Energy and Momentum at All Speeds

Randy S

Abstract Expressions for the energy and momentum of an object in terms of its mass and speed are usually introduced using a low-speed approximation, one that is valid only for (relative) speeds much less than the speed of light. This article introduces expressions for the energy and momentum of a single object that are valid at any speed, from zero up to the speed of light. This article also introduces Lorentz transformations.

Contents

1 Introduction

Flat spacetime is the arena for **special relativity**. Consider an isolated object in flat spacetime, not interacting with anything else. The object's energy E and momentum $\mathbf{p}$ both depend on its velocity $\mathbf{v}$. If the object is moving slowly enough, then we can use the familiar approximations

$$E \approx \text{constant} + \frac{m\mathbf{v}^2}{2} \qquad\qquad \mathbf{p} \approx m\mathbf{v} \tag{1}$$

where m is a $\mathbf{v}$-independent quantity that we call the object's mass. These approximations are valid for objects moving slowly compared to the speed of light. This article introduces better equations for E and $\mathbf{p}$ that are valid over the full range of speeds, from zero up to the speed of light. Article 41182 explains how the better equations can be derived from something deeper, but here they will merely be treated as axioms. Section 5 shows how they lead to the familiar approximations (1) when the object is moving slowly enough.

2 Notation

Momentum and velocity are vectors. In this article, a boldface symbol like $\mathbf{v}$ denotes a vector in three-dimensional space, which may be represented as a list of three numbers called the **components** of the vector, like this:

$$\mathbf{v} = (v_1,\, v_2,\, v_3).$$

In flat space with the usual coordinate system,[1] the **magnitude** of the vector $\mathbf{v}$ is

$$|\mathbf{v}| \equiv \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

The abbreviation

$$\mathbf{v}^2 \equiv |\mathbf{v}|^2$$

will also be used. Here is a summary of symbols that will be used in this article:

$$\begin{aligned} \mathbf{v} &= \text{velocity} & E &= \text{energy} \\ v &= \text{abbreviation for } |\mathbf{v}| & m &= \text{mass} \\ \mathbf{p} &= \text{momentum} & c &= \text{the ``speed of light''} \\ p &= \text{abbreviation for } |\mathbf{p}| \end{aligned}$$

The **speed** of an object is the magnitude $v = |\mathbf{v}|$ of its velocity vector $\mathbf{v}$.

Despite being called the "(vacuum) speed of light," c is a universal constant whose significance is more basic than light. In special relativity, using units in which $c = 1$ is a natural thing to do (article 37431). This article mostly uses units with $c = 1$, but some equations will also be shown with factors of c restored. Starting with an equation written in units with $c = 1$, we can restore factors of c by replacing

$$\mathbf{v} \to \mathbf{v}/c \qquad \mathbf{p} \to \mathbf{p}/c \qquad E \to E/c^2 \qquad m \to m. \tag{2}$$

[1] The "usual" coordinate system for flat space is the coordinate system x_1, x_2, x_3 in which the length ds of a line segment is given by $(ds)^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2$, as explained in article 21808.

3 Energy and momentum at all speeds

In special relativity, the energy and momentum of an isolated object satisfy these conditions:

- The combination $E^2 - \mathbf{p}^2$ cannot be negative. Thanks to this condition, we can define a quantity m by

$$\boxed{m^2 \equiv E^2 - \mathbf{p}^2.} \tag{3}$$

- The combination $E^2 - \mathbf{p}^2$ does not depend on the object's velocity. This implies that m is independent of the object's velocity, so it defines an intrinsic property of the object. We call it the object's **mass**.[2] The mass m may be different for different objects, but the mass of any given object is the same no matter how fast the object is moving.

- The ratio $\mathbf{p}/E$ is equal to the object's velocity:

$$\boxed{\mathbf{v} = \frac{\mathbf{p}}{E},} \tag{4}$$

  where the energy E is understood to be positive:

$$E > 0. \tag{5}$$

The following sections highlight some consequences of these conditions.

