RELATIVISTIC ADDITION AND GROUP THEORY

KEITH CONRAD

1. Introduction

For three particles P, Q, R travelling 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 any 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 velocities is:

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

Example 1.1. If $v_{PQ} = (3/4)c$ and $v_{QR} = (1/2)c$ then

$$v_{PQ} + v_{QR} = \frac{5}{4}c > c, \quad \frac{v_{PQ} + v_{QR}}{1 + (v_{PQ}v_{QR}/c^2)} = \frac{10}{11}c < c.$$

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

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

While $\oplus$ on $(-c, c)$ may seem complicated, it has properties similar to addition on $\mathbf{R}$:

- 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 any 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 any $v_1, v_2, v_3 \in (-c, c)$.

It is left to the reader to check these, of which the first and fourth are the only ones with much content. Note usual addition is *not* closed on $(-c, c)$.

The formula for $v \oplus w$ in (1.2) is reminiscent of 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. Recall the hyperbolic tangent function is given by the formula

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

It is a bijection from $\mathbf{R}$ to $(-1, 1)$, with inverse

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

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 $\varphi(x) = c\tanh(x)$ sends $\mathbf{R}$ to $(-c, c)$ and

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

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:

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

Thus, by a suitably clever transformation, essentially the inverse of the hyperbolic tangent, we replace velocities $v \in (-c, c)$ by rescaled velocities $\psi(v) \in \mathbf{R}$ and find this converts the nonintuitive operation $\oplus$ on $(-c, c)$ into *ordinary addition of real numbers*.

Example 1.2. 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 $\psi(v \oplus w) = \psi((10/11)c) = (1/2)\log 21 = \psi(v) + \psi(w)$.

Of course, the transformation from $\oplus$ to $+$ may seem like a fortuitous accident: could we ever have found the transformation ψ if we were not reminded of the hyperbolic tangent? Yes! We will prove that *every* (continuously differentiable) group law on an open interval of real numbers can be rescaled in an explicit manner to look like ordinary addition on $\mathbf{R}$.

2. Group laws on an interval

Let $I \subset \mathbf{R}$ be an open interval, with a group law $*$. That is, $*$ has the following four properties:

- Closure. $x, y \in I \Longrightarrow x * y \in I$.
- Identity. There is $u \in I$ such that for any $x \in I$, $x * u = u * x = x$.
- Inverses. For any $x \in I$ 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 any 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 any $x \in I$, $F(x, u) = F(u, x) = x$.
- For any $x \in I$ 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 $F(x, y) = x * y$ 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 $\mathbf{R}$.

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:

$$\ell(y) = \int_u^y \frac{\ell'(u)}{F_1(u,t)}dt. \tag{2.1}$$

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 which 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 any $t \in I$, $F_1(u,t) > 0$.*

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,

$$F_1(y,z)F_1(u,y) = F_1(u,F(y,z)) = F_1(u,y*z). \tag{2.2}$$

So if $F_1(u,y) = 0$ for some y, then $F_1(u,y*z) = 0$ for any 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. $\square$

Lemma 2.2 allows us to divide by $F_1(u,t)$ for any $t \in I$, 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 any x in the interval I:

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

By the Fundamental Theorem of Calculus, ℓ is differentiable and

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

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

(If you know about differential forms, the following comments may be of interest. Another way of stating (2.2) is in terms of the differential form $\omega = \mathrm{d}t/F_1(u,t)$. 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.)

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

Proof. We need to check two things:

- $\ell(F(x,y)) = \ell(x) + \ell(y)$.
- $\ell\colon I \to \mathbf{R}$ is a bijection.

For the first item, *fix* $y \in I$. We consider the x-derivatives of the two functions

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

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 and the functions are equal for all x. This verifies the first item: $\ell(F(x,y)) = \ell(x) + \ell(y)$.

For the second item, 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, one positive and the other negative. For any 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 $\ell(I) = \mathbf{R}$. □

Corollary 2.3. *When $x*y = F(x,y)$ in Theorem 2.1 has continuous partial derivatives, it is commutative. In particular, $F_1(x,y) = F_2(y,x)$ for all $x,y \in I$.*

Proof. Commutativity was never used in the proof of Theorem 2.1, so commutativity of addition on $\mathbf{R}$ implies commutativity of $*$ on I. Now differentiate both sides of the formula $F(x,y) = F(y,x)$ with respect to x. □

Corollary 2.4. *The function ℓ determines the operation $*$ by*

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

Proof. We know ℓ is a bijection, so it is invertible. Apply ℓ^{-1} to both sides of $\ell(x*y) = \ell(x) + \ell(y)$. □

Since ℓ is determined by the function $F_1(u,t)$ (see 2.3), and the operation $x * y$ is determined by ℓ (Corollary 2.4), we see the operation $*$ is encoded in the function $F_1(u,t) = F_2(t,u)$.

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:

$$x*y=F(x,y)\approx F(x,u)+F_2(x,u)(y-u)=x+F_1(u,x)(y-u). \tag{2.5}$$

Since $F_1(u,u)=1$, for x and y near u we get from (2.5) that $x*y\approx x+y-u$. Therefore

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

If we changed 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 any x and y near u has been used to reconstruct the operation $x*y$ for any x and any y!

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

$$\ell(x)=\int_0^x dt=x.$$

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

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

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 $.5625v_{QR}=(9/16)v_{QR}$.

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 but also v/c is also small, the coefficient of w is about 1, so $v \oplus w \approx v + w$. But 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 $(9/16)v_{QR}$ in the table above, where $v = (3/4)c$.

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 which converts multiplication on $(0,\infty)$ to addition on $\mathbf{R}$ is

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

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.