Elementary derivation of the expressions of momentum and energy in special relativity

Dedução elementar das expressões do momento e da energia na relatividade restrita

Luca Peliti*

Simons Center for Systems Biology, Institute for Advanced Study, Princeton, NJ, United States of America

The derivation of the expressions of momentum and energy of a particle in special relativity is often less than satisfactory in elementary texts. In some, it is obtained by resorting to quantum or electrodynamic considerations, in others by introducing non-elementary concepts, like that of a four-vector, or even misleading ones, like "relativistic mass". Nevertheless it is possible, following ideas described by Einstein in 1935, to obtain a fully elementary derivation of these expressions based only on the Lorentz transformations, on the conservation laws, and on the Newtonian limit. The resulting argument allows for a clearer and logically consistent introduction to the basic concepts of relativistic dynamics.
Keywords: Special relativity, Lorentz transformations, energy, momentum, conservation laws.

A dedução das expressões do momento e da energia de uma partícula na relatividade restrita é muitas vezes pouco satisfatória nos textos elementares. Em alguns textos, a dedução é feita com recursos a considerações quânticas ou eletrodinâmicos; em outros, através da introdução de conceitos pouco elementares, como quadrivetor, ou até mesmo de conceitos um tanto enganosos, como "massa relativística". No entanto, de acordo com ideias descritas por Einstein em 1935, é possível fazer uma dedução totalmente elementar dessas expressões, com base apenas nas transformações de Lorentz, nas leis de conservação e no limite newtoniano. Os argumentos utilizados permitem uma introdução clara e logicamente consistente aos conceitos básicos da dinâmica relativística.
Palavras-chave: Relatividade restrita; transformações de Lorentz, energia, momento, leis de conservação.

1. Introduction

Several texts provide an elementary derivation of the kinematics of special relativity, eventually based on the first part of Einstein's fundamental paper of 1905. [1] Starting from the two postulates of the total equivalence of inertial reference systems, and of the constancy of the speed of light in all such systems, it is in fact easy to obtain the expression of the Lorentz transformation of space-time coordinates. A particularly simple and appealing derivation is obtained by exploiting Bondi's so-called k-calculus, [2] itself based on the Doppler effect. Nevertheless going from kinematics (the Lorentz transformation) to dynamics, and in particular to the relativistic expressions

of momentum and energy, is often achieved by resorting to more sophisticated concepts, like that of a four-vector, or by the use of quantum considerations. (One example of this approach is the "elementary derivation" suggested by F. Rohrlich, [3] which exploits the expressions of the momentum and energy of a photon of frequency ν.) Einstein himself had originally derived the mass-energy equivalence via an explicit use of electrodynamics, [4] and not just by the kind of kinematic considerations which he had developed in the first part of his 1905 paper, which are based on the constancy of the speed of light, but do not otherwise depend on Maxwell's equations.

Einstein had remarked this problem, and he proposed in 1935 an elementary derivation of the mass-

*Endereço de correspondência: luca@peliti.org.

energy relation, independent of his 1905 argument, motivating it with the following words: [5]

> The special theory of relativity grew out of the Maxwell electromagnetic equations. So it came about that even in the derivation of the mechanical concepts and their relations the consideration of those of the electromagnetic field has played an essential role. The question as to the independence of those relations is a natural one because the Lorentz transformation, the real basis of the special relativity theory, in itself has nothing to do with the Maxwell theory and because we do not know the extent to which the energy concepts of the Maxwell theory can be maintained in the face of the data of molecular physics. In the following considerations, except for the Lorentz transformation, we will depend only on the assumption of the conservation principles for impulse and energy.

