the general science journal

The Dubious Origins of $E=mc^2$

Rothwell Bronrowan
physbron@t-online.de

This paper gives the derivation of $E=mc^2$. It also raises serious doubts as to the validity of the equation.

***

This is not actually a paper I ever intended writing!

All I wanted initially was to check again quickly on why c was involved in Einstein's famous formula in the first place. But this wasn't easy!

I found a number of articles in the Web beginning with statements like "... a full explanation would be too complex." Not what I wanted.

Then there were those that claimed to give something like the easiest, fastest or simplest explanation, but usually also began with something like, "Imagine an object travelling at almost the speed of light ..." Not what I wanted either!

So eventually I dug out an old Encyclopaedia Britannica - and found pretty much what I was looking for, fairly well written, on less than 3 pages.

***

I'll now give a *very short* summary, in my own words, of the essential contents of the Britannica article, until we arrive at $E=mc^2$. This will be followed by a more detailed analysis. Anyone therefore wishing to skip this segment can readily do so.

Treating the historical aspects of relativity, the Britannica article begins with Newton and moves on to Ørsted, Faraday, Maxwell and Hertz, before arriving at Michelson and Morley and a brief outline of their experiment.

FitzGerald and Lorentz are next brought into the picture and their independently-found "solution" to the unexpected zero result of the Michelson-Morley experiment, namely *length contraction*, is introduced. The amount of foreshortening is given as being equivalent to the factor $(1-v^2/c^2)^{1/2}$, which is the same as $\sqrt{(1-v^2/c^2)}$.[1]

This factor is then also said to be responsible for clock-slowing (time dilation).

[1] The factor is actually the inverse of that initially quoted in Encyclopaedia Britannica.

Einstein is next introduced and the findings of FitzGerald and Lorentz are given a new interpretation in the context of observers with different inertial frames of inference, the mathematical equations relating the *space* and *time* measurements of one observer to those of the other being indicated in terms of the *Lorentz transformations*:

$$x' = (1-v^2/c^2)^{-1/2}\,(x-vt),\ y' = y,\ z' = z,$$
$$t' = (1-v^2/c^2)^{-1/2}\,(t-vx/c^2).$$

Two associated diagrams are presented, these supposedly showing the differences between the non-relativistic (Fig. 1) and relativistic (Fig. 2) transformations, for relative motion of magnitude v along the x-axis:[2]

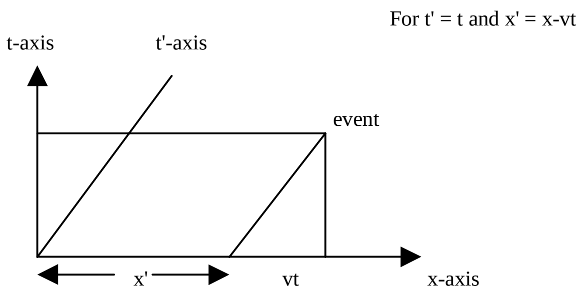

Fig. 1

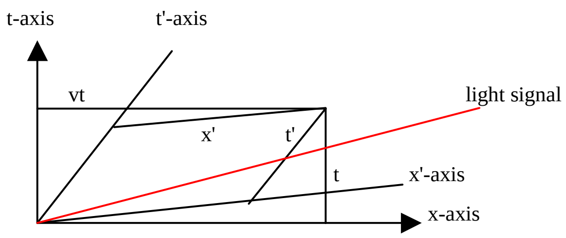

Fig. 2

It is then claimed that, as the relative speed of one frame of reference is increased with respect to the other, its clocks appear slower and its (measuring) rods appear shorter, both of these effects increasing indefinitely as the speed difference between them increases towards c.

The Lorentz transformations are therefore said to take appropriate account of changes in scale-length, in clock-time and in simultaneity.

The article next turns its attention to *relativistic mass*.

[2] The diagrams are given as "From P. Bergmann, *The Riddle of Gravitation*; Charles Scribner's Sons"

The author of the Britannica article argues as follows (and now I will quote his or her words, in order to be sure that no error on my part colours the picture)[3]:

(01) *Variable mass.*

(02) The mass of a material body is a measure of its resistance to a change in its state of motion caused by a given force.

(03) The larger the mass, the smaller the acceleration.

