The Rudiments of the Special Theory of Relativity

by

Bill Smith

1. Measurement

We begin by specifying the units we use in the measurement of physical properties

a) **Length:** The standard unit of length is the metre, as measured against a rigid metre rule.
b) **Time:** The standard unit of time is also the metre, which is obtained as the product of the speed of light and the time light takes to travel one metre. Conversion to true time is obtained by dividing the time in metres by the speed of light. Such times can be measured with a light-clock. A light-clock is an evacuated tube silvered at both ends with some device to count the arrival of photons, as shown in Figure 1.

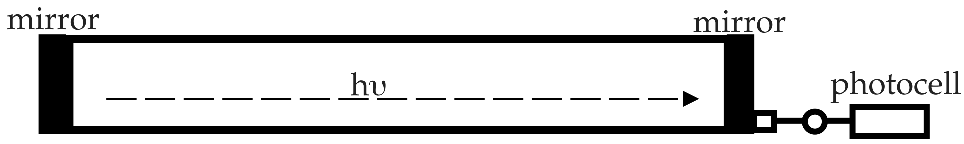

Figure 1. A light-clock

c) **The Velocity of Light:** Using the above defined units of distance and time, the velocity of light c, is 1.

d) **Energy Units:** The dimensions of energy are ML^2T^{-2}, but since time is here measured in metres, this becomes M alone. To convert this into MKS units it is necessary to multiply by c^2 which follows from the conversion of metres of time to seconds. Hence

$$E(MKS) = E(Relativity) \times c^2 . \tag{1}$$

2. Space-Time

Following Minkowski, events occurring in space and time can be marked on a so-called space-time plot, which records event histories. Typical event histories (also known as world-lines) are shown in Figure 2.

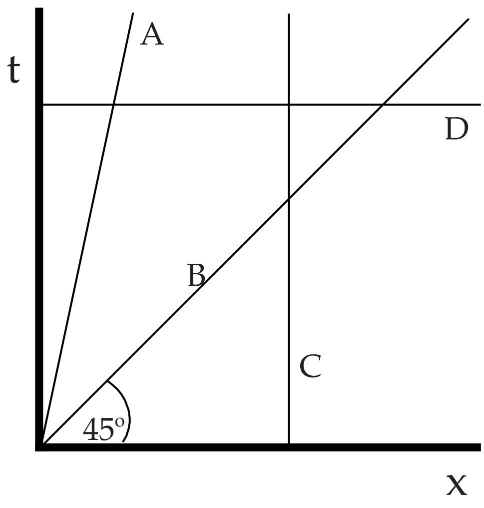

Figure 2. A space-time plot

1. Line A is the world line of an object moving with a constant velocity less than that of light.
2. Line B is the world line of a light photon. It makes an angle 45o to the axes.
3. Line C represents a stationary object.
4. Line D is a line of simultaneity, along which events happen at the same time.

An inertial frame of reference is one which is not accelerated with respect to the distant mass of the Universe. It is a frame of reference in which Newton's laws of motion are valid. There is no universal inertial frame of reference. All inertial frames of reference are in uniform linear motion with respect to each other.
The *Special Theory of Relativity* is concerned with rationalising observations of physical phenomena made with respect to different inertial frames of reference so that the laws of physics may take the same form in all frames. i.e. if F' is a frame of reference moving with a velocity v with respect the frame F, events in the former frame are defined by coordinates (x', y', z', t') and the latter by (x, y, z, t). The mathematical transformation of (x', y', z', t') to (x, y, z, t) and vice versa is a central role in Relativity Theory.

3. The Galilean Transformation

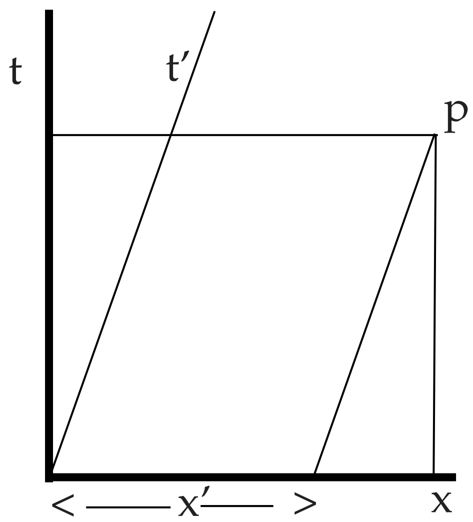

Figure 3. The Galilean Transformation

The Galilean transformation is the classical method of transforming between frames F and F'. It recognises that the position coordinates of an event in one frame will depend on its motion relative to another. i.e. position is assumed to be relative. However it assumes that the time coordinate is universal and is the same in all inertial frames. For convenience we may assume F is a stationary frame and that F' moves at a constant velocity v with respect to F, we may also define the direction of v as the x axis (in both frames) and the origin of the time coordinate as the moment when the Cartesian origins of both F and F' coincide. In this case the equations of the Galilean transformation are :

