THE STRONG EQUIVALENCE PRINCIPLE AND ITS VIOLATION

V. M. Canuto and I. Goldman
NASA, Goddard Institute for Space Studies
2880 Broadway, New York, New York 10025

Abstract. In this paper, we discuss theoretical and observational aspects of an SEP violation. We present a two-times theory as a possible framework to handle an SEP violation and summarize the tests performed to check the compatibility of such violation with a host of data ranging from nucleosynthesis to geophysics. We also discuss the dynamical equations needed to analyze radar ranging data to reveal an SEP violation and in particular the method employed by Shapiro and Reasenberg.

I. The Strong Equivalence Principle, SEP

The Strong Equivalence Principle[1], demanding that local physics (described in a local lorentzian coordinate system) be the same anywhere and at anytime in the universe, assumes that the outcome of local experiments is independent of cosmological influence. This is equivalent to complete decoupling of local physics from the global structure of the universe, which can therefore influence local systems by determining only boundary conditions, i.e. the background space-time structure at large distances. In this sense, the SEP allows only for a very restricted realization of "Mach's principle".

If the SEP holds, the different dynamical physical fields of nature must be coupled in a way that is independent of cosmology, i.e. by coupling coefficients that are constant. On the other hand, if the SEP is violated on a cosmological time scale, there must exist an underlying dynamics connecting the cosmological structure with local physics. Before discussing how to construct a theoretical framework that incorporates an SEP violation, we shall consider an aspect of an SEP violation that has direct observational consequences.

Consider Einstein's equations and electrodynamics. Both dynamical theories are not invariant under scale transformations. Because of this property, each of them defines its own unit of time (i.e. a clock) which serves as a basis for an intrinsic system of units. (On the contrary, equations that are scale invariant, like Maxwell equations for photons, cannot give rise to a clock). Since there exist several non-scale invariant dynamics in nature, each defining its own clock and units, the question arises as to how do the different clocks relate to one another.

The SEP demands that all clocks be equivalent, i.e. that the ratio

G. O. Abell and G. Chincarini (eds.), Early Evolution of the Universe and Its Present Structure, 485–492.

of the intrinsic units of time be constant. On the other hand, if the SEP is violated, the couplings coefficients among different fields are cosmology dependent and so are the ratios of the periods of the various clocks, i.e. one faces the possibility of non-equivalent systems of units.

A minimal form of an SEP violation is one in which all microscopic units are equivalent (e.g. atomic clocks, nuclear clocks and weak interaction clocks have constant ratios of their periods), but gravitational units are different, a choice consistent with the fact that gravity is the only interaction important on a cosmological scale. Taking into consideration the observed homogeneity of the universe, one may restrict the SEP violation to depend only on the time coordinate of the comoving cosmological coordinate system. One may therefore write

$$ds_E = \beta_a(t)ds_a \,, \qquad \beta_a(t_o) = 1 \qquad (1)$$

where ds_E is a physical space-time line element measured with gravitational clocks, (e.g. a planet revolving a star), while ds_a is the corresponding line element measured with atomic clocks. The function β_a epitomizes our ignorance of how the two different dynamics couple: a constant β_a implies that the two systems of units are equivalent, a non-null $\dot{\beta}_a$ implies a violation of the SEP. (In what follows, EU and AU will be used to denote gravitational or Einstein units and atomic units respectively).

Presently available lunar data[2] suggest $\dot{\beta}_a \sim 10^{-11}$ yr^{-1} and radar ranging data[3] to the inner planets provide an upper limit $|\dot{\beta}_a| < 10^{-10} yr^{-1}$.

II. Frameworks that Incorporate an SEP Violation

Let us turn to the formulation of a framework that allows for an SEP violation of the order of H_o. One may follow two alternatives.

1) Local approach. β_a is assumed to be a dynamical field coupled locally to matter sources. This approach, adopted for example by Brans-Dicke[4], has two limitations:

a) Since solar system data[5] require $\omega > 500$, it follows that $\dot{\beta}_a < 10^{-13}$ yr^{-1} since the BD theory predicts[6] $\dot{\beta}_a \sim H_o/(1 + \omega)$. Therefore, within a BD-type of theory, it is not possible to accommodate both solar system data as well as $\dot{\beta}_a \sim H_o$.

b) From a physical point of view, one would expect in gravitational units, the trajectories of macroscopic objects to be described by geodesics i.e that gravity be described by a metric theory[1] in the system of units defined by its own intrinsic clock. However, this is not so in the BD framework, where gravity is described by a metric theory not in gravitational units but rather in atomic units, which however belong to a different dynamics.