(04) If a material body is already moving at a speed approaching the speed of light, it must offer increasing resistance to any further acceleration so as not to cross the threshold of c.

(05) Hence the special theory of relativity leads to the conclusion that the mass of a moving body m is related to the mass that it would have if at rest, m_o, divided by the square root of one minus the fraction v^2/c^2:

(06) $m = (1-v^2/c^2)^{-1/2} m_o$

(07) The changing value of the mass of the moving body, m, is called the relativistic mass.

(08) As v approaches c, the figure within the parentheses approaches zero and, ultimately, $m = m_o/0$, which would be an infinitely large number.

(09) The relativistic mass number may be interpreted as indicating that the relativistic mass of a body exceeds its rest mass m_o by an amount that equals its kinetic energy E, divided by c^2: $m - m_o = E/c^2$.

(10) Hence the hypothesis that generally the energy is c^2 times the mass, or $E = mc^2$, and that energy and mass are, in fact, equivalent physical concepts, differing only by the choice of their units.

(The sentence-numbering is mine.)

***

Having now traced the steps from the Michelson-Morley experiment to deduction of the formula $E = mc^2$, it is time to take a closer look at this content.

We can begin by dividing the previous segment into three sections; the first on the Lorentz factor, the second on the Lorentz transformations, and the third on relativistic mass.

***

So how does the Michelson-Morley experiment relate to the Lorentz factor?

The background to this experiment is that it was conducted to verify a mathematical calculation made by Michelson, this calculation for its part being seen as a proof of the existence of the *ether*.

In the late nineteenth century it was believed that light travelled through this ether in all directions at a constant speed. Knowing that the earth was also in motion, Michelson employed a light-interference technique to detect a difference in the net

[3] Encyclopaedia Britannica, Macropaedia 15, Relativity, P. 583, © 1981.

speed of light in one direction (the direction of the earth's motion), as compared to another (a direction perpendicular to this). A beam of light was split into two pencils and sent out in both directions, one pencil being expected to return later than the other.

Instead of the anticipated result, however, the experiment indicated no clear findings at all. This "zero result", as it is often called, was accepted by Michelson.

By contrast, Hendrik Antoon Lorentz (1853-1928), a Dutch mathematician and physicist and a believer in the ether, was not prepared to accept the interpretation of the result and proposed an alternative solution, namely *length contraction*.

This "proposal" is so important because, as already indicated in connection with the Britannica article, the same *factor* on which it was based was subsequently also used - again primarily by Lorentz - to quantify *time dilation*, *frames of reference* and, yes (and this time by Einstein), even *mass increase*.

At this point, for a closer treatment, I refer the reader to my last article[4], in which I show that the Lorentz factor, as used by Lorentz (and Einstein), is merely one special case of a more general type of factor.

Here, in short, this can be explained as follows. Every right-angled triangle contains a "Lorentz factor" of its own. We can show this in terms of the following illustration:

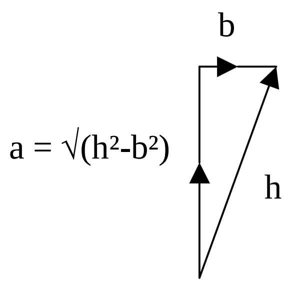

Fig. 3
(h = hypothenus, a & b are the shorter sides)

Here we have a generalized right-angled triangle with sides a, b and h. By Pythagoras we also know that $a = \sqrt{(h^2-b^2)}$, as is shown in the diagram. If we now divide h by a, we get:

	$h / \sqrt{(h^2-b^2)}$	
=	$h / \sqrt{(h^2(1-b^2/h^2))}$	extracting h^2 from the terms in brackets
=	$h / h\sqrt{(1-b^2/h^2)}$	extracting h, as the root of h^2, from the remaining denominator
=	$1 / \sqrt{(1-b^2/h^2)}$	cancelling h, above and below

4 *The Lorentz Triangle*, Rothwell Bronrowan, General Science Journal, 26.07.2010.

This is a generalized result because one can put any values into the triangle and, provided its angles remain unchanged, the factor derived remains *exactly* the same.