Einstein's considerations exploit a conceptual experiment introduced by G.N. Lewis and R.C. Tolman [6] and further discussed by P.S. Epstein, [7] where one considers collisions between pairs of particles in different inertial reference frames, and one looks for the expressions of momentum and energy by postulating their conservation. It is interesting to point out that Lewis and Tolman's paper, as well as Epstein's one, only consider elastic collision and derive the relativistic expression of momentum, while they provide a doubtful argument for the mass-energy equivalence by the consideration of the change of the "relativistic mass" with speed. "Relativistic mass" is in fact a rather problematic concept, [8,9] and there is a growing consensus to avoid its introduction in the teaching of relativity. Einstein derives instead the equivalence by simply extending the argument to inelastic collisions. The advantage of this approach for introducing the basic concepts of special relativity has been well remarked by R.F. Feynman who, in his *Lectures* (Ref. [10, Vol. I. Secs. 16–4, 16–5]), derives the relativistic expressions of momentum and energy in a way that closely resembles Einstein's one. (Feynman's derivation is however marred by his use of the "relativistic mass".) Einstein's argument has been more recently discussed by F. Flores, [11] who identifies three closely related but different claims within the mass-energy equivalence concept, and compares Einstein's 1935 argument with his original

1905 derivation [4] and with M. Friedman's 1983 derivation, [12, p. 142ff] which rests upon the consideration of Newton's equations in special relativity.

In this note, I present this line of thought in the hope that it may be found useful for the presentation of these fundamental concepts of special relativity in introductory courses for students of physics and mathematics. While the derivation of the relativistic expression of momentum in Sec. 3 is close to Epstein's and Feynman's arguments, the discussion of the expression of the kinetic energy and of the mass-energy equivalence is closer to Einstein's one.

2. Lorentz transformations and dynamic postulates

Following Einstein's 1905 paper (Ref. [1, § 2]), the kinematic concepts of special relativity rest on the following postulates:

(i) The laws that govern the transformations of the state of physical systems take the same form in reference frames animated by uniform translational motion one with respect to the other.

(ii) There is a class of such reference frames (the *inertial frames*) in which the speed of light assumes the same value c, independently of the state of motion of its source.

We shall choose from now on units in which $c = 1$ and limit our considerations to inertial frames. Based on these postulates one can easily derive Lorentz transformations in the following form. Let us consider two reference frames, K and K', such that K' is in uniform translational motion in the positive x direction and with speed V with respect to K. Then the event of coordinates (t', x', y', z') in K' has in K the coordinates (t, x, y, z), where

$$
\begin{aligned}
t &= \gamma(V)(t' + Vx'); \\
x &= \gamma(V)(x' + Vt'); \\
y &= y'; \\
z &= z',
\end{aligned}
\tag{1}
$$

and we have defined

$$
\gamma(V) = \frac{1}{(1 - V^2)^{1/2}}.
\tag{2}
$$

The same relation holds for the differentials dt, dx, etc. Dividing by dt we obtain the rules for the trans-

formation of velocities:

$$u_x = \frac{dx}{dt} = \frac{u'_x + V}{1 + u'_x V};$$
$$u_y = \frac{dy}{dt} = \frac{u'_y}{\gamma(V)(1 + u'_x V)}; \qquad (3)$$
$$u_z = \frac{dz}{dt} = \frac{u'_z}{\gamma(V)(1 + u'_x V)}.$$

To introduce dynamical concepts we obviously need supplementary postulates. We shall therefore postulate the following:

(iii) The momentum $\boldsymbol{P}$ and the energy E of a particle possessing the velocity $\boldsymbol{u}$ in the reference frame K have respectively the expressions

$$\boldsymbol{P} = m\boldsymbol{u}\,F(u); \quad E = E_0 + m\,G(u), \qquad (4)$$

where E_0 is a constant, that can be called the *rest energy*, m is a positive constant (which is *relativistically invariant*, i.e., does not change from one inertial frame to another, and thus *does not depend on the particle's velocity*) which we shall simply call its *mass*, and $F(u)$ and $G(u)$ are *monotonically increasing* universal functions of $u = |\boldsymbol{u}|$. The fact that the functions F and G depend only on the magnitude u of the velocity, but not on its direction, can be inferred by the isotropy of space.

(iv) For $u \ll 1$ these expressions reduce to the well-known classical ones. One has in particular

$$F(u) = 1 + \mathrm{O}(u^2); \quad G(u) = \frac{1}{2}u^2 + \mathrm{O}\left(u^4\right). \qquad (5)$$

(v) The total momentum $\boldsymbol{P}^{\mathrm{tot}}$ and the total energy E^{tot} of a system of several particles are respectively given by the sum of $\boldsymbol{P}$ and of E running over all the particles of the system.

