Force and gravitation in special relativity

by

B.L.Coleman

Revised version 18 Sept 2006

Summary

Newton's second law is conventionally extrapolated by postulate from classical physics to special relativity as "force equals rate of change of relativistic momentum vector". Applied to calculate an orbital trajectory it predicts a perihelion advance one-third of the correct magnitude. This paper derives, by descent from general relativity, a slightly different law based on rate of change of the energy-momentum tensor and not of the momentum vector and the correct prediction is given. Deflection of light is also calculated.

Classification

03.30 Special relativity

04.20 Classical general relativity

Notation

Vectors are written $\vec{\mathbf{u}} \equiv (u_1,u_2,u_3)$. Unit vectors or tensors are written $\hat{\mathbf{v}}$, $\hat{\mathbf{f}}$, $\hat{\mathbf{T}}$ *etc*. Greek symbols are scalars. Velocity is measured in units for which *c*, the speed of light is unity. A vector may also be considered as a column matrix. Subscript s means measured in the observer's framework and m in a framework moving with respect to the observer. $\vec{\mathbf{o}}$ is the zero vector, **I** the unit matrix and $\boldsymbol{\eta}$ the matrix whose leading diagonal is (1,-1,-1,-1) and is zero elsewhere. All parameters are real.

1. Introduction

1.1. Background

In classical physics Newton's second law is expressed in two ways

(1.1.1) $\vec{\mathbf{f}} = m\mathrm{d}\vec{\mathbf{v}}/\mathrm{d}t$ (mass times acceleration)

(1.1.2) $\vec{\mathbf{f}} = \mathrm{d}(m\vec{\mathbf{v}})/\mathrm{dt}$ (rate of change of momentum)

Gravitational theory constructs $\vec{\mathbf{f}}$ from the inverse-square law of attraction between bodies of masses *M* and *m*

(1.1.3) $\vec{\mathbf{f}} = (\text{-}GMm/r^3)\hat{\mathrm{r}}$

This paper is restricted to *M* being a massive body at the origin of coordinates and *m* a relatively small satellite. The combination of (1.1.1) and (1.1.3) leads to an accurate description of *m*'s elliptical orbit around *M* save for the very small advance of the perihelion, some 43 seconds of arc per century for Mercury around the Sun.

(1.1.2) and (1.1.3) are extrapolated into postulates of special relativity save that *m* is now taken to mean $m_o\gamma$, the relative mass, where $\gamma = (1 - \vec{\mathbf{v}}^{\mathrm{T}}\vec{\mathbf{v}})^{-1\backslash 2}$. An advance is predicted but of only a numerical sub-multiple of the measured value. At present the correct prediction is found by going to general relativity, formulated around two principles[1]:

(a) The principle of covariance, that the laws of physics must be independent of the space-time coordinates.

(b) The principle of equivalence, that the fundamental tensor $\mathbf{g}_{\mu\nu}$ of the metric can be chosen to account for the presence of a gravitational field.

Einstein applied the principles to formulate the law connecting his tensor $\mathbf{G}^{\mu\nu}$ with the energy-momentum tensor

$$\textbf{(1.1.4)}\quad \mathbf{T}^{\mu\nu} = \rho_{\mathrm{o}}(\mathrm{d}x^{\mu}/\mathrm{d}s).(\mathrm{d}x^{\nu}/\mathrm{d}s)$$

representing the matter in the field, thus

$$\textbf{(1.1.5)}\quad \mathbf{G}^{\mu\nu} = -4\pi\rho_{\mathrm{o}}\,[-\mathbf{g}^{\mu\nu} + 2(\mathrm{d}\boldsymbol{x}^{\mu}/\mathrm{d}\boldsymbol{s}).(\mathrm{d}\boldsymbol{x}^{\nu}/\mathrm{d}\boldsymbol{s})]$$
$$= -4\pi\rho_{\mathrm{o}}\,\hat{\mathbf{T}}^{\mu\nu}$$

so written to relate to analysis below.