2) Global approach[7]. Since there is no a priori reason why one should restrict an SEP violation to be described by a local space-time field, it is conceivable that β_a in eq. (1) is not coupled to local matter sources but rather represents an average over the global distribution of matter in the universe at any given time. One may of course try to construct a complete theory to predict how β_a is related to the global structure of the universe. However, a great deal can be learned even from the use of a framework in which β_a is treated phenomenologically, i.e. its dynamics is not specified; β_a enters such a framework as an external quantity

much as for example viscosity and diffusion coefficients enter the classical hydrodynamic equations. Such phenomenological framework has been meaningfully employed to limit the allowed degree of violation of the SEP using geological, astrophysical and cosmological data[8,11,12,13].

III. Breaking the Symmetry between Atomic and Gravitational Clocks.

The main objective of any theory that tries to incorporate an SEP violation is that of achieving a symmetry breaking between AU and EU, which, according to the SEP, are indistinguishable systems of units. To be more specific, one must construct gravitational and atomic clocks whose periods P and p respectively, satisfy the following symmetry-breaking conditions

$$\text{EU:} \qquad P_E = \text{constant} \qquad p_E \sim \beta_a \tag{2}$$

$$\text{AU:} \qquad p_a = \text{constant} \qquad P_a \sim \beta_a^{-1} \tag{3}$$

The construction of such gravitational and atomic clocks was presented in [7]. The most crucial ingredients are the equations of motion for micro and macroscopic objects, which must clearly be different from one another in order for conditions (2) and (3) to be satisfied. To achieve our goal, it is convenient to introduce a mathematical formalism whereby equations are first expressed in a general system of units. This operation does not introduce any physics and is merely a convenient mathematical tool. In [7], it was shown that in general units, the equations of motion of an object with mass μ (macro, M or microscopic, m) is given by ($\Delta^{\alpha\nu} = u^{\alpha}u^{\nu} - g^{\alpha\nu}$).

$$u^{\alpha}_{;\nu}\, u^{\nu} + \frac{(\mu\beta^{2-g})_{,\nu}}{(\mu\beta^{2-g})}\,\Delta^{\alpha\nu} = \frac{e}{\mu}\, u^{\nu}F^{\alpha}_{\nu} \tag{4}$$

Here, the function β defines the general system of units to which we have arrived starting from EU, taken to be our fiducial system, and performing the transformations

$$L_E = \beta L, \qquad G_E = G\beta^{g}, \qquad M_E = M\beta^{1-g} \tag{5}$$

where g is a constant number. The details of the mathematical solutions of (4) were presented in [7]. Here, we shall follow a different approach. Taking the ϕ component of (4), we obtain for a circular orbit, the relation

$$\Pi/(r^2\mu\beta^{2-g}) = \text{constant} \tag{6}$$

where r is the radius distance and Π the period. Eq (6) is valid in any units and for any mass μ. Let us now consider the following argument. In AU, the period of an atomic clock and the dimension of it, must clearly be constant and so must the ratio r/Π. The same argument must hold true in EU for a gravitational clock. We therefore conclude that the dimensionless ratio r/Π must remain constant for either clock in <u>any</u> units. This changes eq. (6) to

$$\Pi \sim 1/(\mu\beta^{2-g}) \tag{7}$$

Applying (7) to a macroscopic object ($\mu \equiv M$, $\Pi \equiv P$) and a microscopic one ($\mu = m$, $\Pi = p$) and imposing (2) - (3), we obtain

EU, $\beta \equiv 1$	$P_E \sim M_E^{-1}$	i.e.	M_E = constant	(8)
	$p_E \sim m_E^{-1}$		$m_E \sim \beta_a^{-1}$	
AU, $\beta \equiv \beta_a$	$p_a \sim m_a^{-1}\beta_a^{g-2}$		$m_a \beta_a^{2-g}$ = constant	
	$P_a \sim M_a^{-1}\beta_a^{g-2}$		$M_a \beta_a^{1-g}$ = constant	

To proceed further, we must define the AU and EU more precisely. It is physically reasonable to characterize these units and the nature of the physical clocks defining them, by requiring that microscopic masses be constant in AU and that macroscopic masses be constant in EU, i.e. the results (8) will be supplemented by the two conditions (the second of (9) is already contained in (8))

$$\text{AU:} \quad m_a = \text{constant} \qquad\qquad \text{EU:} \quad M_E = \text{constant} \tag{9}$$

which, together with (8) imply that the parameter g must be 2. We therefore conclude that the symmetry breaking conditions (2) and (3) can be met provided the following conditions are satisfied

AU:	$\beta = \beta_a$,	m_a, e_a = const.,	$M_a \sim \beta_a$,	$G_a \sim \beta_a^{-2}$	(10)
EU:	$\beta = 1$,	$m_E \sim \beta_a^{-1}$,	e_E = const.	M_E, G_E = const.	