[2] Some people use a different language, in which the velocity-independent quantity m is called "rest mass" and the energy E is called "relativistic mass." Language is never perfect, but people with more experience tend to reserve the word "mass" for the velocity-independent quantity defined by (3).

4 Energy and momentum as functions of velocity

This section shows how E and $\mathbf{p}$ may both be expressed in terms of only m and $\mathbf{v}$ when $m > 0$. Re-arrange equation (4) to get

$$\mathbf{p} = E\mathbf{v}. \tag{6}$$

Take the magnitude of both sides and use the resulting expression for p on the right-hand side of equation (3) to get

$$m^2 = E^2 - v^2 E^2.$$

Notice that the condition $m^2 > 0$ implies $v \leq 1$. Solve this for E and use the inequality (5) to get

$$\boxed{E = \frac{m}{\sqrt{1 - v^2}}.} \tag{7}$$

Subsitute the expression (7) for E into (6)

$$\boxed{\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - v^2}}.} \tag{8}$$

As the speed approaches 1, the denominator in these equations approaches zero. This says that when $m > 0$, E and $\mathbf{p}$ both grow without bound as $v \to 1$.

5 The low-speed approximation

The low-speed equations (1) can be derived from the all-speed equations (7) and (8), by using the approximation

$$\frac{1}{\sqrt{1-v^2}} = 1 + \frac{v^2}{2} + O(v^4), \tag{9}$$

which will be derived below. The notation $O(v^4)$ means that we are neglecting terms of order v^n for all $n \geq 4$.[3] To derive the approximation (9), square both sides and then multiply both sides by $1 - v^2$ to get the obviously-true equation $1 = 1 + O(v^4)$. We got this obviously-true from equation (9) using reversible operations,[4] so equation (9) is also true.

Use (9) in equations (7) and (8) to get

$$E = m + \frac{mv^2}{2} + O(v^4) \tag{10}$$

$$\mathbf{p} = m\mathbf{v} + O(v^3), \tag{11}$$

which agree with equations (1). These are good approximations when $v \ll 1$, because then $v^2 \ll v$ and $v^3 \ll v^2$ and so on. In the extreme case $\mathbf{v} = \mathbf{0}$, equation (10) reduces to

$$E = m \qquad \text{if } \mathbf{v} = \mathbf{0}. \tag{12}$$

We might worry that this isn't valid, because we can't set $\mathbf{v} = \mathbf{0}$ in equations (7) or (8). Equation (12) is still valid, though, because we can derive it directly by setting $\mathbf{v} = \mathbf{0}$ in (4), which gives $\mathbf{p} = \mathbf{0}$, and then using this in (3) to get (12). After restoring factors of c as explained in section 2, equation (12) becomes the famous equation $E = mc^2$. This tells us how much energy is contained in an object with mass m when the object is not moving.

[3] If v is very small, then v^2 is very very small, and v^3 is very very very small, and so on.

[4] Taking the square is reversible with the understanding that we should use the positive square root. Multiplying by $1 - v^2$ is reversible with the understanding that $v \neq 1$.

6 The speed limit

An object can have arbitrarily large momentum, but its speed cannot be arbitrarily large. To see this directly from equations (3)-(4), take the square root of equation (3) and use (5) to get