Schwarzschild applied (1.1.5) to the case of $\rho_{\mathrm{o}} = 0$, to determine a spherically symmetrical metric of a space empty save for the central mass, a widely used form of which is

$$\textbf{(1.1.6)}\quad \mathrm{d}s^2 = (1-2GM/r)\mathrm{d}t^2 - (1-2GM/r)^{-1}\mathrm{d}r^2 - r^2\mathrm{d}\theta^2 - r^2\sin^2\theta\mathrm{d}\varphi^2$$

Einstein further proposed the geodesic postulate; the path of a small mobile in a space of that metric will be a geodesic. This is found using the calculus of variations and correctly predicts the perihelion advance, though without any reference to the presence of a mobile.

Eddington[2] developed an alternative approach in which the presence of the satellite is acknowledged as a localized energy-momentum tensor located in a small cylindrical tube around its worldline, in an otherwise empty space. Such a tensor has inherent zero divergence and the equations of motion result when that divergence is calculated assuming (1.1.6) to be the metric containing the gravitational potential even within the tube, thereby ignoring the gravitational effect of the mobile itself.

1.2. Object

Special relativity is situated between classical physics and general relativity. As noted above its dynamics are conventionally postulated by extrapolation from (1.1.2) to

(1.2.1) $\vec{\mathbf{f}} = \mathrm{d}(m_0 \gamma\, \vec{\mathbf{v}})/\mathrm{d}t$

The idea behind this paper is to replace (1.2.1) with a force law derived by descent from Eddington's approach but suiting conditions in the Minkowski metric

(1.2.2) $\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}r^2 - r^2\mathrm{d}\theta^2 - r^2\sin^2\theta\mathrm{d}\varphi^2$

considered as the degenerate form of a Schwarzschild metric, where gravitational effects, no longer incorporated into the local metric, are ascribed to the inverse-square action at a distance law (1.1.3). In this way an updated form of Newton's second law is found which will be shown to predict correctly the advance of the perihelion and the deflection of light.

2. Gravitational theory in special relativity

2.1. Reprise of classical theory

In both classical and relativity physics gravitational and inertial masses are taken equal and it is convenient to deal with force per unit mass $\hat{\mathbf{f}}$ and to regard (1.1.1) as

(2.1.1) $\hat{\mathbf{f}} = \mathrm{d}\,\vec{\mathbf{v}}/\mathrm{d}t$

This is shorthand for

(2.1.2) $\hat{\mathbf{f}} = \lim(\vec{\mathbf{v}}_{t+\delta t} - \vec{\mathbf{v}}_{t})/\delta t$

as $\delta t \to 0$. At any stage before the limit is reached (2.1.2) can be rewritten

(2.1.3) $(\hat{\mathbf{f}}\,\delta t) + \vec{\mathbf{v}}_{\mathrm{t}} = \vec{\mathbf{v}}_{\mathrm{t}+\delta\mathrm{t}}$

where $(\hat{\mathbf{f}}\,\delta t)$, written $\delta\,\vec{\mathbf{u}}$ below for convenience, is recognized as a velocity increment adding vectorially to $\vec{\mathbf{v}}_{\mathrm{t}}$ to create $\vec{\mathbf{v}}_{\mathrm{t}+\delta\mathrm{t}}$.

2.2. Conventional update to special relativity

In special relativity the elements of $(\hat{\mathbf{f}}\,\delta t) = \delta\,\vec{\mathbf{u}}$ are unchanged, but the kinematic algorithm, vector addition of velocities, is no longer valid for composing $\delta\,\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}_{\mathrm{t}}$ to give $\vec{\mathbf{v}}_{\mathrm{t}+\delta\mathrm{t}}$. The only other tools currently available for constructing an appropriate algorithm are the four-vector $\mathbf{V}(\gamma, \gamma\,\vec{\mathbf{v}})$ and the boost, expressed here as a proper Lorentz matrix