$$
\begin{aligned}
x' &= x - vt, \\
y' &= y, \\
z' &= z, \\
t' &= t.
\end{aligned} \quad , \qquad (2)
$$

Figure 3 shows these equations at work in a space-time plot.

4. Simultaneity and Special Relativity

The Special Theory of Relativity is founded on two principles:

1. The velocity of light is the same in all inertial frames of reference.
2. The laws of physics are the same in all inertial frames of reference.

These principles quickly lead to contradictions with everyday experience, which Einstein was able to show were only apparent but were otherwise self consistent. The first example of a contradiction is the concept of simultaneity, which is revealed in the following two cases.

a) Light propagation in a stationary frame (Figure 4).
A, B are stationary points in the frame F, C is the mid point. After a time interval t_c a flash of light from C reaches A and B simultaneously.

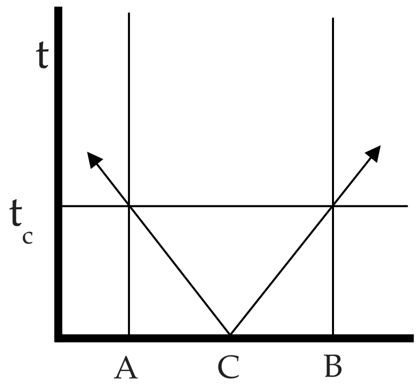

Figure 4. Light propagation in a stationary frame.

b) Light propagation in a moving frame (Figure 5).
Points A,B,C are fixed in the moving frame F' (hence sloped world lines), with distances AC and CB equal. The flash of light from C cuts line A at A′ and line B at B′. The first principle of Special Relativity requires that these two events be simultaneous in the moving frame. A′B′ is thus a *line of simultaneity* in frame F' and must be parallel to the x′-axis in frame F'.

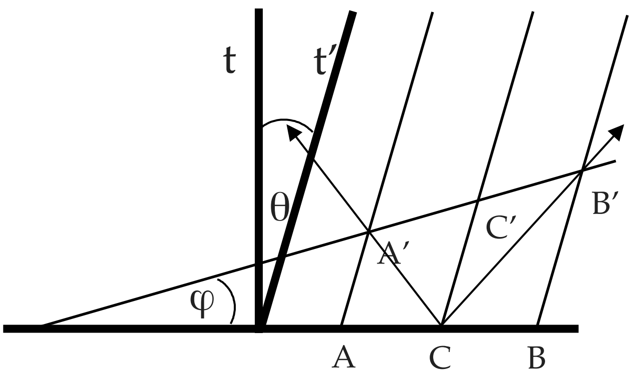

Figure 5. Light propagation in a moving frame

Angle φ defines the slope of the this line, which is calculated as follows:

- Angle A′CB′ is a right angle,
- Points A′CB′ lie on a circle centred on C′,
- Hence A′C′=C′C.

also

- Angle C′A′C=45°+ φ,
- Angle A′CC′=45°+ θ,

but

- Angle C′A′C=A′CC′,
- Hence $\varphi=\theta$.

From this result, we infer that the correct representation of these events on a space-time plot is as shown in Figure 6. A Point P with coordinates (x, y) in the stationary frame F has the coordinates (x', y') in the moving frame F'.

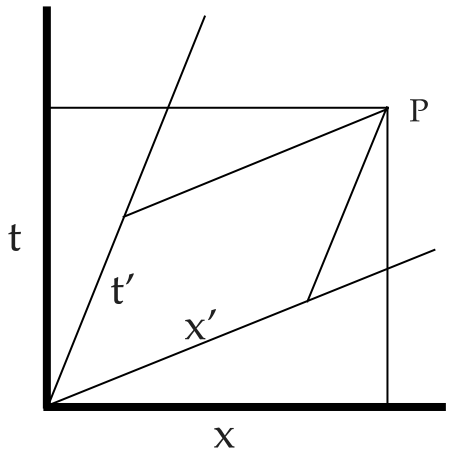

Figure 6. The relativistic transformation between frames F and F'.

