Technical Note

Relativity: 300 years from a principle to reality

H Rahnejat
Wolfson School of Mechanical and Manufacturing Engineering, Loughborough University, Loughborough, Leicestershire LE11 3TU, UK

Abstract: This technical note provides a brief background to the historical developments that led to the generalization of Galileo's principle of relativity, leading to Einstein's general theory of relativity. Therefore, its primary purpose is to mark the contributions made in physics of motion on this occasion—the marking of the new Millennium. To keep with the tradition set in this journal, this note contains a new analysis in respect of reported super-luminal observations in gain-assisted light propagation, in particular vis-à-vis their concordance with the general theory of relativity. The incompleteness principle, as an axiom of observation, is introduced here to resolve this issue.

Keywords: principle of relativity, physics of motion, incompleteness principle

1 INTRODUCTION

Three major principles gave rise to the theories of special relativity and general relativity: instantaneity, simultaneity and equivalence. The former stemmed from the purported, and historically questioned, free falling masses experiment by Galileo from the Leaning Tower of Pisa in 1593. Galileo noted that, in accelerated motions, the instantaneous velocity should be described by the average velocity in a diminutive period of time, referred to ever since as 'infinitesimal'. The definition and measurement of such small quantities was of course resolved in 1821 by Cauchy, with the use of differential limits, taken for granted in the analysis of dynamic systems today (e.g. for Galileo's experiment: $v = \lim_{t\to 0} \Delta h/\Delta t = \mathrm{d}h/\mathrm{d}t$). The profound value of this contribution and the significance of tangential properties of differential calculus were immediately realized. In a way the problems posed for 228 years (1593–1821) were resolved, together with a good description of the concept of instantaneity. One problem, however, remained in this regard, the apparent instantaneous action of the 'force of gravity' in Newton's law of universal gravitation [**1**]. In fact, the use of limits in the differential calculus gave the opportunity of assigning curvature properties to surfaces. This was later exploited by Gauss in the nineteenth century in his *Theorema Egregium*, briefly discussed below, which revolutionized the understanding of geometry and in time provided one of the key elements in understanding the mechanism of gravitational action through the general theory of relativity [**2**]. However, before this could be achieved a physical understanding of the relative motion of bodies in a fundamental manner was required.

It was long appreciated that the Newtonian laws of motion hold true for, as Einstein put it, 'privileged' inertial frames of reference and when an 'absolute' fixed frame of reference can be assumed somewhere in space. Since such a static position cannot be assumed, particularly at speeds encountered in celestial events, the principle of relativity must be adhered to.

As early as the seventeenth century Galileo appreciated that no observer could distinguish between the states of absolute motion and absolute rest. It had become clear to him that a concept such as 'absolute' cannot exist and that observers in relative uniform motion should describe the laws of nature in precisely the same manner. This astute and yet natural deduction was reaffirmed three centuries later by Einstein, who declared the Galileo's principle of relativity as a law of physics, which he termed 'the principle of covariance', as described later on. What had troubled the physicists in the late nineteenth century was the fundamental basis for the understanding of the Michelson–Morley experiment [**3**] (readers should refer to any text on modern physics) and Maxwell's equations for electromagnetic wave propagation. A fundamental explanation for these necessitated the assumption that the speed of light would remain

The MS was received on 12 December 2000 and was accepted after revision for publication on 24 January 2001.

constant, irrespective of the state of motion of the observer. However, the Galilean transformation from one frame of reference of observation to another does not preserve the light cone. To resolve this conundrum, Einstein reasoned that a new transformation that preserves the light cone should be used and that the constancy of the speed of light *in vacuo* be maintained in all inertial frames of reference as a law of physics; this is referred to as the principle of causality. Such a transformation, fortunately, was already at hand, thanks to Fitzgerald's explanation in respect of the Michelson–Morley experiment—that the measured distances contract in the direction of motion by a significant amount at very high speeds. This explanation had paved the way for the derivation of transformation laws that preserve the light cone, known as the Lorentzian transformation [**4**].