(vi) **Conservation of momentum and energy:** Let us assume that (elastic or inelastic) collisions occur in a system of particles. Then, in each reference frame, $\boldsymbol{P}^{\mathrm{tot}}$ and E^{tot} maintain the same values before and after each collision.

As a corollary, the velocity of each particle remains constant as long as no collision occurs, because one can consider systems made up of single, independent particles.

3. Elastic collisions and relativistic momentum

Let us now consider a *particle pair*, i.e., a system made of two particles with equal values of m. Let us assume that in a reference frame K they have opposite velocities $\boldsymbol{u}_1$, $\boldsymbol{u}_2 = -\boldsymbol{u}_1$, where $\boldsymbol{u}_1 = (V, v, 0)$, with $|v| \ll V$, $V > 0$. Thus $|\boldsymbol{u}_{1,2}| \simeq V$. Let us moreover assume that the particles undergo an elastic collision, and take on respectively the velocities $\boldsymbol{w}_1 = (W, w, 0)$ and $\boldsymbol{w}_2 = (W', w', 0)$ after it. By the conservation of momentum one must have $W' = -W$ and $w' = -w$, *independently of the form of the function $F(u)$*. Indeed, one has $\boldsymbol{P}^{\mathrm{tot}}_{\mathrm{in}} = 0$ before the collision. By the conservation law one must have

$$\boldsymbol{P}^{\mathrm{tot}}_{\mathrm{out}} = 0 = m\boldsymbol{w}_1\,F(w_1) + m\boldsymbol{w}_2 F(w_2). \qquad (6)$$

Thus the vectors $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are antiparallel, and one has

$$\frac{|\boldsymbol{w_1}|}{|\boldsymbol{w_2}|} = \frac{F(w_2)}{F(w_1)}. \qquad (7)$$

Since the function $F(u)$ is monotonically increasing, this equation can only be satisfied if $|\boldsymbol{w}_1| = |\boldsymbol{w}_2|$, and we have therefore $\boldsymbol{w}_1 = -\boldsymbol{w}_2$. Now energy conservation imposes $w_i = u_i$. Indeed, the total kinetic energy before the collision is given by $2mG(u)$, and after it is given by $2mG(w)$. Since $G(u)$ is a monotonically increasing function of u, this condition can only be satisfied if $w = u$.

Let us now consider the special case where the velocity change is parallel to the y axis. (This rules out "head-on collisions", where the particles simply exchange their velocity.) Let the particles' velocities be given by $\boldsymbol{u}'_1 = (V', v', 0)$ and $\boldsymbol{u}'_2 = -\boldsymbol{u}'_1$ (cf. fig. 1) in the K' reference frame. Let us now look at the same collision in a reference frame K moving with a velocity V' in the x direction with respect to K'.

In this reference frame, the velocities $\boldsymbol{u}$ of the particles are respectively given by