It is clear that the Galilean transformation does not properly describe the transformation indicated in Figure 6. The correct form was first given by Lorentz.

5. The Lorentz Transformation

In place of the Galilean transformation we try

$$x'=\gamma(x-vt). \tag{3}$$

For which, by the second principle of relativity, the inverse transform is

$$x=\gamma(x'+vt'), \tag{4}$$

in which it is now assumed that the time coordinate is different in the two frames
Substituting (4) into (3) gives

$$\begin{aligned} x'&=\gamma(\gamma(x'+vt')-vt),\\ x'&=\gamma^2 x'+\gamma^2 vt'-\gamma vt,\\ t&=(\gamma^2 vt'-(1-\gamma^2)x')/\gamma v,\\ t&=\gamma(t'-(1-\gamma^2)x'/\gamma^2 v). \end{aligned} \tag{5}$$

The value of the parameter γ is obtained as follows.
First, by the first principle of Special Relativity, for a light flash:

$$c=\frac{x}{t}=\frac{x'}{t'}=1. \tag{6}$$

Hence we have $x=t$ and $x'=t'$ for light propagation. From the latter we have

$$\begin{aligned}
&\gamma(x'+vt')=\gamma(t'-(1-\gamma^2)x'/\gamma^2 v),\\
&\gamma^2 v(1+v)=\gamma^2 v-(1-\gamma^2),\\
&\gamma^2 v^2=-1+\gamma^2,\\
&\gamma=1/\sqrt{(1-v^2)}.
\end{aligned} \tag{7}$$

Substituting this result into (5) gives

$$\begin{aligned}
&t=\gamma\left\{t'+\frac{\gamma^2(\gamma^{-2}-1)x'}{\gamma^2 v}\right\},\\
&t=\gamma\left\{t'+\frac{v^2 x'}{v}\right\},\\
&t=\gamma(t'+vx').
\end{aligned} \tag{8}$$

Note the resemblance between equations (4) and (8) in these units.
The inverse transform of (8) is

$$t'=\gamma(t-vx). \tag{9}$$

Equations (3), (4), (8) and (9) are the Lorentz Transformations. These are summarized in (10).

$$\begin{aligned}
&x'=\gamma(x-vt),\\
&y'=y,\\
&z'=z,\\
&t'=\gamma(t-vx),
\end{aligned}\quad \text{, and inverse:}\quad
\begin{aligned}
&x=\gamma(x'+vt'),\\
&y=y',\\
&z=z',\\
&t=\gamma(t'+vx').
\end{aligned} \tag{10}$$

The conventional form (i.e. in conventional units where $c\neq1$ and the replacement $t\rightarrow ct$ is made) is presented in (11).

$$\begin{aligned}
&x'=\gamma(x-vt),\\
&y'=y,\\
&z'=z,\\
&t'=\gamma(t-vx/c^2),
\end{aligned}\quad \text{and inverse:}\quad
\begin{aligned}
&x=\gamma(x'+vt'),\\
&y=y',\\
&z=z',\\
&t=\gamma(t'+vx'/c^2).
\end{aligned} \tag{11}$$

The factor γ in this case is

$$\gamma=1/\sqrt{1-v^2/c^2}. \tag{12}$$

6. The Lorentz-Fitzgerald Contraction

A rod of length L' in a uniformly moving frame F' can be measured from a stationary frame if both ends, x_1' and x_2', can be located simultaneously in the stationary frame. From equations (10) we have

$$\begin{aligned}
&x_1'=\gamma(x_1-vt_1),\\
&x_2'=\gamma(x_2-vt_1),\\
&x_2'-x_1'=\gamma(x_2-x_1),\\
&L'=\gamma L.
\end{aligned} \tag{13}$$

Since $\gamma>1$, this means that a moving rod measured from a stationary frame F must appear reduced from its *proper length* (measured in the frame in which it is stationary) by a factor $1/\gamma$. This is the *Lorentz-Fitzgerald* contraction.
This phenomenon may be expressed as: *moving rods appear shorter than their proper length in the moving frame.*

7. Time Dilation

A clock at rest in a moving frame F', marks off time intervals at a different rate from an identical clock in the stationary frame F. Thus a clock at location $x'=0$ in F', measures a time interval of $\Delta t'=(t_2'-t_1')$ which transforms to the frame F as follows.