Einstein used the uniform motion in inertial frames of reference in order to simplify somewhat the physical studies of motion, without regard to applied forces and inertial properties. This particular branch of dynamics is, of course, known as the kinematic analysis of motion. He noted that unlike the Galilean transformation, which has a simple geometric interpretation within the Euclidean geometry with rigid rotations, translations and reflections, all of which preserve images that are congruent to the original, it appeared that this could not be held to be true for the Lorentzian transformation. Therefore, the Lorentzian transformation had to be put on a similar footing that would be affine to the concept of congruent transformation. The developments in the use of hyperbolic geometry had paved the way for this. In particular Minkowski showed in 1907 that the Lorentz transformation preserved each of the hyperbolas in the H-plane by the notional orthogonal properties of the Minkowski norm: $\|E\|$, where $c^2t^2 - x^2 - y^2 - z^2 = 0$ [**5**]. Therefore, the Minkowski norm (having the dimensions of time) and the interval between two events are absolute quantities, meaning that all uniformly moving observers assign to them the same values. This firmly established the concept of space–time and confirmed the validity of Galileo's principle of relativity. Einstein was then able to resolve the contentious issue of simultaneity of two events at approximately the same place, by observing that since Galilean observers need not agree that any two events had occurred at the same time, the notion of simultaneity lost its physical meaning. This was replaced in effect with 'the time of a stationary system', which determined the constancy of the velocity of light in empty space [**6**].

Thus far, it has been noted how two of the triad of aforementioned principles, instantaneity and simultaneity, led to the establishment of the special theory of relativity. To see the significance of the third principle, equivalence, one should return to the alleged experiment performed by Galileo and note that at the time the effect of gravity was not fully understood. Galileo merely concluded that the rate of fall remained independent of the mass of the object. A century later Newton formulated his law of universal gravitation, arrived at empirically through observation. In Newton's view gravity was a non-discriminatory force that applied to everything instantaneously at a distance without a prescribed contact. Although his empirical deduction remains remarkably accurate, it suffered from two major shortcomings. It was and remains devoid of a physical description and embodied the then vague concept of instantaneity. The preceding paragraph has already dealt with the latter problem. The former observation was unsatisfactory to Newton, when he noted: 'It is inconceivable, that inanimate brute matter, should, without the mediation of something else, which is not material, operate upon and affect other matter without mutual contact. … I believe no man who has in philosophical matters a competent faculty of thinking can ever fall into it (i.e. this conclusion).'

The problem for Newton partly arose in the duality of his description of mass. In his law of universal gravitation, he referred to the mass of an object as the gravitational mass, whereas in his second law of motion the mass of an object was described as the inertial mass. This anomaly was to become clear to Einstein as unnatural. He described the indistinguishable nature of motion in a gravitational field with respect to a fixed frame of reference K, from that in a uniformly accelerated frame of reference K′ in a region of space–time devoid of gravitational action. This led Einstein to the realization that a local gravitational field can be studied using a coordinate frame such as K′ with respect to an inertial frame such as K. This he termed the equivalence principle, which resolved the duality in the definition of mass put forward by Newton. Einstein argued that inertial frames should not hold the privileged position endowed to them in Newtonian physics and that all laws of physics should be covariant (principle of covariance). He noted: 'The laws of physics must be of such a nature that they apply to systems of reference in any motion' [**2**]. Therefore, they should also hold true for any arbitrarily chosen non-inertial (accelerated) frames of reference. Gravity is the most important law of nature. In effect Einstein introduced a generalization of Galileo's principle of relativity—the principle of general relativity.

The equivalence principle, noted by Einstein as his 'happiest thought', opened the way for a physical description of gravity: '… in pursuing the general theory of relativity we shall be led to a theory of gravitation, since we will be able to "produce" a gravitational field merely by changing the system of coordinates'. He noted that a massive body (e.g. the Sun) causes the fabric of space to warp, rather like the effect of a bowling ball placed on a rubber sheet. The difference from the description of gravity by Newton (where the field is described by a force) is that Einstein has described the mechanism by which gravity is transmitted, this being the warping of space [**7, 8**]. Accelerated motions therefore undertake curvilinear paths in space–time. Now, for the first time since the experiment of Galileo in 1593, a theory existed in 1911 that linked firmly with the facts of astronomical observations.