It is important to stress the fundamental role played by (9) in achieving a symmetry breaking.

IV. Baryons and Photons

From (10), it follows that the number of particles N must vary like

$$N \sim \beta_a \tag{11}$$

independently of the system of units, as indeed expected for N is a pure number. We therefore conclude that, within the present framework, a violation of the SEP can be realized provided the number of particles varies with time. This in turn implies that in AU, many-body thermodynamic relations will also depend on β_a[8], a result that can be visualized in terms of a "viscosity-like" role played by β_a. In fact, while at the one body level, say Dirac equation, β_a does not appear, it does so at the many-body level, as a consequence of the fact that the conservation laws are changed, since in AU the Einstein equations have changed[9]. It is in fact only in EU that Einstein equations are assumed to be unaltered.

As for photons, it has recently been shown[8] that the number of photons N_γ is constant, independently of the value of g. This, together with the value g = 2, are sufficient to allow the construction of a photon theory in which all the photon relations remain unaltered, in spite of a possible violation of the SEP.

This result leads us to suggest a scenario in which β_a was constant during the radiation dominated era. The argument is as follows. The function β_a represents a cosmological influence on local physics. The dynamics of β_a is determined by the global structure of the universe whose dynamics is governed by the energy density (and pressure) of matter and radiation. Since matter is affected by β_a, while radiation is not, one may expect that during the matter dominated era the rate of variation of β_a to be comparable with the rate of expansion $\dot\beta_a/\beta_a \sim 1/t$, i.e. $\beta_a \sim t^{-n}$, $n \sim 1$. However, during the radiation era, such a variation must have been considerably smaller $\dot\beta_a/\beta_a << 1/t$, since the dynamical role played by matter was negligible. This is in agreement with the results of a recent numerical study of nucleosynthesis[10] showing that in order to explain the present abundance of light elements, $\dot{\beta}_a/\beta_a \ll 1/t$.

V. Compatibility Tests

The compatibility of an SEP violation has been tested against a host of geophysical, astrophysical and cosmological data. The global result is that a variation of the type $\beta_a \sim t^{1/2}$ during matter dominated era and $\beta_a \sim$ constant during the radiation era, is consistent with the following set of tests: 1. Nucleosynthesis[8], 2. The 3K black-body radiation[8], 3. The m vs. z relation for QSOs, radio galaxies and elliptical galaxies[11], 4. Angular diameters for radio galaxies[11], 5. The isophotal angles for optical galaxies[11], 6. The N(S) vs. S relation for radio galaxies[12], 7. Ages of stars and globular clusters[8], 8. The early temperature of the Earth[8], 9. The Earth's paleoradius[13]

VI. Testing the SEP by Radar Ranging.

A planet orbiting the Sun is a gravitational clock. Therefore, by tracking the motion of such a planet by means of atomic clocks, one compares directly gravitational vs. atomic clocks and can therefore determine the value of β_a. The observations consist of measuring with precision atomic clocks the round trip time of radio beams sent from earth and bounced back by the planet surface. A series of such measurements have been analyzed over the years by Shapiro and Reasenberg[3]. Since these measurements and their analysis predated by several years the appearance of the two-times theory presented here, it may be instructive to analyze the mathematical framework employed by SR vis á vis the present framework.

Let us begin with Newton's equations in their standard form, i.e. in dynamical units (subindex E)

$$\frac{d^2\vec{x}_E}{dt_E^2} = -\frac{G_o M}{x_E^3}\vec{x}_E \tag{12}$$

where G_o is a constant. With the solutions of these equations, one can construct the "range" $\rho_E(t_E)$, i.e. half the round trip time,

$$\rho_E(t_E) = (x_{1E}^2 + x_{2E}^2 - 2x_{1E}x_{2E}\cos\theta_{12})^{1/2} \tag{13}$$

which is a quantity of direct observational interest. Let us now perform a transformation from dynamical units x_E and t_E to the corresponding atomic ones x and t, by writing in general

$$x_E(t_E) = \Lambda_1(t)x(t), \qquad dt_E = \Lambda_2(t)dt, \tag{14}$$
$$\Lambda_k(t) = 1 + \lambda_k t + \ldots \qquad t_E = t + 1/2\lambda_2 t^2 + \ldots$$

where $t = t_E = 0$ denote the initial time of the ranging data Substitution in (12) yields (up to first order in λ_1, λ_2)

$$\frac{d^2x^k}{dt^2} = -\frac{G_o M}{x^3}x^k - (2\lambda_2 - 3\lambda_1)\frac{G_o M}{x^3}x^k t - (2\lambda_1 - \lambda_2)\frac{dx^k}{dt} \tag{15}$$