From equations (10) we have

$$\begin{aligned} &t_1=\gamma(t_1'+vx_1'),\\ &t_2=\gamma(t_2'+vx_1'),\\ &t_2-t_1=\gamma(t_2'-t_1'),\\ &\Delta t=\gamma\,\Delta t'. \end{aligned} \qquad (14)$$

Since $\gamma>1$, this result means that the measured time interval in frame F is greater than that in frame F' i.e. $\Delta t>\Delta t'$. This is the relativistic time dilation effect, summarised by the statement: moving clocks run slow".

8. The Space-Time Invariant

The quantity

$$\tau^2=t^2-x^2 \qquad (15)$$

is an invariant between different inertial frames, as is simply proved. Substituting the Lorentz inverse transform (10) into (15) gives:

$$\begin{aligned} &\tau^2=\gamma^2((t'+vx')^2-(x'+vt')^2),\\ &\tau^2=\gamma^2(t'^2+2\text{vt}'x'+v^2x'^2-x'^2-2\text{vt}'x'-v^2t'^2),\\ &\tau^2=\gamma^2(t'^2(1-v^2)-x'^2(1-v^2)),\\ &\tau^2=t'^2-x'^2. \end{aligned} \qquad (16)$$

In general τ is known as the *proper time*, which is the time recorded by a clock that is stationary in the moving frame F'.
The generalisation to 3D also holds:

$$\tau^2=t^2-x^2-y^2-z^2. \qquad (17)$$

The most important expression of this relationship is the *space-time interval*:

$$\Delta\tau^2=\Delta t^2-\Delta x^2-\Delta y^2-\Delta z^2, \qquad (18)$$

which measures the displacement between two events in space-time. This measure is conserved across all inertial frames and is an important property of space-time. It is the space-time equivalent of Pythagoras' theorem. Note that expressions (15) to (17) hold (in part) because we have defined the frames F and F' in a particular way with regard to orientation and time origin. However (18) holds for all inertial frames regardless of how they are set up with respect to each other.

The corresponding form for (18) using conventional units is

$$\Delta\tau^2=\Delta t^2-(\Delta x^2+\Delta y^2+\Delta z^2)/c^2. \tag{19}$$

9. The Velocity Addition Formula

If space and time have different measures in different inertial frames, how does the velocity of an object transform between frames F and F' ?

From the Lorentz equations we may write the differential forms

$$\begin{aligned} dx' &= \gamma(dx-vdt), & dx &= \gamma(dx'+vdt'), \\ dy' &= dy, & dy &= dy', \\ dz' &= dz, & dz &= dz', \\ dt' &= \gamma(dt-vdx), & dt &= \gamma(dt'+vdx'). \end{aligned} \tag{20}$$

An object moving with velocity $\vec{u}$ and $\vec{u}'$ with respect to frames F and F' respectively has components

$$\begin{aligned} u_x &= dx/dt, & u_x' &= dx'/dt', \\ u_y &= dy/dt, & u_y' &= dy'/dt', \\ u_z &= dz/dt, & u_z' &= dz'/dt'. \end{aligned} \tag{21}$$

Hence we may write (20) directly as

$$\begin{aligned} u_x' &= \frac{dx'}{dt'}=\frac{(u_x-v)}{(1-u_x v)}, & u_x &= \frac{dx}{dt}=\frac{(u_x'+v)}{(1+u_x' v)}, \\ u_y' &= \frac{dy'}{dt'}=\frac{u_y}{\gamma(1-u_x v)}, & u_y &= \frac{dy}{dt}=\frac{u_y'}{\gamma(1+u_x' v)}, \\ u_z' &= \frac{dz'}{dt'}=\frac{u_z}{\gamma(1-u_x v)}, & u_z &= \frac{dz}{dt}=\frac{u_z'}{\gamma(1+u_x' v)}, \end{aligned} \tag{22}$$

which represents the transformation of the velocity $\vec{u}$ in frame F to the velocity $\vec{u}'$ in frame F' and *vice versa*. It is useful to note that

$$\vec{u}\cdot\vec{v}=u_x v \quad \text{and} \quad \vec{u}'\cdot\vec{v}=u_x' v, \tag{23}$$