Because gravity could now be explained in terms of the curved space–time, its formulation needed a parametric definition of surfaces which included their intrinsic properties. These properties would ideally describe the curved surfaces in a unique manner, and can therefore be regarded as metric properties. Therefore, in developing his theory of gravitation, leading to the general theory of relativity, Einstein had to return to the intrinsic properties of differential geometry [**9**]. The metric tensor $g_{ij} = W_i \cdot W_j$, defined as the first fundamental form by Gauss, is one of the intrinsic features of a curved surface, which can be used to calculate the local curvature, the length of the curve, the area of a local region of the space, etc. Gauss has shown in *Theorema Egregium* that curvature can be expressed entirely in terms of the derivatives of the metric tensor. Note that, for example, $W_i = \mathrm{d}W/\mathrm{d}q^i$ and $W_j = \mathrm{d}W/\mathrm{d}q^j$ describe the tangent plane at any point P, where map $W{:}R^2 \rightarrow R^3{:}(q^i,q^j) \rightarrow (x, y, z)$. Other fundamental forms can be obtained by the use of the multi-indexed tensorial quantities in *Theorema Egregium*. Therefore, $W_{jk} = \Gamma^i_{jk} W_i + a_{jk} N_j$, where N can be expressed as $N_j = -a^i_j W_i$. Thus, $a_{jk} = a^i_j g_{ik}$, which is referred to as the second fundamental form.

The Christoffel symbol, defined by Γ^i_{jk}, was used by Einstein, who argued that the tensorial quantities are generally covariant for the development of relativity on surfaces. The principle of covariance also determines that the freely falling objects in a gravitation field follow a geodesic path. Therefore, in general,

$$\frac{\mathrm{d}^2\xi^k}{\mathrm{d}t^2} = -\Gamma^k_{ij}\frac{\mathrm{d}\xi^i}{\mathrm{d}t}\frac{\mathrm{d}\xi^j}{\mathrm{d}t}$$

is the geodesic relativistic equation of motion for a free falling object.

It is now possible to return to the topical subject of super-luminal light propagation, claimed to have been observed in anomalous dispersion when a gain-assisted pulse of light has travelled through an atomic vapour cell [**10–12**]. In particular, there is a need to ascertain the validity of such observations in relation to the general theory of relativity.

2 SUPER-LUMINAL EVENTS AND RELATIVITY

Detection of a light pulse at the exit of an atomic vapour cell has been claimed, before its instance of entry, the implication being that a gain-assisted light pulse can exceed the speed of light *in vacuo* [**12**]. The investigators have observed that their finding ‘is not at odds with causality’. This claim, while not in line with the special theory of relativity, is quite justified in the case of accelerated motions, as in the aforementioned experiment. Einstein himself noted that: ‘the law of constancy of the velocity of light *in vacuo*, which constitutes one of the two fundamental assumptions in the special theory of relativity … cannot claim any unlimited validity’ [**13**]. However, it should be noted that the accelerated motion of a pulse of light constitutes its variable propagation with position, which may be interpreted as an anomalous dispersion. In keeping with the theory of general relativity, such a pulse of light would have to undertake a curvilinear path, analogous to accelerated motion in a gravitational field. By virtue of the principle of equality of inertial-induced and gravitational fields within the postulate of relativity (i.e. the equivalence principle), it can be inferred that the gain-assisted pulse of light takes a uniformly accelerated motion in such a curvilinear path and that the deflection of light by such a gravitational field would be infinitesimal, given the finite length of the light path and the strength of the field. Therefore, given an inappreciable angle of deflection from a straight-line path and a short path length, the observed super-luminal behaviour of the light pulse appears to be plausible and in line with the general theory of relativity.

There are, however, some critical observations that make the aforementioned hypotheses not as safe as they appear at first glance. To promulgate on these observations, a simple theoretical experiment can be used, in which a particle in uniform motion at the speed of light is observed by a ‘stationary’ observer. As the field of view of the observer is narrowed to a vanishing point, the world-line of the motion of the particle can become indistinct as viewed by this observer, who views the event at the limit of observations (i.e. the speed of light). An infinitesimal particle with a momentary local existence without an appreciable duration (but nevertheless of finite value—note the resolution of the concept of instantaneity) can be described by a large number of coordinate values in a unidimensional continuum, which deploys a very narrowed field of observation. Therefore, the world-line of the particle can be one of a manifold of parallel lines that cross the field of view of the stationary observer in this unidimensional continuum. This argument can be rephrased by quoting Einstein himself: ‘If this point (a material point) had a momentary existence without duration, then it would be described in space–time by a single system of values. Thus, its permanent existence must be characterised by an infinitely large number of such systems of values, …, thus we have a (unidimensional) line in the four dimensional continuum’ [**14**]. A fast-moving particle of infinitesimal dimension can be viewed, in a narrow frame of observation, to be infinitely far or equally well as infinitesimally close, given an inappreciable duration of observation. There lies the doubt about the validity of the foregone hypotheses and the safety of interpretation of super-luminal observations, noting that the points raised herein are pertinent to such experiments, while they are in line with general relativity. The qualifications for these reservations are that the field of view of the detector must necessarily be narrow and directed, the duration of the event is momentary (but finite) and the speed of detection/observation cannot exceed the speed of light itself.