(2.2.1) $$\mathbf{B}(\gamma, \gamma\,\vec{\mathbf{v}}) \equiv \begin{bmatrix} \gamma & \gamma\,\vec{\mathbf{v}}^{\mathrm{T}} \\ \gamma\,\vec{\mathbf{v}} & \mathbf{I} + (\gamma\,\vec{\mathbf{v}})(\gamma\,\vec{\mathbf{v}}^{\mathrm{T}})/(1+\gamma) \end{bmatrix}$$

The conventional law (1.2.1) can now be written

(2.2.2) $\hat{\mathbf{f}} = \gamma^{-1}\mathrm{d}(\gamma\,\vec{\mathbf{v}})/\mathrm{d}t$

and the presence of γ^{-1} rules out an algorithm based on $\mathbf{V}$, the four-vector, which has no intrinsic algebra other than vector addition.

It is shown in the Appendix that the successive boost algorithm, regarded here as a purely formal device for composing velocities,

(2.2.3) $\mathbf{B}(\varepsilon, \varepsilon\delta\,\vec{\mathbf{u}})\mathbf{B}(\gamma_{\mathrm{t}}, \gamma_t\,\vec{\mathbf{v}}_{\mathrm{t}}) = \mathbf{B}(\gamma_{t+\delta t}, \gamma_{t+\delta\,\mathrm{t}}\,\vec{\mathbf{v}}_{\mathrm{t}+\delta\mathrm{t}})\mathbf{R}$

where **R** is some rotation matrix needed for congruity, leads to

(2.2.4) $\varepsilon\delta\,\vec{\mathbf{u}}/(1+\varepsilon) = [(\gamma_{\mathrm{t}+\delta\,\mathrm{t}}\ \vec{\mathbf{v}}_{\mathrm{t}+\delta\,\mathrm{t}}) - (\gamma_t\,\vec{\mathbf{v}}_{\mathrm{t}})]/(\gamma_{\mathrm{t}+\delta\,\mathrm{t}} + \gamma_{\mathrm{t}})$

and in the limit to (2.2.2). It may be noted that (2.2.3) with appropriate interpretation is used in calculating the Thomas precession from the rotation matrix.

2.3. Proposed update

The proposed algorithm is based on using, in place of the boost, the unit energy momentum tensor

(2.3.1) $\hat{\mathbf{T}}^{\mu\nu} = -\mathbf{g}^{\mu\nu} + 2(\mathrm{d}x^{\mu}/\mathrm{d}s).(\mathrm{d}x^{\nu}/\mathrm{d}s)$

which in the Minkowski metric reads

(2.3.2) $\hat{\mathbf{T}}(\gamma,\gamma\,\vec{\mathbf{v}}) = -\boldsymbol{\eta} + 2\,\mathbf{V}(\gamma,\gamma\,\vec{\mathbf{v}})\,\mathbf{V}^{\mathrm{T}}(\gamma,\gamma\,\vec{\mathbf{v}})$

Remarkably, when written as a matrix

(2.3.3) $\hat{\mathbf{T}}(\gamma,\gamma\,\vec{\mathbf{v}}) \equiv \begin{bmatrix} 2\gamma^2 - 1 & 2\gamma^2\,\vec{\mathbf{v}}^{\mathrm{T}} \\ 2\gamma^2\,\vec{\mathbf{v}} & \mathbf{I} + 2(\gamma\,\vec{\mathbf{v}})(\gamma\,\vec{\mathbf{v}}^{\mathrm{T}}) \end{bmatrix}$

direct multiplication using (2.2.1) shows that

(2.3.4) $\hat{\mathbf{T}}(\gamma,\gamma\,\vec{\mathbf{v}}) = \mathbf{B}(\gamma,\gamma\,\vec{\mathbf{v}})\,\mathbf{B}(\gamma,\gamma\,\vec{\mathbf{v}})$
$= \mathbf{B}(2\gamma^2 - 1, 2\gamma^2\,\vec{\mathbf{v}})$

and is therefore a proper Lorentz matrix, though not a boost the name reserved for the transformation operator of special relativity kinematics. It may be noted that it is a tensor and as such is form invariant under Lorentz transformation. A boost is not.