The velocity transforms in conventional units are

$$\begin{aligned} u_x' &= \frac{(u_x-v)}{(1-u_x v/c^2)}, & u_x &= \frac{(u_x'+v)}{(1+u_x' v/c^2)}, \\ u_y' &= \frac{u_y}{\gamma(1-u_x v/c^2)}, & u_y &= \frac{u_y'}{\gamma(1+u_x' v/c^2)}, \\ u_z' &= \frac{u_z}{\gamma(1-u_x v/c^2)}, & u_z &= \frac{u_z'}{\gamma(1+u_x' v/c^2)}. \end{aligned} \tag{24}$$

From (22) we can easily obtain a useful expression relating the magnitudes u and u' of the velocities $\vec{u}$ and $\vec{u}'$, which is

$$u'^2=1-\frac{(1-u^2)}{\gamma^2(1-\vec{u}\cdot\vec{v})^2}, \qquad u^2=1-\frac{(1-u'^2)}{\gamma^2(1-\vec{u}'\cdot\vec{v})^2}. \tag{25}$$

and in conventional units this becomes

$$u'^2 = c^2\left(1 - \frac{(1-u^2/c^2)}{\gamma^2(1-\vec{u}\cdot\vec{v}/c^2)}\right), \qquad u^2 = c^2\left(1 - \frac{(1-u'^2/c^2)}{\gamma^2(1-\vec{u}'\cdot\vec{v}/c^2)}\right). \tag{26}$$

In equations (24) and (26) note that γ is given by equation (12).

The velocity transform has some interesting properties. For example, a photon travelling with a velocity 1 (i.e. at light speed) along the x coordinate in frame F' has a velocity in frame F of

$$u_x = \frac{(1+v)}{(1+1.\mathrm{v})} = 1. \tag{27}$$

So the velocity of the photon in frame F is still only the velocity of light. This is true even if the velocity of frame F' is itself travelling at light speed.

10. The Headlight Effect

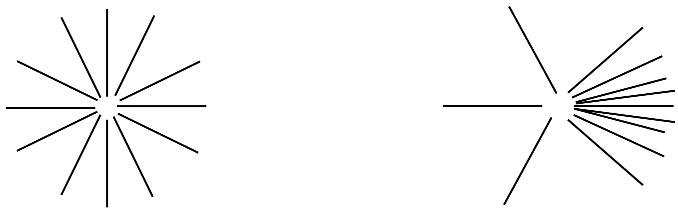

Figure 7. The headlight effect on a moving radiant light source

An interesting consequence of the velocity transform (22) is the *headlight effect*. This involves a radiant light source moving at relativistic speed with respect to an observer. The effect comes about as follows. A light ray in frame F' has a velocity vector (u_x', u_y', u_z') such that

$$u_x'^2 + u_y'^2 + u_z'^2 = 1 \tag{28}$$

If we assume that $v \approx 1$ then the velocity transform (22) becomes

$$\begin{aligned} u_x &\approx (1+u_x')/(1+u_x') \approx 1, \\ u_y &\approx u_y'/\gamma(1+u_x') \approx 0, \\ u_z &\approx u_z'/\gamma(1+u_x') \approx 0, \end{aligned} \tag{29}$$

which show that, in frame F, the x-component of the photon velocity is almost the speed of light, but that it is very small in the perpendicular directions. Thus in the frame F' light from a source may be emitted isotropically in all directions, while in the observer's frame F it is beamed in the direction of the motion of the source. See Figure 7.

The headlight effect occurs in synchrotron radiation, which arises in astrophysics and laboratory synchrotrons.