So now let us substitute-in the terms with which Lorentz was concerned, in the context of the Michelson-Morley experiment (namely velocities affecting the outbound journey in the perpendicular direction):

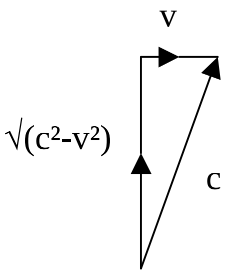

Fig. 4
Outbound journey in the perpendicular direction

Here, v stands for the velocity of the earth and c is the speed of light. Repeating our calculations, as above, we get:

$$c / \sqrt{(c^2-v^2)}$$
$$= c / \sqrt{(c^2(1-v^2/c^2))}$$ extracting c^2 from the terms in brackets
$$= c / c\sqrt{(1-v^2/c^2)}$$ extracting c, as the root of c^2, from the remaining denominator
$$= 1 / \sqrt{(1-v^2/c^2)}$$ cancelling c, above and below

Notice that we have just derived the Lorentz factor!

Now let us substitute-in terms from the experiment for distance travelled, rather than for velocity:

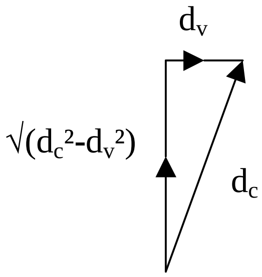

Fig. 5
Outbound distances in the perpendicular direction
(d_c = the distance travelled by the light [in the given time],
d_v = the distance travelled by the earth [in the given time], and
$\sqrt{(d_c^2-d_v^2)}$ = the perpendicular equivalent)

And again, repeating our calculations, as above, we get:

$$d_c / \sqrt{(d_c^2-d_v^2)}$$
$$= d_c / \sqrt{(d_c^2(1-d_v^2/d_c^2))}$$ extracting d_c^2 from the terms in brackets

=	$d_c / d_c \sqrt{(1-d_v^2/d_c^2)}$	extracting d_c, as the root of d_c^2, from the remaining denominator
=	$1 / \sqrt{(1-d_v^2/d_c^2)}$	cancelling d_c, above and below

The next important thing to notice is that each of these *factors* takes no terms, i.e. they are simple ratios. If they are still "expressed in terms", this is only because we need variables to substitute for the numerals not yet specified. So the following equation is no contradiction, for the triangle given:

$$1 / \sqrt{(1-b^2/h^2)} = 1 / \sqrt{(1-v^2/c^2)} = 1 / \sqrt{(1-d_v^2/d_c^2)}$$

We will return to this point again shortly.

Notice first that, in order to derive his *length contraction*, Lorentz *divides* the distance travelled in the direction of the earth's motion (d_v) by his Lorentz factor, $1 / \sqrt{(1-v^2/c^2)}$. Doing this division for Fig. 5, we have:

	$d_v / 1 / \sqrt{(1-v^2/c^2)}$	
=	$d_v / 1 / \sqrt{(1-d_v^2/d_c^2)}$	substituting, since $1 / \sqrt{(1-v^2/c^2)} = 1 / \sqrt{(1-d_v^2/d_c^2)}$, see above
=	$d_v \sqrt{(1-d_v^2/d_c^2)}$	since dividing by 1/x = multiplying by x
=	$d_v\, d_c \sqrt{(1-d_v^2/d_c^2)} / d_c$	multiplying by d_c, both above and below
=	$d_v \sqrt{(d_c^2(1-d_v^2/d_c^2))} / d_c$	taking d_c^2, as the square of d_c, inside the square-root term
=	$d_v \sqrt{(d_c^2-d_v^2)} / d_c$	multiplying the terms in brackets by d_c^2

Now notice, in Fig. 5, that this is just the *length*, as represented by d_v, multiplied by the ratio of the perpendicular side to the hypothenus.[5]

But what on earth is the mathematical - or even physical - justification for obtaining a length contraction in that one divides the given length by a factor that is essentially always present in every triangle?

I can find no good answer to this question.

And if there is none, then the length in question remains unaltered and *length contraction* is non-existent.[6]

Returning now to the above equation

$$1 / \sqrt{(1-b^2/h^2)} = 1 / \sqrt{(1-v^2/c^2)} = 1 / \sqrt{(1-d_v^2/d_c^2)}$$

it is obvious that, since the terms apparently included are interchangeable, they are in fact no terms at all!

---

[5] i.e. b / (h/a) = b (a/h).