It is possible to proceed further with basic mathematical justification for the reservations as to the validity of the conclusions attributed to the experiments, indicating observation of super-luminal events. Reference should be made back to the aforementioned 'simple theoretical experiment' in the preceding section where the parallel world-lines in Minkowski's hyperbolic geometry are considered. The motion of the light pulse would be on a geodesic path. In the *H*-plane such paths are described by vertical lines and semi-circles, which would have their centres on the spatial axis in a narrow field-of-view experiment. The spatial axis is denoted by q^1. Therefore, $q^2 = t$ represents the temporal domain. Referring to Fig. 1, the point of observation/detection can be denoted as P on a vertical line given as $q^1(t) = k$, $q^2(t) = t$, $t \succ 0$ (or $q = k$, $q^2 = e^s$ in terms of the arc length). This can be considered as the directed line of detection in a narrow-field observation. Such a line in the *H*-plane represents a geodesic line, with the speed of observation being that of the speed of light. The gain-assisted pulse can be represented by any geodesic line that must necessarily intersect this vertical line for an observation/detection to be made. There is clearly such a family of semi-circles, as shown in Fig. 1. Furthermore, all these world-lines may be parallel, indicating particle motion at different speeds and in fact with motion taking place in either sense. The family of semi-circular lines can be represented by geodesic parameterization: $q^1 = k + r \tanh s$, $q^2 = r \operatorname{sech} s$, $-\infty < s < \infty$. Note that, for a short duration t and at speeds $\boldsymbol{v} \geqslant c$, the position P along the vertical line $q^1 = k$, $q^2 = t$ cannot be expected with certainty. The postulate outlined here for the narrow field of view observation of an event, purported to be at or above the speed of light, is in line with the fundamentals of quantum mechanics, while remaining firmly within the fold of a special case of induced gravity in the general theory of relativity.

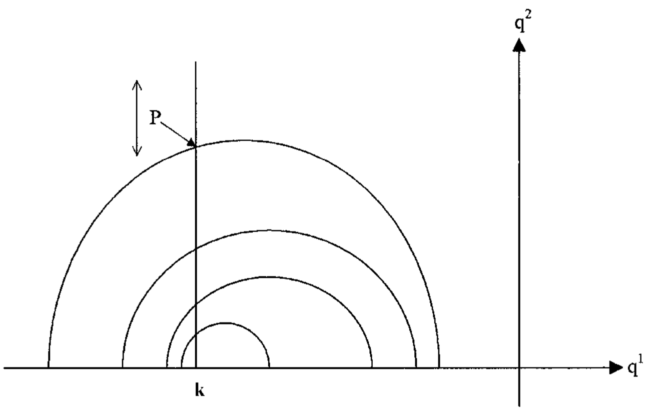

Fig. 1 Observation model of geodesic propagation events

3 THE PRINCIPLE OF INCOMPLETENESS

The reservations expressed with regard to the reported super-luminal observations are, therefore, mainly confined to the multiplicity of likely outcomes, rather than the validity of the actual experiments. In a sense the certainty of a distinct outcome, from an observational rather than the actual event, cannot be assured. In this sense, narrow-field observation of high-speed events represents a problem in both quantum mechanics and general relativity, and paradoxically unites them in this respect. A significance of such reported experiments as that of Wang *et al.* [**12**] has been to bring this observational problem to the fore. At such limits of observation a unidimensional and 'continuous' event becomes indistinct from its possible alternatives. In a sense the event is too 'complete' to be discerned, meaning that its incomplete feature (i.e. its changing nature) is beyond discrimination. This means that the events are discerned by their incomplete or discontinuous nature. In this way, uncertainty at the core of quantum mechanics is not in conflict with the general theory of relativity, but in concordance with it according to the thus established principle of incompleteness.