Replacing the arguments of $\mathbf{B}$ in (2.2.3) by those of $\hat{\mathbf{T}}$ in (2.3.4), *e.g.* $(\gamma,\gamma\,\vec{\mathbf{v}}) \rightarrow (2\gamma^2 - 1, 2\gamma^2\,\vec{\mathbf{v}})$, and amending (2.2.4) accordingly results in

(2.3.5) $\delta\,\vec{\mathbf{u}} = [(\gamma^2{}_{\mathrm{t}+\delta\,\mathrm{t}}\,\vec{\mathbf{v}}_{\mathrm{t}+\delta\mathrm{t}}) - (\gamma^2{}_{t}\,\vec{\mathbf{v}}_{\mathrm{t}})]/(\gamma^2{}_{\mathrm{t}+\delta\,\mathrm{t}} + \gamma^2{}_{\mathrm{t}} - 1)$

and in the limit a law of force

(2.3.6) $\hat{\mathbf{f}} = (2\gamma^2 - 1)^{-1}\mathrm{d}(\gamma^2\,\vec{\mathbf{v}})/\mathrm{d}t$

Although (2.3.6) appears to differ widely from (2.2.2) a better comparison is afforded if they are split into components parallel to and normal to the unit directional velocity vector. Thus, writing v for the speed and including the classical (2.1.1) in the comparison

$$(2.3.7)\qquad \mathrm{d}\vec{\mathbf{v}}/\mathrm{d}t = (\mathrm{d}v/\mathrm{d}t)\hat{\mathbf{v}} + v\mathrm{d}\hat{\mathbf{v}}/\mathrm{d}t$$
$$(2.3.8)\qquad \gamma^{-1}\mathrm{d}(\gamma\,\vec{\mathbf{v}})/\mathrm{d}t = (1-v^2)^{-1}(\mathrm{d}v/\mathrm{d}t)\hat{\mathbf{v}} + v\mathrm{d}\hat{\mathbf{v}}/\mathrm{d}t$$
$$(2.3.9)\qquad (2\gamma^2-1)^{-1}\mathrm{d}(\gamma^2\,\vec{\mathbf{v}})/\mathrm{d}t = (1-v^2)^{-1}(\mathrm{d}v/\mathrm{d}t)\hat{\mathbf{v}} + (1+v^2)^{-1}v\mathrm{d}\hat{\mathbf{v}}/\mathrm{d}t$$

(2.3.8) and (2.3.9) have the same component parallel to the velocity in which the factor $(1-v^2)^{-1}$ generates the rotating ellipse of the motion. They differ in the normal component by the factor $(1+v^2)^{-1}$ which determines the correct rate of advance. Both factors are insignificant when v<<1 as in most gravitational calculations where the classical (2.3.7) applies.

3. Perihelion advance

In (2.3.6), taking

$$(3.1)\qquad \hat{\mathbf{f}} = (-GM/r^2,\ \ 0,\ \ 0)$$
$$(3.2)\qquad \vec{\mathbf{v}} = (\mathrm{d}r/\mathrm{d}t,\ r\mathrm{d}\vartheta/\mathrm{d}t,\ r\sin\vartheta\mathrm{d}\varphi/\mathrm{d}t)$$

the conventional device of supposing the initial motion to be in the plane $\vartheta = \pi/2$ with $\mathrm{d}\vartheta/\mathrm{d}t = 0$ reduces the motion to two-dimensional in that plane. With $\sin\vartheta = 1$ it follows that

$$(3.3)\qquad r^{-1}\mathrm{d}(\gamma^2 r^2\mathrm{d}\varphi/\mathrm{d}t)/\mathrm{d}t = 0$$
so that
$$(3.4)\qquad \gamma^2 r^2\mathrm{d}\varphi/\mathrm{d}t = h$$

where h is a generalization of the areal constant. Using (2.3.6), (3.1) and (3.2) and writing u=1/r leads to

$$(3.5)\qquad \hat{\mathbf{f}}^{\mathrm{T}}\,\hat{\mathbf{v}} = GM\mathrm{d}u/\mathrm{d}t = \mathrm{d}(\ln\gamma)/\mathrm{d}t$$
whence
$$(3.6)\qquad \gamma = \mathrm{d}t/\mathrm{d}s = K\exp(GMu)$$

where K is a constant of integration. Using (3.4) and (3.6) in the two-dimensional form of the Minkowski metric gives the differential equation for the trajectory as

