x.$$

Example 3.2. If $F(v, w) = v \oplus w = \dfrac{v + w}{1 + vw/c^2}$ on $(-c, c)$, then $u = 0$, $F_1(0, v) = 1 - v^2/c^2$, and

$$
\begin{aligned}
\ell(v) &= \int_0^v \frac{\mathrm{d}t}{F_1(0, t)} \\
&= c^2 \int_0^v \frac{\mathrm{d}t}{c^2 - t^2} \\
&= \frac{c}{2} \int_0^v \left(\frac{1}{c - t} + \frac{1}{c + t} \right) \mathrm{d}t \\
&= \frac{c}{2} \log \left(\frac{1 + v/c}{1 - v/c} \right).
\end{aligned}
$$

---

[2]This is a special case of Hilbert's Fifth Problem.

This is the same as the rescaling function $\psi(v)$ we met at the beginning, up to a factor of c. Of course if the rescaling function $\ell(x)$ in Theorem 2.1 is multiplied by a nonzero constant, it has the same relevant properties (except $\ell'(u) \neq 1$.)

The following table, where v_{PR}^{rel} is v_{PR} computed according to the relativistic formula (1.1), gives in the last column the difference between classical and relativistic formulas for v_{PR}. Note there is significant relative error not only in the first row, when both v_{PQ} and v_{QR} are substantial fractions of the speed of light, but even in the second and third rows, when only v_{PQ} is near c. In the last row, the fourth column entry is about $.5625 v_{QR} = (9/16) v_{QR}$.

v_{PQ}	v_{QR}	v_{PR}^{rel}	$(v_{PQ} + v_{QR}) - v_{PR}^{\mathrm{rel}}$
$(3/4)c$	$(1/2)c$	$(10/11)c$	$.341c$
$(3/4)c$	$(1/100)c$	$.75434c$	$.00566c$
$(3/4)c$	$(1/1000)c$	$.750437c$	$.000563c$

For $v, w \in (-c, c)$, take v to be fixed and think of $v \oplus w$ as a function of w. Set $h(w) = v \oplus w = (v + w)/(1 + vw/c^2)$. For w/c small, a Taylor expansion of $h(w)$ at $w = 0$ yields

$$v \oplus w \approx h(0) + h'(0)w = v + \left(1 - \frac{v^2}{c^2}\right) w.$$

If not just w/c is small, but also v/c is small, then the coefficient of w is about 1, so $v \oplus w \approx v + w$. If v/c is not small, e.g., $v = (3/4)c$, then we see a deviation from the classical addition formula $v + w$ by an error of around $(v/c)^2 w$. This explains the error $\approx (9/16)v_{QR}$ in the table above, where $v = (3/4)c$ and $9/16 = (3/4)^2$.

Example 3.3. As a final example, consider $F(x, y) = xy$ on $I = (0, \infty)$. This is the group of positive real numbers under multiplication. (The multiplicative group of non-zero real numbers, rather than just the positives, is not an interval. Where does the proof of Theorem 2.1 break down if we try to apply it to all non-zero reals?) Here $u = 1$ and $F_1(1, x) = x$, so Theorem 2.1 tells us that a rescaling function that converts multiplication on $(0, \infty)$ to addition on $\mathbf{R}$ is

$$\ell(x) = \int_1^x \frac{\mathrm{d}t}{F_1(1, t)} = \int_1^x \frac{\mathrm{d}t}{t} = \log x.$$

Of course we already knew that $\log(xy) = \log x + \log y$, but it is interesting to see how we have rediscovered the logarithm by applying calculus to algebra.

Our discussion here has been entirely one-dimensional. In three dimensions, velocities and reference frames are no longer collinear, and the relativistic formula for velocity addition is rather complicated. See https://en.wikipedia.org/wiki/Velocity-addition_formula.

APPENDIX A. TECHNICAL REMARKS

We want to give a more conceptual perspective on the way $\ell(x)$ in Theorem 2.1 is related to the structure of the group law $*$ on I. First of all, since $\ell \colon I \to \mathbf{R}$ is a bijection, and thus invertible, we can apply ℓ^{-1} to both sides of $\ell(x * y) = \ell(x) + \ell(y)$ to see that

$$(A.1) \qquad\qquad x * y = \ell^{-1}(\ell(x) + \ell(y)).$$

So the group law $*$ can be reconstructed from ℓ and ordinary real addition.

By Corollary 2.3, $F(x, y) = F(y, x)$, and differentiating both sides of this formula with respect to x shows $F_1(x, y) = F_2(y, x)$ for all $x, y \in I$. Since ℓ is determined by the function

$F_1(u, t)$ (see 2.3), and the operation $x * y$ is determined by ℓ (see A.1), we see the operation $*$ is encoded in the function $F_1(u, t) = F_2(t, u)$.