$$\boldsymbol{u}_1 = (0, v, 0),$$
$$\boldsymbol{u}_2 = (-V, -w, 0), \qquad (8)$$

where

$$V = \frac{2V'}{1 + V'^2},$$
$$v = \frac{v'}{\gamma(V')(1 - V'^2)} \qquad (9)$$
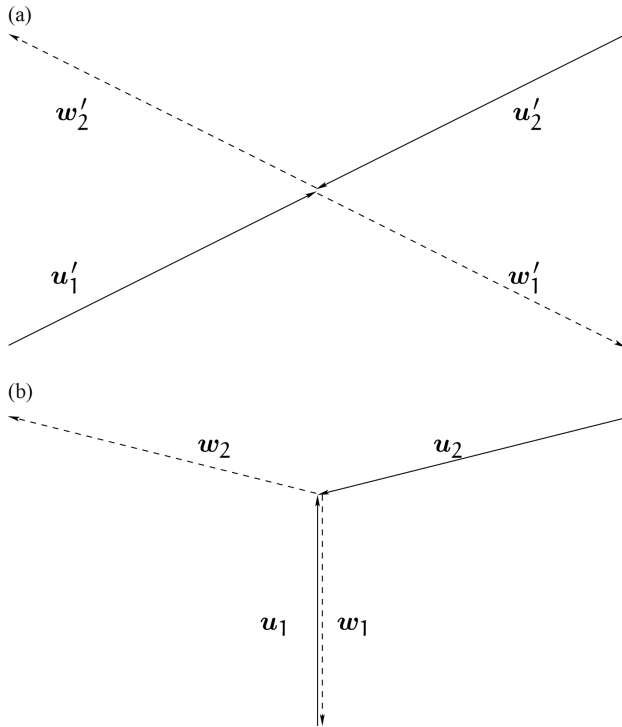$$w = \frac{v'}{\gamma(V')(1 + V'^2)}.$$

(a)

(b)

Figure 1: Above: Collision of a particle pair in the K' reference frame. Below: The same collision as seen in the K reference frame. Figures are schematic and no quantitative relations are implied.

One can easily check that

$$\gamma(V) = \frac{1}{\sqrt{1-V^2}} = \frac{1+V'^2}{1-V'^2}, \qquad (10)$$

and therefore that

$$v = w\,\gamma(V). \qquad (11)$$

One can also obtain this result by applying Eq. (3) to the transformation from the K' reference frame to a frame K'' animated, with respect to K', by a uniform translational motion with velocity $-V'$ parallel to the x axis. In this frame the x component of the velocity $\boldsymbol{u}_2''$ vanishes and that of the $\boldsymbol{u}_1''$ velocity is equal to V, and thus the speeds of the two particles are interchanged.

Let us now consider the conservation of momentum. The change $\delta\boldsymbol{P}_1$ of the momentum of particle 1 is given by

$$\delta\boldsymbol{P}_1 = -2mv\,F(v)\,\boldsymbol{e}_y, \qquad (12)$$

where $\boldsymbol{e}_y$ is the y axis versor. The corresponding quantity for particle 2 is given by

$$\delta\boldsymbol{P}_2 = 2mw\,F(u_2)\,\boldsymbol{e}_y, \qquad (13)$$

where $u_2 = \sqrt{V^2 + w^2}$. Let us momentarily assume $v, w \ll V$: then $F(v) \simeq 1$ and $u_2 \simeq V$. From momentum conservation we obtain $\delta\boldsymbol{P}_1 + \delta\boldsymbol{P}_2 = 0$, which implies

$$mv = mw\,F(V). \qquad (14)$$

Since $w = v/\gamma(V)$, we obtain

$$F(V) = \gamma(V) = \frac{1}{\sqrt{1-V^2}}. \qquad (15)$$

Having obtained this result for $v, w \ll V$, it is easy to see that it also holds for larger values of v and w, by substituting V with the speed of the corresponding particle. We have in fact

$$mv\,F(v) = mw\,F(u_2) = \frac{mv}{\gamma(V)}\,F(u_2), \qquad (16)$$

from which it follows

$$F(u_2) = F(v)\gamma(V), \qquad (17)$$

namely

$$\frac{1}{\sqrt{1-u_2{}^2}} = \frac{1}{\sqrt{1-v^2}}\frac{1}{\sqrt{1-V^2}}, \qquad (18)$$

an identity which is easy to check directly.

4. Kinetic energy conservation

Let us consider a particle having the velocity $\boldsymbol{u}' = (u', 0, 0)$, parallel to the x axis, in the K' reference frame. Its velocity $\boldsymbol{u}$ in the K frame, in uniform translational motion with respect to K' with speed V in the direction of the positive x-axis, is given by $\boldsymbol{u} = (u, 0, 0)$, with

$$u = \frac{u' + V}{1 + u'V}. \qquad (19)$$

One can easily see that

$$\gamma(u) = (1 + u'V)\,\gamma(u')\gamma(V). \qquad (20)$$

If $\boldsymbol{u}'$ is not parallel to the x axis, but one has instead $\boldsymbol{u}' = (u'_x, u'_y, u'_z)$, one has the more general relation

$$\gamma(|\boldsymbol{u}|) = (1 + u'_x V)\gamma(u')\gamma(V), \qquad (21)$$

which can be obtained with a little algebra. It is also easy to check that

$$\begin{aligned} u_x\gamma(u) &= (u'_x + V)\gamma(u')\gamma(V); \\ u_y\gamma(u) &= u'_y\gamma(u'); \\ u_z\gamma(u) &= u'_z\gamma(u'). \end{aligned} \qquad (22)$$

