# RELATIVISTIC ADDITION AND REAL ADDITION

KEITH CONRAD

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:

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

For example, if $v_{PQ} = (3/4)c$ and $v_{QR} = (1/2)c$ then

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

There is an interesting algebraic similarity between the classical and relativistic velocity addition formulas. The classical model for velocity 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

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

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

- Closure, i.e. $v, w \in (-c, c) \Rightarrow v \oplus w \in (-c, c)$.
- 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: $(v \oplus w) \oplus x = v \oplus (w \oplus x)$ for $v, w, x \in (-c, c)$.

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 $v \oplus w$ 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 (2).

The *hyperbolic* tangent is better. It 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 (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 combines by ordinary 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 $\oplus$ on $(-c, c)$ into ordinary addition of real numbers. For example, 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 $\psi(v \oplus w) = \psi((10/11)c) = (1/2)\log 21 = \psi(v) + \psi(w)$.

It turns out that the transformation of $\oplus$ on $(-c, c)$ into $+$ on $\mathbf{R}$ is not so special, even though it is arising in special relativity. *Every* (continuously differentiable) "addition law" on an open interval of real numbers can be rescaled to ordinary addition on $\mathbf{R}$.

We work in the general setting of any open interval. Let $I \subset \mathbf{R}$ be an open interval, with an "addition 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 $x, y, z \in I$, $(x * y) * z = x * (y * z)$.

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 $x, y, z \in I$, $F(F(x, y), z) = F(x, F(y, z))$.

**Theorem 1.** *If $F(x, y) = x * y$ has continuous partial derivatives, then there is a differentiable rescaling function $f\colon I \to \mathbf{R}$ that converts $*$ on $I$ to ordinary addition on $\mathbf{R}$. That is, $f$ is a differentiable bijection with $f(x * y) = f(x) + f(y)$.*

Theorem 1 will be proved by giving an explicit recipe for $f$. To discover $f$, rewrite the desired equation $f(x * y) = f(x) + f(y)$ as

$$f(F(x, y)) = f(x) + f(y).$$

Let's differentiate this equation with respect to $x$:

$$f'(F(x, y))F_1(x, y) = f'(x),$$

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

$$f'(y)F_1(u, y) = f'(u),$$

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

$$f(y) = \int_u^y \frac{f'(u)}{F_1(u,t)} dt. \tag{3}$$

This is a formula for the rescaling function $f$. The constant $f'(u)$ is just an arbitrary factor, which we will set equal to 1. When we integrated, we didn't introduce an additive constant since we want $f(u) = 0$ and the integral formula already takes care of that. Of more pressing interest is the validity of dividing by $F_1(u,t)$. Why is it never zero?

**Lemma 1.** *For any $t \in I$, $F_1(u,t) > 0$.*

*Proof.* We differentiate the associative law with respect to $x$:

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

Setting $x = u$,

$$F_1(y,z)F_1(u,y) = F_1(u, F(y,z)) = F_1(u, y * z). \tag{4}$$

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 $y = u$, it must always be positive by the Intermediate Value Theorem. □

This lemma 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.

Note the use of $(d/dx)|_{x=u}$ in the discussions of (3) and (4).

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 $I$:

$$f(x) \stackrel{\text{def}}{=} \int_u^x \frac{dt}{F_1(u,t)}. \tag{5}$$

By the Fundamental Theorem of Calculus, $f$ is differentiable and

$$f'(x) = \frac{1}{F_1(u,x)}. \tag{6}$$

In particular, $f'(u) = 1$.

(High-powered remark: another way of stating (4) 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 $\tau_z^*$ on differential forms on $I$, and (4) says $\tau_z^*\omega = \omega$. So $f(x) = \int_u^x \omega$ is simply the integral of the "invariant" differential form $\omega$ along the path from the identity to $x$.)

With the choice of (5) for our rescaling function we now prove Theorem 1.

*Proof.* We need to check two things:

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

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

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

By (6), the derivative of the first function is

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

Does this equal the $x$-derivative of the second function, namely $f'(x) = 1/F_1(u,x)$? 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 (4) with $x, y, z$ relabelled as $u, x, y$. Therefore $f(F(x,y))$ and $f(x) + f(y)$ have equal $x$-derivatives for all $x$. Since they are equal at $x = u$, they are equal for all $x$. This verifies the first item.

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

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

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

**Corollary 1.** *When the operation $x * y = F(x,y)$ in Theorem 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 1, so commutativity of ordinary 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.** *The rescaling function $f$ determines the operation $*$ by*

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

*Proof.* Apply $f^{-1}$ to both sides of $f(x * y) = f(x) + f(y)$. □

Since $f$ is determined by the function $F_1(u,t)$ (Theorem 1), and the operation $x * y$ is determined by $f$ (Corollary 2), we see the operation $*$ is encoded in the function $F_1(u,t) = F_2(u,t)$.

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 $y$:

$$x * y = F(x,y) \approx F(x,u) + F_2(x,u)y = x + F_1(u,x)y. \tag{7}$$

Since $F_1(u,u) = 1$, for small $x$ and $y$ we get from (7) that $x * y \approx x + y$. But for small $y$ and not-so-small $x$ 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 small $y$ has been used to reconstruct the operation $x * y$ for any $x$ and any $y$!

Let's look at some examples, to see which functions rescale various "addition laws" on some interval to the real numbers with ordinary addition.

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

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

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

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

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 $f(x)$ in Theorem 1 is multiplied by a nonzero constant, it has the same relevant properties (except $f'(u) \neq 1$.)

The following table, where $v_{PR}$ is computed according to (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 $v_{PQ}/c$ is near 1. For $v_{QR}/c$ small, the fourth column entry is about $.5625 v_{QR} = (9/16) v_{QR}$.

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

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

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

If $v/c$ 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 Newton's $v + w$ by an error of around $(v/c)^2 w$. This explains the error $(9/16)v_{QR}$ in the table above, where $v = (3/4)c$.

As a third and final example, consider $F(x, y) = xy$ on $I = (0, \infty)$. This is the set of positive real numbers under multiplication. (The set of all nonzero real numbers is not an interval, and the reader should check to see where the proof of Theorem 1 breaks down in this case.) Here $u = 1$ and $F_1(1, x) = x$, so Theorem 1 tells us that the rescaling function which converts multiplication on $(0, \infty)$ to addition on $\mathbf{R}$ is

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

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