Now it should also be noted that the curvilinear motion of a pulse of light in the Euclidean geometry of observation can only take place as the result of an accelerated motion. A ray of light can undertake such a motion in a gravitational field. The deflection of the ray, as shown by Einstein, is necessarily very small, even in the presence of significant gravitational fields [**7**]. If the 'observed' super-luminal behaviour of the light pulse is accepted and at the same time there is a commitment to the postulate of relativity (as there should be), the following scenario must be accepted. A gain-assisted pulse of light assumes a curvilinear motion in a field induced by a centrifugal effect, and the induced field deflects the path of travel by an inappreciable amount as viewed by an observer. Furthermore, a real and discrete change of physical description with uniquely defined unidimensional space–time, constituting an event of inappreciable but nevertheless finite duration, has been observed. It is this last condition that casts a shadow of doubt upon the observational findings of such reported experiments, which is in contravention of general relativity. It would mean that such events (observations at or exceeding the limit of observation) in a unidimensional continuous frame can be uniquely described in space–time.

Consider a small particle undertaking a curvilinear path, shown in Fig. 2, with a velocity $\boldsymbol{v}$ with respect to a frame of reference K'. It has already been established that motion of such a particle will be on a curvilinear path

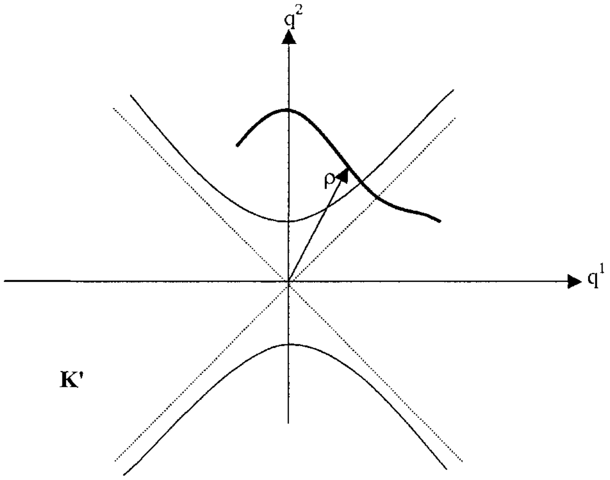

Fig. 2 Description of an event in the Minkowskian hyperbolic geometry

when the particle is subject to a field or if the frame of reference undergoes accelerated motion. The motion of the particle can be described in Minkowski's geometry, with a space-like acceleration of $\boldsymbol{v}^2/\rho$, with ρ being the local path curvature with respect to the hyperbolic frame. It can be noted that as $\rho \to \infty$, the particle acceleration diminishes, and the steady motion becomes rectilinear from the point of view of the observer. This corresponds to the vertical geodesic line, previously described in Fig. 1. When $\boldsymbol{v} \to c$, the event will be viewed as light-like. With $c \to \infty$, the light-like behaviour takes place along q^2 in Fig. 1. At the other extreme, photons of light expand in a waveform in space. The space–time representation of this is a light cone, with $\rho = 0$ everywhere on its surface, rendering also a light-like behaviour. If $\boldsymbol{v} \to c \to \infty$, the same solution is obtained as in the previous case. Significantly, the semi-circles in Fig. 1, representing the geodesic lines for the motion of the light pulse, have tended to a limiting solution for $c \to \infty$. Now note that for an observation to be made, the geodesic lines (one for the propagated light pulse and the other for the line of view of the observer) must necessarily intersect. However, it has already been shown, by almost elementary deduction, that as $c \to \infty$ and for $0 \leqslant \rho \leqslant \infty$, the aforementioned geodesics either have no solution (i.e. they are parallel) or are coincident (i.e. a manifold of solutions in a unidimensional temporal continuum). The event, thus described, is 'continuous' in time and can be detected at a given location, but also at any time. Therefore, a unique and distinct system of coordinates cannot be assumed in the detection of this event. In other words, a unique solution may not be found, as such a 'material point', in the words of Einstein, can be described by a manifold of space–time coordinates.