At the same time, the perturbation in the range $\delta\rho \equiv \rho(t) - \rho_E(t)$ can

https://doi.org/10.1017/S0074180900039589 Published online by Cambridge University Press

easily be evaluated using $\rho(t) = \Lambda_1^{-1}(t)\rho_E(t_E)$. The result is

$$\delta\rho = -\lambda_1 t\rho_E(t) + \frac{1}{2}\lambda_2 t^2\dot{\rho}_E(t) \tag{16}$$

Defining the finite differences $\Delta t \equiv t_2 - t_1$, $\Delta t_E = t_{2E} - t_{1E}$, as well as $2t = t_1 + t_2$, we obtain

$$\Delta x_E = (1 + \lambda_1 t)\Delta x, \qquad \Delta t_E = (1 + \lambda_2 t)\Delta t \tag{17}$$

which give the transformation laws for finite lengths and time intervals. With these results, we shall investigate two choices for λ_1 and λ_2.

A) $\lambda_1 = \lambda_2 \equiv \lambda_A$. This corresponds to requiring the same scaling for lengths and times, i.e. velocities remain invariant under scaling, whereas angular momenta per unit mass, J, do not, since the dimensions of J are $\sim$ (length)^2/time. Eqs. (15) and (16) now become

$$\frac{d^2x^k}{dt^2} = -\frac{G_oM}{x^3}x^k + \lambda_A\frac{G_oM}{x^3}x^k t - \lambda_A\frac{dx^k}{dt} \tag{18}$$

$$\delta\rho = -\lambda_A t\rho_E(t) + \frac{1}{2}\lambda_A t^2\dot{\rho}_E(t)$$

The first of (18) is precisely the form of Newton's equation derived from the ingeodesic equations in atomic units, proposed several years ago[9]. To make the identification complete, we must call $\lambda_A \equiv \dot{\beta}_a/\beta_a$. Because of the presence of the "viscosity term" $\sim dx^k/dt$ in (18), the angular momentum J cannot be conserved in atomic units.

B) $\lambda_1 = 1/2\ \lambda_2 \equiv \lambda_B$. This corresponds to requiring that J remains constant under scaling, i.e. that lengths and times scales differently. With this choice eqs. (15) and (16) become

$$\frac{d^2x^k}{dt^2} = -\frac{G_oM}{x^3}x^k - \lambda_B\frac{G_oM}{x^3}x^k t \tag{19}$$

$$\delta\rho = -\lambda_B t\rho_E(t) + \lambda_B t^2\rho_E(t)$$

Let us note that the viscosity term has disappeared and therefore J may be constant. The first of Eqs. (19) are exactly the relations employed by SR in their analysis.

Now the question is: is the choice $\lambda_1 = 1/2\ \lambda_2 \equiv \lambda_B$ acceptable? While there may be physical situations in which times and lengths scale differently, in the experimental situation at hand, ranging from planets, the units of length <u>and</u> time are not independent, distances being in fact defined as round trip times. It follow that only the choice $\lambda_1 = \lambda_2$ is acceptable, i.e. eqs. (19) are not compatible with the present two-times theory <u>and</u> the accompanying assumption that in gravitational units Newton's equation holds unchanged.

Given the unphysical nature of the scaling $\lambda_1 = \lambda_2/2$, what can one say about eqs. (19) from the theoretical and experimental point of view?

Let us first consider the theoretical aspect. One may not want to adopt the present two-times theory and the assumption that eq. (12) is valid in EU. In that case, the derivation of (19) through scaling would no longer exist. For example, the Brans-Dicke-type theories <u>do not adopt eq. (12) in EU</u>, rather they demand that in AU Newton's law should read

$$\frac{d^2 \vec{x}}{dt^2} = - \frac{G(t)M}{x^3} \vec{x} \tag{20}$$

i.e., (19) with $\lambda_B = (\dot{G}/G)_0$. Therefore, the starting equation in the SR analysis can be said to belong to the Brans-Dicke-like theoretical frameworks. However, because of the arguments given in IIa, one may expect that $(\dot{G}/G)_0 \ll H_0$.