[6] Another point that might be made here is that the *only* reason that Lorentz initially introduced *length contraction* was as a means of saving his concept of the *ether* from widespread rejection, following the result of the Michelson-Morley experiment. Less than 15 years later, however, Einstein came along and simply dismissed the ether concept as "unnecessary" - though retaining *length contraction* and the Lorentz factor for his own purposes!

This can be readily demonstrated by substituting "real values" taken from the Michelson-Morley experiment into the Lorentz factor:

$$1 / \sqrt{(1-v^2/c^2)}$$

$$= 1 / \sqrt{(1-(30 \text{ km/h})^2/(3 \times 10^8 \text{ km/h})^2)}$$ for v = velocity of the earth and c = speed of light

$$= 1 / \sqrt{(1-(30)^2/(3 \times 10^5)^2)}$$ cancelling the term "km/h", above and below

in which case we have no remaining *terms*, but simply a number.

Even if this were not the case, though, the justification for dividing the length of one side by *the ratio of the hypothenus to the other side* is still missing.

But now let's now take look at things from a new angle.

As referred to in the Britannica article, the Lorentz factor is applied not only in *length contraction*, but also in *time dilation*. Whereas the length is said to contract, the time is said to slow, or its rate is said to decrease. In both cases, moreover, this is said to apply simultaneously and to the same extent, i.e. by the value of the Lorentz factor.

Knowing therefore that

velocity = distance travelled / time taken

i.e.

$$v = d_v / t$$

we can now illustrate this by the equation

$$v = (d_v/L_F) / (t/L_F)$$

where L_F is the Lorentz factor. This, however, is equivalent to

$$v = d_v L_F / t L_F$$

which, after cancelling L_F above and below, brings us back to

$$v = d_v / t$$

In other words, since the Lorentz factor is said to effect both *length contraction* and *time dilation* simultaneously and to the same extent, the end result is that these effects cancel and the velocity remains unchanged![7]

[7] An interesting point here is that this *could* be said to present a new solution to the result of the Michelson-Morley experiment, namely *length extension!* To understand this, remember that length contraction applies in the direction of motion. This can hardly be the case for time dilation, however. If in the Michelson-Morley experiment, therefore, the length in the perpendicular direction remains unchanged whereas the time dilates (or its rate decreases), this in turn means that the velocity (d/t) effectively increases. Since the velocity in question is that of light itself, though, it can't increase! So what is left is an effective elongation - or extension - in the length, making the *perpendicular* journey longer.

And since I am not convinced of winning much support for the above claim that the Lorentz factor is just a number and can't be used as a basis for argumentation over "terms", let's now take an if-you-can't-beat-'em-joint-'em approach with the following argument:

> Knowing from the Lorentz factor, $1/\sqrt{(1-v^2/c^2)}$, that as the value of v approaches that of c, ... [any old argument here] ..., we can now say, having shown that v does not approach c, that this argument does not hold.

In fact, if one really wanted to be a stickler about this, one could even make the point that, in a relativistic universe, no change in velocity (i.e. acceleration or deceleration) is possible.

***

In view of the conclusions of the last segment, all I wish to say additionally on the *Lorentz transformations* is that the diagrams shown in Fig. 1 and Fig. 2, taken from the Britannica article, are given by me in the hope that the reader may understand them. I don't.

***

Which brings us at last to *relativistic* or
(01) *variable mass.*

This is a term that already raises a number of uncertainties, probably because our concept of "mass" is so unclear.

(02) The mass of a material body is a measure of its resistance to a change in its state of motion caused by a given force.

This is certainly not a definition of mass, but it can be accepted as a behavioural property.

(03) The larger the mass, the smaller the acceleration.

Already things are getting difficult. This can probably be accepted for most known forces applied to known masses, but it doesn't - for example - apply to gravity.

(04) If a material body is already moving at a speed approaching the speed of light, it must offer increasing resistance to any further acceleration so as not to cross the threshold of *c*.

This is a highly doubtful assertion. It seems to me to be based on a false assumption, namely that it is the mass of a material body that is invariably responsible for the postulated speed limitation, never the force.

Ignoring *fields* of force, for the moment, I would instead argue that our experience tells us a different story. Every force acting on a body has *its* limitations.