Let us now consider a particle pair, which have opposite velocities $\boldsymbol{u}_1'$, $\boldsymbol{u}_2' = -\boldsymbol{u}_1'$ in the K' frame. Thus $u' = |\boldsymbol{u}_1'| = |\boldsymbol{u}_2'|$ is the common value of the particles' speed in the K' frame. Let us denote by $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ the corresponding velocities in the K frame. We obtain

$$\gamma(u_1) + \gamma(u_2) = 2\gamma(u')\gamma(V). \qquad (23)$$

We have seen that an elastic collision in the K frame cannot change the common value u' of the particles' speed. Thus the right-hand side of this equation cannot change in the collision. But then neither can its left-hand side. If we denote by $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ the particles' speeds in the K frame after the collision, we obtain

$$\gamma(u_1) + \gamma(u_2) = \gamma(w_1) + \gamma(w_2). \qquad (24)$$

As Einstein (Ref. [5, p. 227]) points out, these equations have the form of conservation laws. We can thus interpret [13] $m(\gamma(u) - 1)$ as the kinetic energy T of a particle with mass m animated by a velocity $\boldsymbol{u}$. For small values of u this quantity is thus given by

$$T = m(\gamma(u) - 1) \simeq \frac{1}{2}mu^2, \qquad (25)$$

in agreement with the classical limit. We can thus set

$$G(u) = \gamma(u) - 1. \qquad (26)$$

Let us remark moreover that, by applying Eqs. (22) to a particle pair, we obtain

$$\boldsymbol{u}_1\gamma(u_1) + \boldsymbol{u}_2\gamma(u_2) = 2\boldsymbol{V}\gamma(u')\gamma(V). \qquad (27)$$

We can thus derive the following relation:

$$\boldsymbol{u}_1\gamma(u_1) + \boldsymbol{u}_2\gamma(u_2) = \boldsymbol{w}_1\gamma(w_1) + \boldsymbol{w}_2\gamma(w_2), \qquad (28)$$

which can be interpreted as the conservation law for the momentum. We thus recover the relativistic expression of the momentum derived in sec. 3.

5. Mass-energy equivalence

Let us now consider a totally inelastic collision in a particle pair. In the reference frame K' in which $\boldsymbol{P}^{\text{tot}} = 0$, the total kinetic energy before the collision is given by

$$T_{\text{in}}' = 2m(\gamma(u') - 1), \qquad (29)$$

For simplicity, we shall assume that the particles velocities $\boldsymbol{u}_{1,2}'$ are parallel to the y axis. The total kinetic energy after the collision vanishes, but the energy of the resulting particle has increased by T_{in}'. By momentum conservation the resulting particle is at rest in the reference frame K' and has thus velocity $\boldsymbol{V}$ parallel to the x axis in the K frame. Let us denote by M its mass. In the K frame the total momentum before the collision is given by

$$\boldsymbol{P} = m(\boldsymbol{u}_1\gamma(u_1) + \boldsymbol{u}_2\gamma(u_2)) = 2m\boldsymbol{V}\gamma(u')\gamma(V), \qquad (30)$$