Let us now consider the experimental situation. Suppose that a particular set of ranging data is used for which the quadratic term in $\delta\rho$ dominates over the linear one (e.g. the Mercury data). If so, eq. (18), representative of the present two times theory, and (19), representative of the BD-type theories, give

$$\delta\rho \simeq \frac{1}{2} \dot{\rho}_E \lambda_A t^2, \qquad \delta\rho \simeq \dot{\rho}_E \lambda_B t^2 \tag{21}$$

i.e. two expressions which are equally good to detect whether the coefficient of the quadratic term (independently of how we call it) is non-null, a result which would signal an SEP violation.

This conclusion is in accord with the claim made by SR sometime ago[14] (on the basis of a similar but less general analysis), that the goal of first "smoking out" the presence of an SEP violation could be achieved in a "theory independent" fashion.

The differentiation between the present two times theory and the BD-like theories can be carried out if one employs the Viking data, since in that case the short time interval (since 1976) together with the long period of Mars, make the linear term in $\delta\rho$ important. The Viking data are at present analyzed by SR with this aspect in mind; it is hoped that the results will be soon forthcoming.

Acknowledgements. It is a pleasure to thank Drs. I. I. Shapiro and R. D. Reasenberg for interesting discussions.

References

1. Will, C. M., in "General Relativity", (eds. Hawking, S. W. and Israel, W. (Cambridge Univ. Press, New York, 1979). See also: Thorne, K. S., Lee, D. L. and Lightman, A. P., Phys. Rev. D7, 3563 (1973).
2. For a summary of the results see: Van Flandern, T. C., Ap. J. 248, 813 (1981).
3. Reasenberg, R. D. and Shapiro, I. I., in "On the Measurements of Cosmological Variations of G", (ed. by L. Halpern), (Univ. Presses of Florida, Gainsville, Fla. 1978).
4. Dicke, R. H., "The Theoretical Significance of Experimental Relativity", (Gordon and Breach, Science Publ., N. Y., N. Y., 1964).
5. Reasenberg, R. D., Shapiro, I. I., Macneil, P. E., Goldstein, R. B., Breidenthal, J. C., Brenkle, J. P., Cain, D. L., Kaufman, T. M., Konarek, T. A., and Zygielbaum, A. I., Ap. J. Lett. 234, L219 (1979).
6. Weinberg, S., "Gravitation and Cosmology: principles and applications of the General Theory of Relativity", pp. 628-629 (Wiley, N.Y. 1972).
7. Canuto, V. M. and Goldman, I., Nature 296, 709 (1982).
8. Canuto, V. M. and Goldman, I., 1982 Nature (submitted).
9. Canuto, V. M., Adams, P.J., Hsieh, S.-H. and Tsiang, E., Phys. Rev. D16, 1643 (1977).

10. Rothman, T. and Matzner, R., Ap. J. 257, 450 (1982).
11. Canuto, V. M., Hsieh, S.-H. and Owen, J.R., Ap. J. Suppl. 41, 263 (1979).
12. Canuto, V. M. and Owen, J. R., Ap. J. 41, 301 (1979).
13. Canuto, V. M., Nature 290, 739 (1981).
14. Shapiro, I. I. and Reasenberg, R. D., private communication (1982).

Discussion

Spyrou: Could you please comment on the systematic post-Newtonian uncertainties and errors you mentioned concerning the solar system?

Goldman: In analyzing the planetary radar ranging data, one employs a post-Newtonian model of the gravitational dynamics of the solar system. The model contains parameters such as solar and planetary masses, planetary orbital elements, post-Newtonian gravity parameters, and in our case also, $\dot{\beta}_a$. These parameters are fitted so as to minimize the residuals between the computed and measured range, as a function of time.

Prior to the use of the Viking Mars lander, the main source for uncertainty in $\dot{\beta}_a$ determination came from unknown topography of the planetary surfaces from which the radar waves were reflected. With the Viking lander there is a systemic uncertainty due to the modeling of the asteroid belt. Another limitation to the accuracy stems from possible correlations of $\dot{\beta}_a$ with other parameters; e.g., an error in Mars' semimajor axis can, for a short observing time, mimic a $\dot{\beta}_a$ effect. Nevertheless, it is estimated that the Viking data will provide an accuracy of $\sim$ 1 part in 10^{11} yr^{-1}. good enough to detect a cosmological effect of magnitude $\sim H_o$.

https://doi.org/10.1017/S0074180900039589 Published online by Cambridge University Press