The observations made thus far are in complete accord with general relativity. The phenomenon described should be assumed to be attributed to the incompleteness of observations, when $c \to \infty$, and particularly with a narrow field of view of an infinitesimal region of the space–time. A special class of events that defy distinct observation points by their nature has thus been established. It can be concluded that observable events possess a discontinuous or incomplete nature within the frame of observation, leading to the aforementioned fundamental principle of incompleteness.

4 SYMMETRIC COMPLETENESS

The same arguments can be extended even further by noting that particle motions on the geodesic semi-circular paths in Fig. 1 can indeed take place in either sense, resulting in the same observational conundrum. The aforementioned arguments apply equally well to the case where $c \to -\infty$, $-\infty \leqslant \rho \leqslant 0$, at least in a mathematical sense. This incidentally shows that symmetric events with an observation period less than or equal to the frequency of observation at the speed of light would also manifest the same problem. It can therefore be established that the principle of incompleteness may be extended to those events that may be regarded as symmetrically complete.

This extension of the principle can help to deal with the uncertainty at the core of quantum physics in a tangible manner. It can be seen that the theoretical experiment described earlier on can in fact easily be regarded as the narrow directed observation of an electron, undertaking its geodesic curvilinear motion with respect to the stationary observer. It should be remembered that the described theoretical experiment led to the establishment of the principle of incompleteness as an axiom of observation and the extension of the same to all symmetrically complete events in the vanishing regions of space–time. The axiom, thus described, is in concordance with the uncertainty in quantum mechanics. In this case it should be noted that the curvilinear wave motion of the electron may be viewed at any given position, but with an indeterminate momentum (i.e. indistinct in a temporal sense). This fact is in line with the aforementioned principle of incompleteness, which has been shown to be in turn in accord with general relativity. Therefore, for unidimensional observations the principle of incompleteness and its extension, symmetrical completeness, apply and the motion of material points irrespective of their particulate or waveform nature concur with both general relativity and quantum mechanics in a unified manner. At a first glance, it appears that the principle of incompleteness is merely the re-statement of the 'point of absurdity' at the fringes of both theories of relativity and quantum mechanics. However, a moment of contemplation would reveal this not to be true, as the said principle was arrived at in describing a ‘real’ physical phenomenon, investigated on numerous occasions in experimental physics.

REFERENCES

1 **Newton, I.** *Philosophie Naturalis Principia Mathematica,* 1687 (Royal Society, London).

2 **Einstein, A.** Die Grundlage der allgemeinen Relativitätstheorie. *Annln der Physik*, 1916, **49**.

3 **Lorentz, H. A.** *Versuch einer Theorie der Elektrischen und optischen Erscheinungen in bewegten Körpern,* 1895 (Leiden).

4 **Einstein, A.** *Relativity: The Special and the General Theory*, 1961, Ch. 11 and Appendix 2 (Wings Books, New York).

5 **Minkowski, H.** Space and time. An address to the 80th Assembly of German Natural Scientists and Physicians, Köln, Germany; Translation in *The Principle of Relativity (Lorentz, Einstein, Weyl and Minkowski)*, 1952 (Dover Publications).

6 **Einstein, A.** Zur elektrodynamik bewegter Körper. *Annln der Physik*, 1905, **17**.

7 **Einstein, A.** Uber den Einfluss der Schwerkraft auf die Ausbreitung des Lichtes. *Annln der Physik*, 1911, **35**.

8 **Weyl, H.** *Gravitation und Elektriticitat,* 1918 (Sitzungsberichte der Preussischen Akad. D. Wissenschaften).

9 **Einstein, A.** On the generalized theory of gravitation. *Scient. Am.*, 1950, **182**(4).

10 **Bolda, E., Garrison, J. C.** and **Chiao, R. Y.** Optical pulse propagation at negative group velocities due to a nearby gain line. *Phys. Rev.*, 1994, **A49**.

11 **Akulshin, A. M., Barreiro, S.** and **Lezama, A.** Steep anomalous dispersion in coherently prepared Rb vapor. *Phys. Rev. Lett.*, 1999, **83**.

12 **Wang, L. J., Kuzmich, A.** and **Dogariu, A.** Gain-assisted superluminal light propagation. *Nature Lett.*, 2000, **406**.

13 **Einstein, A.** *Relativity: The Special and the General Theory*, 1961, Ch. 22 (Wings Books, New York).

14 **Einstein, A.** *Relativity: The Special and the General Theory*, 1961, Ch. 27 (Wings Books, New York).