$$E = \sqrt{m^2 + \mathbf{p}^2}.$$

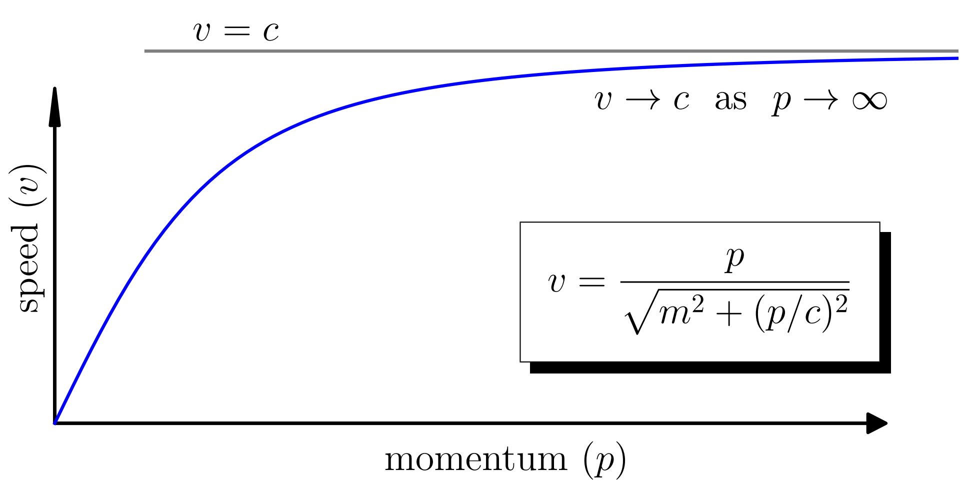
Substitute this into equation (4) to get

$$\mathbf{v} = \frac{\mathbf{p}}{\sqrt{m^2 + \mathbf{p}^2}}. \tag{13}$$

This implies $v \leq 1$, so the speed of an object never exceeds 1, no matter how large its momentum happens to be. In particular:

- If the object is massless ($m = 0$), then equation (13) says that its speed is automatically equal to the limiting value $v = 1$.

- If the object is massive ($m > 0$), then equation (13) says that the speed approaches its limiting value $v \to 1$ as $p \to \infty$. The harder you throw an object ($p \to \infty$), the closer it comes to moving at the limiting speed ($v \to 1$). This is illustrated below, with factors of c restored:



7 Can anything travel faster than light?

You may have heard that "nothing can travel faster than light." Here are some clarifications:

- The limiting speed c is usually called the "speed of light," but light travels at this speed only in empty space. In a medium like air or water, light moves slower than c, so other entities can outrun light in such a medium. The limiting speed c is still the same as it is in a vacuum, though: nothing can move faster than c, regardless of the ambient medium.

- A better name for the speed limit c is the "speed of information." To see why, imagine pointing a laser pointer at the left side of the moon and then quickly re-pointing it at the right side of the moon. The illuminated spot can move across the moon faster than c, but the illuminated spot isn't carrying any information across the moon, just like using two laser pointers to illuminate spots on both sides of the moon at the same time doesn't carry any information across the moon. It does carry information from the pointer to the moon, and that speed is limited to c.

- The speed limit c applies to the relative speeds at which *co-located* entities can pass by each other. It does not apply to the relative speeds of entities that are far away from each other, like two galaxies in different parts of the universe, whatever "relative speed" even means in that case. The *local* nature of the speed limit c, and the ambiguity of the naive idea of "relative speed" between distant objects, is easier to appreciate after studying the geometry of not-necessarily-flat spacetime. Article 48968 takes a first step in that direction.

8 Lorentz symmetry

A **Lorentz transformation** is any linear transformation of the four quantities (E, p_1, p_2, p_3) that leaves the right-hand side of equation (3) invariant. The mass m is invariant under Lorentz transformations.

One example of a Lorentz transformation is

$$\begin{pmatrix} E' \\ p_1' \\ p_2' \\ p_3' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & C & S & 0 \\ 0 & -S & C & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} E \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} \tag{14}$$

where C and S are any two real numbers satisfying

$$C^2 + S^2 = 1. \tag{15}$$

This qualifies as a Lorentz transformation because it satisfies $(E')^2 - (\mathbf{p}')^2 = E^2 - \mathbf{p}^2$. We recognize (14) as an ordinary rotation in the 1-2 plane, with $C = \cos\theta$ and $S = \sin\theta$ for some angle θ. An ordinary rotation changes the direction of the object's velocity $\mathbf{v} = \mathbf{p}/E$ (because it changes the direction of $\mathbf{p}$), but it does not change the object's speed $v = p/E$ (because it doesn't change p or E).