Take, for example, a locomotive that can push a wagon in front of it. This locomotive can generate motion up to a certain speed, let's say 100 km/h. At this speed its motive force reaches an upper limit, being output on its motion and on overcoming friction, air resistance, etc. When pushing the wagon additionally, it can usually *almost* reach the same speed, though its acceleration is slower to the extent of the increased mass (as stated under point 03, above). Normally, however, neither of these vehicles will ever break the "light barrier".

To move the wagon faster, another - faster - locomotive is needed, i.e. the force generated by the locomotive, or by its engine, is the limiting factor. Or put another way, the greater the motion required, the greater the force needed.

If you want to travel at a speed greater than the speed of light, therefore, you must have a force that can be applied - or can travel - at speeds greater than the speed of light.

It's not that the wagon becomes too heavy to push any faster, but that the locomotive can't travel any faster. Or put another way, *potential force decreases with increasing velocity*.

(05) Hence the special theory of relativity leads to the conclusion that the mass of a moving body m is related to the mass that it would have if at rest, m_o, divided by the square root of one minus the fraction v^2/c^2:
(06) $m = (1-v^2/c^2)^{-1/2}m_o$
(07) The changing value of the mass of the moving body, m, is called the relativistic mass.
(08) As v approaches c, the figure within the parentheses approaches zero and, ultimately, $m = m_o/0$, which would be an infinitely large number.

Notice first of all that dividing m_0 by $\sqrt{(1-v^2/c^2)}$ - or by $(1-v^2/c^2)^{1/2}$, as under point 05 - is the same as *multiplying* by the inverse of this term (06), which means that the product $(1-v^2/c^2)^{-1/2}m_0$, *increases* in value as compared to m_0.

One question immediately arises in this connection, however: why *multiply* this time instead of *divide*[8]? We should try to ignore the fact that mass increase is *wanted*, here, since this is not a mathematical justification.

And how can we illustrate this in terms of our generalized Lorentz triangle[9]?

---

[8] As in the case of *length contraction*.
[9] *Not* to be confused with the *Lorentz triangle of inequalities!*

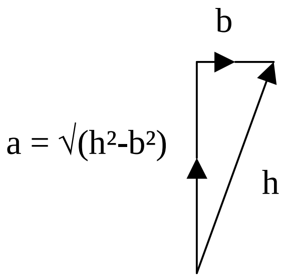

Fig. 6
The Lorentz triangle
(h = hypothenus, a & b are the shorter sides)

If we now substitute-in the values for mass, in keeping with the use generally made of the Lorentz factor, we must put b = m_0, i.e.:

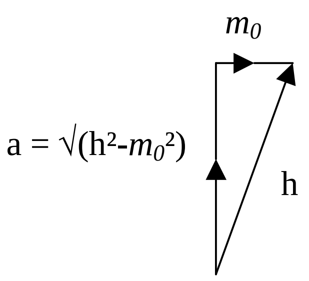

Fig. 7
The Lorentz triangle for mass
(h = hypothenus)

Now, deriving the Lorentz factor for these terms, we have

$$h / a$$

$$= h / \sqrt{(h^2 - m_0^2)}$$ substituting for a, as in Fig. 7

$$= h / \sqrt{(h^2(1 - m_0^2/h^2))}$$ extracting h² from the terms in brackets

$$= h / h \sqrt{(1 - m_0^2/h^2)}$$ extracting h, as the root of h², from the remaining denominator

$$= 1 / \sqrt{(1 - m_0^2/h^2)}$$ cancelling h, above and below

And multiplying m_0 by our Lorentz factor, as stated above, we have:

$$(m_0)\ (1 / \sqrt{(1 - m_0^2/h^2)})$$

$$= m_0 / \sqrt{(1 - m_0^2/h^2)}$$ since $1 \times m_0 = m_0$

But how are we to conceive this and how does it help us? Have we perhaps put our term for rest mass, m_0, in the wrong place? And where does the velocity aspect come in?