Now we take a look at Taylor expansions. The function $F_1(u, x)$ appears in the first term of the Taylor expansion at $y = u$ of $x * y = F(x, y)$ for small y:

$$(A.2) \qquad x * y = F(x, y) \approx F(x, u) + F_2(x, u)(y - u) = x + F_1(u, x)(y - u).$$

Since $F_1(u, u) = 1$, for x and y near u we get from (A.2) that $x * y \approx x + y - u$. Therefore

$$x * y - u \approx (x - u) + (y - u).$$

If we change variables to make 0 the $*$-identity, then this says $*$ is approximately just addition when both variables are small. However, for y near u and x not-so-near u there is a deviation of $x * y$ from the simple law $x + y$, measured by the function $F_1(u, x)$. This deviation for all x and y near u has been used to reconstruct the operation $x * y$ for all x and all y. That is, the local structure of $*$ near the identity completely determines the global group law.

If you know about differential forms, then the following is of interest. Another way of stating (2.2) is in terms of the differential form $\omega = dt / F_1(u, t)$. For each $z \in I$ we have the function $\tau_z \colon I \to I$ given by right translation by z: $\tau_z(x) = F(x, z)$. This induces a map τ_z^* on differential forms on I, and (2.2) says $\tau_z^* \omega = \omega$. In other words, ω is a $*$-invariant differential form, and $\ell(x) = \int_u^x \omega$ is the integral of this $*$-invariant differential form along the path from the identity element to x. This description of the function ℓ ties it more closely to the group structure than (2.3) alone suggests.

Appendix B. Deriving relativistic momentum and kinetic energy

When a particle of mass m and velocity v moves along a line, its energy is entirely kinetic energy and is given by $E = \frac{1}{2}mv^2$. Its momentum is $p = mv$. These are formulas from classical mechanics, where two fundamental laws are conservation of energy and momentum for every closed system. The relativistic formulas for kinetic energy and momentum of a particle with mass m moving along a line at velocity v look quite different from their classical versions:

$$(B.1) \qquad E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} - mc^2, \quad p = \frac{mv}{\sqrt{1 - v^2/c^2}}.$$

Taking power series expansions at $v = 0$, the first nonzero term in each formula is $\frac{1}{2}mv^2$ and mv respectively, so the classical formulas are limiting versions of the relativistic formulas when $v \ll c$. Using arguments from [3] we will derive the relativistic formulas (B.1), in the sense that we will show these are reasonable definitions for kinetic energy and momentum in relativity.

Assume for a particle moving along a line with velocity v that it has an "energy" $E(v)$ and a "momentum" $p(v)$. How should E and p be related to each other? In classical mechanics Newton's second law says $F = dp/dt$. Also in classical mechanics, the work done on a particle by a force F over a distance dx is equal to both $F(x)dx$ and dE, which is the change in the energy of the particle over the distance dx. We assume all of this remains true relativistically, so

$$(B.2) \qquad dE = F(x)dx = \frac{dp}{dt}dx = \frac{dx}{dt}dp = v\,dp.$$

We will come back to this equation after we carry out a thought experiment to find another relation between p and E.

Suppose two particles P and Q are moving along a line with fixed velocities v_P and v_Q relative to an observer, and after undergoing a collision their velocities measured by that observer are $\widetilde{v}_P$ and $\widetilde{v}_Q$. If the energy of the system of these two particles is to be the same before and after the collision then we require

$$(B.3) \qquad\qquad E_P(v_P) + E_Q(v_Q) = E_P(\widetilde{v}_P) + E_Q(\widetilde{v}_Q),$$

where E_P and E_Q are the energies of P and Q.

Velocity is a relative notion. If the observer we already discussed is moving at velocity $w \neq 0$ with respect to a second observer (w could be positive or negative) then the second observer initially measures the velocities of the particles as $v_P \oplus w$ and $v_Q \oplus w$ and later measures them as $\widetilde{v}_P \oplus w$ and $\widetilde{v}_Q \oplus w$. The principle of relativity says the laws of physics, like conservation of energy, should be the same for two observers moving at constant velocity relative to each other, so an analogue of (B.3) should be true for the second observer:

$$(B.4) \qquad\qquad E_P(v_P \oplus w) + E_Q(v_Q \oplus w) = E_P(\widetilde{v}_P \oplus w) + E_Q(\widetilde{v}_Q \oplus w).$$