while after the collision it has the value

$$\boldsymbol{P} = M\boldsymbol{V}\gamma(V). \qquad (31)$$

We obtain therefore

$$M = 2m\gamma(u') = 2m + T_{\text{in}}'. \qquad (32)$$

Therefore when two particles collide inelastically to form a new particle, the mass of the resulting particle is larger than the sum of the masses of the colliding particles by a quantity exactly equal, in our units, to the kinetic energy transformed by the collision into other energy forms. One can thus introduce a "natural" choice of the energy at rest E_0 of a particle, by equating it (in our units) to its mass, also because, "from the nature of the concept, [that] is determined only to within an additive constant, one can stipulate that E_0 should vanish together with m." [5, p. 229]. We can then identify $m\gamma(u)$ with the total energy of a particle, and associate the change δE of energy from the kinetic to a different form with a change $\delta m = \delta E$ of the mass of the particle. In the more general case, in which the two particles do not coalesce, but are animated by velocities with the same magnitude w' in the K' frame after the collision, one can then express the conservation of energy by

$$2m\gamma(u') = 2\bar{m}\gamma(w'), \qquad (33)$$

where $\bar{m}$ is the mass of each particle after the collision. The same applies in the K frame, if we take into consideration the relation

$$\gamma(u) = \gamma(u')\gamma(V), \qquad (34)$$

which holds since both $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are parallel to the y axis.

It is then a simple matter to verify that the energy E and the momentum $\boldsymbol{P}$ of a particle of mass m satisfy the relation

$$E^2 - P^2 = m^2, \qquad (35)$$

where the right-hand side is relativistically invariant. This relation also holds in the limit of a zero-mass particle, for which it assumes the form

$$E = |\boldsymbol{P}|. \tag{36}$$

As a corollary, *mass is not additive*: the mass of a system of particles depends on its total energy content in the vanishing-momentum frame. Thus when light waves (of vanishing mass) with opposite momenta are emitted from a particle at rest, the mass of the particle changes (as argued in Einstein's 1905 work, Ref. [4]). Since the right-hand side of Eq. (35) is invariant, it can be taken as the starting point to show that the energy E and the momentum $\boldsymbol{P}$ combine into a relativistic four-vector, a property which generalizes to particle systems.

One should point out that this line of thought does not rest on Maxwell's equations (keeping only the constancy of the speed of light) and neither does on other mechanical concepts, in particular on the concept of force, that is difficult to justify in special relativity. Einstein criticizes the use of the concept of force in the derivation of the relativistic expression of momentum, contained in the book by G.D. Birkhoff and R.E. Langer, *Relativity and Modern Physics*, [14] exactly for this reason, as made clear by the closing paragraphs of his 1935 paper: [5]

> Thus, in the book just mentioned, essential use is made of the concept of force, which in the relativity theory has no such direct significance as it has in classical mechanics. This is connected with the fact that, in the latter, the force is to be considered as a given function of the coordinates of all the particles, which is obviously not possible in the relativity theory. Therefore I have avoided introducing the force concept.
>
> Furthermore, I was concerned with avoiding making any assumption concerning the transformation character of impulse and energy with respect to a Lorentz transformation.

References

[1] A. Einstein, Annalen der Physik **17**, 891 (1905).
[2] H. Bondi, *Relativity and Common Sense* (Dover, New York, 1980).
[3] F. Rohrlich, Am. J. Phys. **58**, 348 (1990).
[4] A. Einstein, Annalen der Physik **18**, 639 (1905).
[5] A. Einstein, Bull. Am. Math. Soc. **41**, 223 (1935).
[6] G.N. Lewis and R.C. Tolman, Phil. Mag. **6**, 510 (1909).
[7] P.S. Epstein, Annalen der Physik **36**, 779 (1911).
[8] C.G. Adler, Am. J. Phys. **55**, 739 (1987).
[9] L.B. Okun, Physics Today **42**, 31 (1989); Am. J. Phys. **77**, 430 (2009).
[10] R. Feynman, R. Leighton and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Boston, 1966).
[11] F. Flores, Stud. Hist. Philos. Mod. Phys. **29**, 223 (1998).
[12] M. Friedman, *Foundations of Spacetime Theories* (Princeton University Press, Princeton, 1983).
[13] Equation (24) implies the conservation of $\sum(\gamma(u) + \text{const})$. The value of the constant is chosen so that the expression vanishes for $u \to 0$.
[14] G.D. Birkhoff and R.E. Langer, *Relativity and Modern Physics* (Harvard University Press, Cambridge, 1923).