(3.7) $(\mathrm{d}u/\mathrm{d}\vartheta)^2 + u^2 = h^{-2}\{[K\exp(GMu)]^4 - [K\exp(GMu)]^2\}$

Expanding in powers of GMu, which for Mercury is O(2x10^{-8})

(3.8) $(\mathrm{d}u/\mathrm{d}\vartheta)^2 + u^2 = K^2h^{-2}\{(K^2\text{-}1) + (4K^2\text{-}2)(GMu) + (8K^2\text{-}2)(GMu)^2 + \ldots\}$

The equations of a rotating ellipse are

(3.9) $lu = 1 + e.\cos\alpha\vartheta$

where l is the semi latus-rectum and e the eccentricity, so that

(3.10) $(\mathrm{d}u/\mathrm{d}\vartheta)^2 + u^2 = \alpha^2(e^2 - 1)/l^2 + \alpha^2(2/l)\,u + (1 - \alpha^2u^2$

Equating coefficients between (3.8) and (3.10) and noting that

(3.11) $4(K^2 - 1) - 3(4K^2 - 2) + (8K^2 - 2) \equiv 0$

gives as the best fit

(3.12) $\alpha = [1 + 6GM/l + 4(1 - e^2)(GM/l)^2]^{-1/2}$

The first two perihelia occur at $\alpha\vartheta = 0, 2\pi$, an increase of $2\pi/\alpha = 2\pi(1 + 3GM/l)$ between occurrences, to the first order in GM/l, so that the advance of the perihelion per orbital revolution is $2\pi.3GM/l$, the result given by Eddington [3] from general relativity theory.

Repeating the calculation with (2.2.2) as the law of force gives

(3.13) $\gamma r^2\mathrm{d}\varphi/\mathrm{d}t = h$

in place of (3.4) for the equation normal to the motion, and predicts an advance just 1/3 of the correct value.

4. Deflection of light

(3.6), (3.7) and (3.8) are the general equations in plane polar coordinates for motion of a particle in a conic section with the gravity-generating mass at the focus, the origin. If the particle velocity approaches that of light γ, and therefore K, tends to infinity and (3.8) degenerates into

(4.1) $(\mathrm{d}u/\mathrm{d}\vartheta)^2 + u^2 = K^4 h^{-2}\{1 + 4GMu + ...\}$

The appropriate model is now the static hyperbola

(4.2) $lu = 1 + e\cos\vartheta$

with ***e***>>1. The corresponding differential equation is

(4.3) $(\mathrm{d}u/\mathrm{d}\vartheta)^2 + u^2 = (e^2 - 1)/l^2 + (2/l)u$

Equating coefficients between (4.1) and (4.3)

(4.4) $4GM = 2l/(e^2 - 1)$

The physical picture is of a ray of light from a star in the plane of the ecliptic reaching Earth after grazing the Sun's disc of radius R. Because of the pull of gravity the trajectory is not a straight line but that branch of the hyperbola farther away from the focus, the centre of the Sun, and bending away from it.. For a terrestrial observer the star suffers a displacement equal to the angle between the asymptotes of (4.2).