Another example of a Lorentz transformation is

$$\begin{pmatrix} E' \\ p_1' \\ p_2' \\ p_3' \end{pmatrix} = \begin{pmatrix} C & S & 0 & 0 \\ S & C & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} E \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} \tag{16}$$

where now C and S are any two real numbers satisfying

$$C^2 - S^2 = 1. \tag{17}$$

The quantities C and S can be written as $C = \cosh\theta$ and $S = \sinh\theta$ (section 10). This qualifies as a Lorentz transformation because it satisfies $(E')^2 - (\mathbf{p}')^2 = E^2 - \mathbf{p}^2$. This type of Lorentz transformation is called a **Lorentz boost**. A Lorentz boost changes the object's speed.

9 Lorentz boosts and speed, part 1

Equation (4) tells us how Lorentz boosts are related to changes in the object's velocity $\mathbf{v}$. If we start with $\mathbf{v} = \mathbf{0}$ ($\mathbf{p} = \mathbf{0}$), then a Lorentz boost along the 1-direction gives

$$\begin{pmatrix} E' \\ p'_1 \\ p'_2 \\ p'_3 \end{pmatrix} = \begin{pmatrix} C & S & 0 & 0 \\ S & C & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} E \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} CE \\ SE \\ 0 \\ 0 \end{pmatrix}$$

with $C^2 - S^2 = 1$. According to equation (4), the new velocity has magnitude

$$v' = \frac{p'}{E'} = \frac{S}{C}. \tag{18}$$

This is consistent with $v \leq 1$, because equation (17) implies that the magnitude of S/C cannot exceed 1.

10 The functions sinh and cosh

We know from trigonometry that real numbers C, S satisfying $C^2 + S^2 = 1$ can be written $C = \cos\theta$ and $S = \sin\theta$, for some real number θ. The sine and cosine functions can be defined by the conditions

$$e^{i\theta} = \cos\theta + i\sin\theta \qquad\qquad \begin{aligned} \cos(-\theta) &= \cos\theta \\ \sin(-\theta) &= -\sin\theta, \end{aligned}$$

where i is the imaginary unit ($i^2 = -1$). In words: $\cos\theta$ and $\sin\theta$ are the even and odd parts of $e^{i\theta}$, normalized to be real-valued.

Similarly, real numbers C, S satisfying $C^2 - S^2 = 1$ can be written $C = \cosh\theta$ and $S = \sinh\theta$, for some real number θ. These functions are defined by

$$e^{\theta} = \cosh\theta + \sinh\theta \qquad\qquad \begin{aligned} \cosh(-\theta) &= \cosh\theta \\ \sinh(-\theta) &= -\sinh\theta. \end{aligned}$$

In words: the **hyperbolic cosine** function $\cosh\theta$ and **hyperbolic sine** function $\sinh\theta$ are the even and odd parts of e^{θ}, respectively. Explicitly:

$$\cosh\theta = \frac{e^{\theta} + e^{-\theta}}{2} \qquad\qquad \sinh\theta = \frac{e^{\theta} - e^{-\theta}}{2}.$$

The trigonometric functions and their hyperbolic relatives satisfy similar-looking relations, but with different patterns of minus-signs. Examples:

$$(\cos\theta)^2 + (\sin\theta)^2 = 1 \qquad\qquad (\cosh\theta)^2 - (\sinh\theta)^2 = 1$$

and

$$\begin{aligned} \frac{d}{d\theta}\sin\theta &= \cos\theta \\ \frac{d}{d\theta}\cos\theta &= -\sin\theta \end{aligned} \qquad\qquad \begin{aligned} \frac{d}{d\theta}\sinh\theta &= \cosh\theta \\ \frac{d}{d\theta}\cosh\theta &= \sinh\theta. \end{aligned}$$

The next section highlights another example.