Let's try another option, this time for momentum, which is mass x velocity:

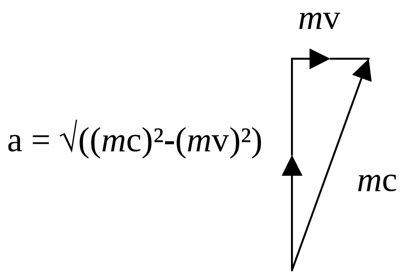

Fig. 8
The Lorentz triangle for momentum

This time our Lorentz factor is

	$mc / \sqrt{((mc)^2-(m\mathrm{v})^2)}$	
=	$mc / \sqrt{((mc)^2(1-(m\mathrm{v})^2/(mc)^2))}$	extracting $(mc)^2$ from the terms in brackets
=	$mc / mc \sqrt{(1-(m\mathrm{v})^2/(mc)^2)}$	extracting mc, as root $(mc)^2$, from the remaining denominator
=	$1 / \sqrt{(1-(m\mathrm{v})^2/(mc)^2)}$	cancelling mc, above and below
=	$1 / \sqrt{(1-\mathrm{v}^2/c^2)}$	cancelling m in the denominator term, above and below

According to points 05 and 06, however, $m = m_0 / \sqrt{(1-\mathrm{v}^2/c^2)}$. So where do we find m_0 in our illustration, with which to multiply our Lorentz factor?

One easy option here would appear to be to simply substitute-in this term for m in $m\mathrm{v}$ (Fig. 8) and already we have our result (multiplied by v):

	$m\mathrm{v}$	taken from Fig. 8
=	$m_0\mathrm{v} / \sqrt{(1-\mathrm{v}^2/c^2)}$	substituting $m_0 / \sqrt{(1-\mathrm{v}^2/c^2)}$ for m

(09) The relativistic mass number may be interpreted as indicating that the relativistic mass of a body exceeds its rest mass m_o by an amount that equals its kinetic energy E, divided by c^2: $m - m_o = E/c^2$. [10]

(10) Hence the hypothesis that generally the energy is c^2 times the mass, or $E = mc^2$, and that energy and mass are, in fact, equivalent physical concepts, differing only by the choice of their units.

This fact (see footnote 10) means that the definition of mass (m), as included in $E=mc^2$, does not embrace the entire mass of the body but only the *mass increase*, i.e. total mass *less* rest mass (m-m_0). So why, if mass *is* energy, does the rest mass play no part in the final energy calculation, $E=mc^2$?

[10] Notice that the m in m-m_0 here is not the m in $E=mc^2$, the latter being the *result* of m-m_0. The first m in this subtraction is known as the *relativistic mass* - sometimes denoted as m_{rel}- whereas the m in $E=mc^2$ is the *mass increase*.

This would instead suggest to me that the increase in energy is not stored as mass, but in some other form, e.g. velocity.

In the end, whatever it is that is supposed to represent *mass increase* again falls foul of our above analysis of the Lorentz factor, i.e. that the velocity (v) does *not* approach c, in which case the argument in favour of mass increase does not apply.

***

Before concluding I would like to briefly consider an object being accelerated in a *field* of force such that the same magnitude of force acts on the object at all times. This would imply something like a force that is itself *accelerating*. I know of no such force.

Must we, then, accept that nothing can travel faster than the speed of light?

Consider the following example. An aircraft in the air can travel at 300 km/h with respect to the earth. There is a jet stream, however, that travels at 400 km/h with respect to the earth. If the aircraft travels with the jet stream it therefore ends up travelling at 700 km/h. Now imagine this jet stream travelling at c-200 km/h.[11] If the aircraft travels with the jet stream can it travel faster than c?

I would say yes, it can.

***

This paper has raised serious doubts as to the general application of the Lorentz factor and calls on supporters of its use in relativistic physics to *explain* this application to the "terms" it changes (length, time, and mass). Such an explanation would also have to define *when* and *why* to multiply or divide.

It has shown that the terms appearing in the Lorentz factor - such as v and c - are no *terms* at all, but are merely not-yet-substituted-in numbers.

It has shown that, if *length contraction* and *time dilation* "exist", they take affect simultaneously and to the same extent, with the result that they cancel each other out. The implications for relativity are devastating.

It has also cast doubts on the assumption of *mass increase* with increasing velocity, as opposed to *potential force decrease* with increasing velocity.

And it has shown that $E=mc^2$ also needs rethinking on the basis that - although energy is claimed to *be* mass - the *rest mass* component is entirely ignored by the equation!

****

[11] Not something that is practically feasible, of course, but it is theoretically conceivable.