Geometrically this is easily calculated on the (physically non-existent) branch of the hyperbola nearer to the focus. Were there no distortion this would have been a straight line stretching from $\vartheta = -\pi/2$ on Earth to $\vartheta = +\pi/2$ on the star with the nearest point to the focus being at $\vartheta = 0$. With distortion it stretches from $\vartheta = -(\pi/2 + \Delta)$ to $\vartheta = (\pi/2 + \Delta)$ for a total deflection of 2Δ where

$$
\begin{aligned}
(4.5)\quad 0 &= 1 + e\cos(\pi/2 + \Delta) \\
&= 1 - e\sin\Delta
\end{aligned}
$$

giving

(4.6) $1/e = \sin\Delta \cong \Delta$

The physical branch corresponds in the geometrical picture to the analytic continuation of the first branch, varying from $\vartheta = (\pi/2 + \Delta)$ on Earth, through π at the grazing point, to $(3\pi/2 - \Delta)$, $[\equiv -(\pi/2 + \Delta)]$, on the star. However[4], since the right-hand side of (4.2) is now negative u on this branch must also be posited negative so that at the grazing point

$$(4.7) \quad -l/R = 1 + e\,\mathbf{cos}\,\pi = 1\text{-}e$$

Substituting in (4.4)

$$\mathbf{(4.8)} \quad 4GM\,/\,\mathrm{R} = 2\,/(e+1) \cong 2\,/\,e \cong 2\Delta$$

agreeing with Eddington[5].

References

1. Spain B., 'Tensor calculus', 2nd edition, Oliver and Boyd, 1956, p111.

2. Eddington A.S., 'The mathematical theory of relativity', Cambridge, 1957, p125.

3. Eddington A.S., loc cit, p88.

4. Smith C., 'Conic sections', Reprint, MacMillan, 1912, p216.

5. _Eddington A.S., loc cit, p90.

Appendix

Einstein's law for transformation of velocity readings between inertial frameworks with parallel axes is

$$\text{(A1)} \quad \mathbf{V}(\gamma_s, \gamma_s\,\vec{\mathbf{v}}_s) = \mathbf{B}(\gamma_u, \gamma_u\,\vec{\mathbf{u}})\mathbf{V}(\gamma_m, \gamma_m\,\vec{\mathbf{v}}_m)$$

where S_m is moving with velocity $\vec{\mathbf{u}}$ with respect to S_s. Multiplied out this gives

(A2) $\gamma_s = \gamma_u \gamma_m + (\gamma_u \vec{\mathbf{u}})^T (\gamma_m \vec{\mathbf{v}}_m)$

(A3) $\gamma_s \vec{\mathbf{v}}_s = \gamma_m \vec{\mathbf{v}}_m + [\gamma_m + (\gamma_u \vec{\mathbf{u}})^T (\gamma_m \vec{\mathbf{v}}_m)/(1+\gamma_u)](\gamma_u \vec{\mathbf{u}})$

$$= \gamma_m \vec{\mathbf{v}}_m + [\gamma_m + \gamma_m \gamma_u + (\gamma_u \vec{\mathbf{u}})^T (\gamma_m \vec{\mathbf{v}}_m)](\gamma_u \vec{\mathbf{u}})/(1+\gamma_u)$$

Substituting (A2) into the second form of (A3) and rearranging

(A4) $(\gamma_u \vec{\mathbf{u}})/(1+\gamma_u) = (\gamma_s \vec{\mathbf{v}}_s - \gamma_m \vec{\mathbf{v}}_m)/(\gamma_s + \gamma_m)$

The successive boost equation of interest here is

(A5) $\mathbf{B}(\gamma_u, \gamma_u \vec{\mathbf{u}}) \mathbf{B}(\gamma_m, \gamma_m \vec{\mathbf{v}}_m) = \mathbf{B}(\gamma_s, \gamma_s \vec{\mathbf{v}}_s) \mathbf{R}$

Noting that the zero velocity four - vector $\mathbf{V}(1, \vec{\mathbf{o}})$ is unaffected by $\mathbf{R}$ and that

(A6) $\mathbf{B}(\gamma_u, \gamma_u \vec{\mathbf{u}}) \mathbf{V}(1, \vec{\mathbf{o}}) = \mathbf{V}(\gamma_u, \gamma_u \vec{\mathbf{u}})$

etc., multiplying each side of (A5) into $\mathbf{V}(1, \vec{\mathbf{o}})$ reduces it to (A1), so that (A4) applies.