11 Combining boosts in the same direction

In the context of the functions $\cos\theta$ and $\sin\theta$, the quantity θ is called an angle. To rotations in the same plane[5] can be combined by adding their angles, thanks to the identity

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\theta' & \sin\theta' \\ -\sin\theta' & \cos\theta' \end{pmatrix} = \begin{pmatrix} \cos(\theta+\theta') & \sin(\theta+\theta') \\ -\sin(\theta+\theta') & \cos(\theta+\theta') \end{pmatrix}.$$

Similarly, in the context of the functions $\cosh\theta$ and $\sinh\theta$, the quantity θ is called a **rapidity** – a name motivated by its relationship to speed in equation (18). Two boosts in the same direction (better: in the same time-space plane) may be combined by adding their rapidities, thanks to the identity

$$\begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}\begin{pmatrix} \cosh\theta' & \sinh\theta' \\ \sinh\theta' & \cosh\theta' \end{pmatrix} = \begin{pmatrix} \cosh(\theta+\theta') & \sinh(\theta+\theta') \\ \sinh(\theta+\theta') & \cosh(\theta+\theta') \end{pmatrix}. \tag{19}$$

This type of linear transformation is called a **hyperbolic rotation**. A Lorentz boost is a hyperbolic rotation in a time-space plane, analogous to an ordinary rotation in a space-space plane.

[5] Traditional language says "about the same axis," but that only makes sense in three-dimensional space. Saying "in the same plane" makes sense in any number of dimensions (articles 33629, 12707, and 81674).

12 Lorentz boosts and speed, part 2

An ordinary rotation in a space-space plane changes the orientation from which an object is viewed. In contrast, section 9 showed that a hyperbolic rotation in a time-space plane changes the *velocity* from which an object is viewed.

To appreciate one implication of the identity (19), suppose we have 3 objects A, B, and C. From A's point of view, B is moving in the $+x$ direction with speed $\tanh\theta$. From B's point of view, C is moving in the $+x$ direction with speed $\tanh\theta'$. The identity (19) says that, from A's point of view, object C is moving in the $+x$ direction with speed $\tanh(\theta+\theta')$. This is summarized below, where the table on the left is A's point of view, and the table on the right is B's point of view:

direction	velocity
A	0
$B\rightarrow$	$\tanh\theta$
$C\rightarrow$	$\tanh(\theta+\theta')$

direction	velocity
$\leftarrow A$	$-\tanh\theta$
B	0
$C\rightarrow$	$\tanh\theta'$

The important message is that we don't add the speeds. We add the "angles" (rapidities) instead. These are not the same thing, because

$$\tanh(\theta+\theta')\neq\tanh\theta+\tanh\theta'. \tag{20}$$

Adding speeds works well enough for most everyday purposes, because

$$\tanh\theta\approx\theta \qquad \text{if } |\theta|\ll 1,$$

so the inequality (20) becomes an *approximate* equality if θ and θ' are both small. Example: if you throw a tomato forward at 70 km/hour while driving forward at 65 km/hour, the billboard will see the tomato approaching at approximately 135 km/hour. This approximation only works for low speeds, though. Instead of a tomato hitting a billboard, consider light from a flashlight hitting your eyes. In a vacuum, light travels with speed $1=\tanh\infty$, so its rapidity is infinite. Since $\infty+\theta=\infty$, the speed with which the light hits your eyes is always $\tanh\infty=1$ regardless of how fast the flashlight itself is rushing towards you (or away from you). The limiting speed is the same from all points of view.

13 References in this series

Article **12707** (https://cphysics.org/article/12707):
"Rotational Motion in Higher-Dimensional Space" (version 2022-02-05)

Article **21808** (https://cphysics.org/article/21808):
"Flat Space and Curved Space" (version 2022-01-16)

Article **33629** (https://cphysics.org/article/33629):
"Conservation Laws and a Preview of the Action Principle" (version 2022-02-05)

Article **37431** (https://cphysics.org/article/37431):
"How to Think About Units" (version 2022-02-05)

Article **41182** (https://cphysics.org/article/41182):
"Energy and Momentum at All Speeds: Derivation" (version 2022-02-18)

Article **48968** (https://cphysics.org/article/48968):
"The Geometry of Spacetime" (version 2022-01-16)

Article **81674** (https://cphysics.org/article/81674):
"Can the Cross Product be Generalized to Higher-Dimensional Space?" (version 2022-02-06)