11. Relativistic Mass and Momentum

Momentum plays a central role in relativistic dynamics as it does in Newtonian dynamics, but it must be couched in an appropriate form. First we derive a useful relation from the invariance of the space-time interval (18), which we write as

$$\tau^2 = \Delta t^2 - \Delta r^2 = \Delta t'^2 - \Delta r'^2 . \tag{30}$$

From this we can proceed as follows:

$$\begin{aligned}
&\Delta t^2 (1-u^2) = \Delta t'^2 (1-u'^2),\\
&\Delta t^2 (1-u^2) = \gamma^2 (\Delta t - v \Delta x)^2 (1-u'^2),\\
&\Delta t^2 (1-u^2) = \gamma^2 \Delta t^2 (1-v u_x)^2 (1-u'^2),\\
&(1-u^2) = \gamma^2 (1-\vec{v}\cdot\vec{u})^2 (1-u'^2),\\
&\gamma_u'^2 = \gamma^2 (1-\vec{v}\cdot\vec{u})^2 \gamma_u^2,\\
&\gamma_u' = \gamma\, \gamma_u (1-\vec{v}\cdot\vec{u}).
\end{aligned} \tag{31}$$

where the 3_d line employs the inverse Lorentz transform of $\Delta t'$ (10) and

$$\gamma_u = 1/\sqrt{1-u^2} \quad \text{and} \quad \gamma_u' = 1/\sqrt{1-u'^2}. \tag{32}$$

Using the final identity of (31) we can easily modify the velocity transforms (22) to obtain

$$\begin{aligned}
&\gamma_u' u_x' = \gamma_u \gamma (u_x - v)\\
&\gamma_u' u_y' = \gamma_u u_y\\
&\gamma_u' u_z' = \gamma_u u_z
\end{aligned} \quad . \tag{33}$$

We now define the *relativistic mass* and *relativistic momentum* for the frames F' and F as follows:

$$\begin{aligned}
&p_x' = \gamma_u' m_o u_x', \qquad && p_x = \gamma_u m_o u_x,\\
&p_y' = \gamma_u' m_o u_y', && p_y = \gamma_u m_o u_y,\\
&p_z' = \gamma_u' m_o u_z', && p_z = \gamma_u m_o u_z,\\
&m' = \gamma_u' m_o, && m = \gamma_u m_o,
\end{aligned} \tag{34}$$

where m_o is the so-called *rest mass* (i.e. the mass possessed by a stationary object.) Using the relations (33) and the definitions (34) we obtain

$$\begin{aligned}
&p_x' = \gamma (p_x - m v),\\
&p_y' = p_y,\\
&p_z' = p_z,\\
&m' = \gamma (m - p_x v).
\end{aligned} \tag{35}$$

We immediately see that the masses and momenta defined in (34) are interrelated through the Lorentz transformation (10), with mass replacing time and the momentum vector replacing the position vector. The inverse transform of (35) is similarly obtained as:

$$\begin{aligned}
&p_x = \gamma (p_x' + m' v),\\
&p_y = p_y',\\
&p_z = p_z',\\
&m = \gamma (m + p_x' v).
\end{aligned} \tag{36}$$

The mass an momentum defined in this way are the proper forms for a relativistic description of dynamics.

12. Mass and Energy

The expression for relativistic mass permits the derivation of an important relationship between mass and energy. The kinetic energy of a moving body is equal to the work done in increasing its speed from zero to v. It may be calculated from the classical integral

$$K=\int_0^{\infty} F\,dx=\int_0^{v}\frac{d(mu)}{dt}=\int_0^{v} u\,d(mu). \tag{37}$$

Using (42) this may be written as

$$K=\int_0^{v} u\,d(m_o\gamma u), \tag{38}$$

which is easily integrated by parts to give

$$K=[m_o u^2\gamma]_o^v-\int_0^{v}\frac{m_o u}{\sqrt{1-u^2}}\,du \tag{39}$$

Hence

$$\begin{aligned} K&=[m_o u^2\gamma]_o^v+[m_o u\sqrt{1-u^2}]_0^v,\\ &=m_o v^2/\sqrt{1-v^2}+m_o\sqrt{1-v^2}-m_o,\\ &=\frac{(m_o v^2+m_o(1-u^2))}{\sqrt{1-v^2}}-m_o,\\ &=\frac{m_o}{\sqrt{1-v^2}}-m_o,\\ K&=m-m_o. \end{aligned} \tag{40}$$

This is in relativistic units. It is useful to rewrite the equation (40) in MKS units:

$$mc^2=K+m_o c^2, \tag{41}$$

where we identify the left hand side (mc^2) as the *total* energy of the body, which is comprised of the kinetic energy K plus $m_o c^{2,}$ which is interpreted as the rest *energy* of the body i.e. energy the body possesses by virtue of its rest mass $m_{o.}$ This is the origin of the famous equation:

$$E=mc^{2.} \tag{42}$$