For our unknown energy function E, we'll form the Taylor expansion of $E(v \oplus w)$ around $w = 0$ and pay attention to the first two terms. Writing $v \oplus w = (v + w)/(1 + vw/c^2)$ in functional form as $F(v, w)$,

$$E(v \oplus w) = E(F(v, w)) = E(F(v, 0)) + \left.\frac{\partial E(F(v, w))}{\partial w}\right|_{w=0} w + O(w^2)$$

and the chain rule implies $\partial E(F(v, w))/\partial w = E'(F(v, w))(\partial F/\partial w) = E'(F(v, w))F_2(v, w)$. Therefore

$$\begin{aligned} E(v \oplus w) &= E(F(v, 0)) + E'(F(v, 0))F_2(v, 0)w + O(w^2) \\ &= E(v) + E'(v)\left(1 - \frac{v^2}{c^2}\right)w + O(w^2). \end{aligned}$$

Applying this to each term in (B.4), equating the constant terms gives us (B.3), which is not new. Equating the coefficients of w gives us

$$(B.5) \quad E_P'(v_P)\left(1 - \frac{v_P^2}{c^2}\right) + E_Q'(v_Q)\left(1 - \frac{v_Q^2}{c^2}\right) = E_P'(\widetilde{v}_P)\left(1 - \frac{\widetilde{v}_P^2}{c^2}\right) + E_Q'(\widetilde{v}_Q)\left(1 - \frac{\widetilde{v}_Q^2}{c^2}\right).$$

Setting $f(v) = E'(v)(1 - v^2/c^2)$, equation (B.5) says

$$f_P(v_P) + f_Q(v_Q) = f_P(\widetilde{v}_P) + f_Q(\widetilde{v}_Q),$$

which is a conversation law for the system of particles P and Q. Having already introduced a conversation of energy in (B.3), we *define* $f(v)$ to be relativistic momentum so that (B.5) can be called conversation of momentum. Having called this momentum, we should write it as $p(v)$ instead of $f(v)$, so set

$$(B.6) \qquad\qquad p(v) = E'(v)\left(1 - \frac{v^2}{c^2}\right) = \frac{dE}{dv}\left(1 - \frac{v^2}{c^2}\right).$$

Feeding (B.2) into this,

$$p(v) = v\frac{dp}{dv}\left(1 - \frac{v^2}{c^2}\right),$$

which can be rearranged in physicist's style to

$$\frac{dp}{p} = \frac{dv}{v(1 - v^2/c^2)}.$$

Decomposing the right side into partial fractions,

$$\frac{dp}{p} = \left(\frac{1}{v} + \frac{1}{2c}\frac{1}{1 - v/c} - \frac{1}{2c}\frac{1}{1 + v/c} \right) dv.$$

Forming antiderivatives,

$$\log|p| = \log|v| - \frac{1}{2}\log\left|1 - \frac{v}{c}\right| - \frac{1}{2}\log\left|1 + \frac{v}{c}\right| + \text{const.} = \log\left|\frac{v}{\sqrt{1 - v^2/c^2}}\right| + \text{const.},$$

so

$$p(v) = \frac{\pm e^{\text{const.}} v}{\sqrt{1 - v^2/c^2}}$$

The first nonzero term in the Taylor expansion of the right side at $v = 0$ is $\pm e^{\text{const.}} v$, so for consistency with classical mechanics the coefficient of v must be the mass of the particle. (Thus mass arises from a constant of integration.) This is why we set relativistic momentum to be $p(v) = mv/\sqrt{1 - v^2/c^2}$, which is the second formula in (B.1).

Feeding the formula for $p(v)$ into (B.6), we proceed to solve for $E(v)$:

$$\frac{mv}{\sqrt{1 - v^2/c^2}} = \frac{dE}{dv}\left(1 - \frac{v^2}{c^2}\right) \implies \frac{dE}{dv} = \frac{mv}{(1 - v^2/c^2)^{3/2}} = \frac{d}{dv}\left(\frac{mc^2}{\sqrt{1 - v^2/c^2}}\right).$$

Therefore

$$E(v) = \frac{mc^2}{\sqrt{1 - v^2/c^2}} + \text{const.}$$

Desiring the normalization condition $E(0) = 0$, we need the constant to be $-mc^2$ and we are led to the first formula in (B.1).

REFERENCES

[1] A. Einstein, Zur Elektrodynamik bewegte Körper (On the Electrodynamics of moving bodies), *Annalen der Physik* **17** (1905), 891–921. URL https://ia803002.us.archive.org/13/items/zurelektrodynami00aein/zurelektrodynami00aein.pdf (original), URL https://www.fourmilab.ch/etexts/einstein/specrel/www/ (translation).

[2] M. M. Schiffer and L. Bowden, "The Role of Mathematics in Science," Math. Assoc. America, Washington, DC, 1984.

[3] S. Sonego and M. Pin, Deriving relativistic momentum and energy, *Eur. J. Phys.* **26** (2005), 33–45. URL http://arxiv.org/pdf/physics/0402024.pdf.