

Varying G

V. Canuto* and S.-H. Hsieh† *NASA Goddard Institute for Space Studies,
Goddard Space Flight Center New York, New York 10025, USA*

J. R. Owen *Department of Physics, CCNY, New York, USA*

Received 1979 February 28; in original form 1978 July 26

Summary. We critically analyse the problem of the variation of the gravitational constant with cosmological time.

Since Einstein's equation does not allow G to vary on any time-scale, no observational data can be analysed within the context of the standard theory. The recently proposed scale covariant theory, which allows (but does not demand) G to vary, and which has been shown to have passed several standard cosmological tests is employed to discuss some recent non-null observational results which indicate a time variation of G .

1 Introduction

McCrea (1978) has pointed out that there is no sense in which the standard general relativity (GR) can admit a variable 'gravitational constant'. It was also suggested (McCrea 1974) that with improved observational confirmation of Einstein's results, thus establishing the correctness of his theory of gravitation, any variation of the gravitational constant can be ruled out by inference. While we agree with the former remark, we do take issue with the latter inference. In the present paper, we shall explain how Einstein's *theory of gravitation* can be reconciled with a varying gravitational constant.

The value of any *dimensional* physical constant depends on the units one employs. In a space-time theory such as GR, the fundamental unit is a length, which is provided by some measuring procedure. However, *any measuring instrument, being a physical system itself, must obey certain dynamical laws*. Thus, for example, if we use the distance between orbiting gravitational bodies as a reference, we would have a gravitational unit (or Einstein unit) of length. On the other hand, if atomic instruments, whose governing physical law is quantum electrodynamics, are used, we have an atomic unit of length. *A priori*, there is no reason to believe that the two units of length must be a constant multiple of each other. Consequently, when the 'gravitational constant' is a constant in one system of unit, it is not necessarily a constant in the other system of unit.

* Also with the Department of Physics, CCNY, New York.

† NAS–NRC Research Associate.

The idea can be better illustrated. If we write ds_A, ds_E for the line elements as measured in atomic and gravitational units respectively, we would have in general

$$ds_E = \beta(x) ds_A \quad (1)$$

it follows then that all dimensional physical quantities in the two respective units are similarly related:

$$Q_E = \beta^\pi Q_A \quad (2)$$

Where Q_E and Q_A may be scalars, vectors or tensors and the exponent π is given by the dimensions of Q . In particular, the 'gravitational constants' in the two units are related by

$$G_E = \beta_E^\pi G_A. \quad (3)$$

We wish to note here a subtle difference between our use of the terms 'general relativity' and 'Einstein's theory of Gravitation'. The former assumes the strong principle of equivalence which dictates that β must be strictly constant. If one assumes only the weak equivalence principle, Einstein's theory of gravitation remains intact, and β can in general be a function of space-time. In the geometrical framework of Einstein's theory of gravitation, the Bianchi identities along with the conservation of energy and momentum demand a constant G_E , which is a proportionality factor between the geometrical Einstein tensor and the energy-momentum tensor. We note that it is G_E , and not the 'gravitational constant' in any other units which is required to be a constant, because Einstein's theory governs the dynamics of gravitational phenomena only, and it provides a geometrodynamical unit of length. Hence Einstein's field equations must be understood as written in Einstein units. In standard GR, people use these equations as though they are also valid in atomic units. It should be recognized, however, that this amounts to making an extra assumption, namely, that the scaling function $\beta(x)$ in equation (1) is a constant. Thus, *one goes beyond the realm of gravitational physics and stipulates a specific relation between gravitational dynamics and atomic dynamics*. If the gravitational interaction strength changes relative to electromagnetic interaction, we must expect β to be varying and therefore according to equation (3) G_A must also be varying. It is in this sense that we can accommodate and interpret a varying 'gravitational constant'. From this viewpoint of scaling between two kinds of dynamical units, G must be expressible as a functional of β as is clear from equation (3). However, it is important to note that for a complete determination of the variation of G , one must know not only the variation of β but also the value of π_g , as we shall explain in more detail below.

With this understanding, it becomes evident that observational confirmation of Einstein's results in purely gravitational experiments can be compatible with experiments which purport to measure the variation of G , provided in the latter, atomic units are used. This in fact is what some observers have been attempting to do in the past few years: measuring the non-gravitational variation of the orbital period of the Moon in terms of atomic time. Whatever theoretical prejudice one may have for preferring a null result in the above experiments, one should keep an open mind and allow for the possibility of a non-null result.

Historically, Milne had long ago anticipated the possibility of equation (1). When Dirac introduced his Large Numbers Hypothesis (LNH) and proposed a varying G , he also had equations (1)–(3) in mind. Unfortunately, when people study the effects of a varying gravitational constant, Newton's or Einstein's dynamical equations are used with only the modification that G_E is allowed to be a variable. As was pointed out, this is a logically inconsistent procedure.

A consistent formulation of the gravitational equations in non-gravitational units has been given in an earlier paper (Canuto *et al.* 1977) in which we developed the ideas outlined

above: gravitational dynamics remains unchanged. When described in atomic units, the dynamical equations are obtained by conformal transformation, as required by equation (1), from the corresponding ones in standard GR. Clearly the *a priori* undetermined function β would appear in these equations. If, e.g. the observational results are analysed with these equations, β can be determined observationally. On the other hand, one could apply theoretical considerations such as Dirac's LNH to fix the functional form of β . In this case, the conformally transformed dynamical equations are fully deterministic, and one can predict results from experiments using atomic units. In this manner, one can have a valid observational check on Dirac's ideas.

In the next section, we briefly review the framework of scale covariant gravitation introduced in an earlier paper and then illustrate the types of theoretical considerations one can use for the determination of β . In Section 3, we shall show in some detail how the scale covariant framework can be used to interpret and analyse data from atomic measurements of gravitational phenomena, thus giving a description of the observational determination of β .

2 Theoretical determination of β

With the premise that Einstein's theory correctly describes the gravitational phenomena, the following field equations and geodesic equations are assumed valid:

$$G_{\mu\nu}^{(E)} = -8\pi G_E T^{\mu\nu (E)} \quad (4a)$$

$$\frac{d^2 x^\mu}{d\lambda_E^2} + \Gamma_{\nu\rho}^\mu \frac{dx^\nu}{d\lambda_E} \frac{dx^\rho}{d\lambda_E} = 0 \quad (5a)$$

where $d\lambda_E$ is the differential affine parameter, identifiable with the differential path length for non-null geodesics. It is therefore a length measured in Einstein units. We adopt the convention that the coordinate differential is dimensionless. Thus, given that

$$ds_E^2 = g_{\mu\nu}^{(E)} dx^\mu dx^\nu; \quad ds_A^2 = g_{\mu\nu}^{(A)} dx^\mu dx^\nu \quad (6)$$

equation (1) implies

$$g_{\mu\nu}^{(E)} = \beta^2 g_{\mu\nu}^{(A)}, \quad (7)$$

which can be readily recognized as a conformal transformation. The geometric parts of equations (4a) and (5a) are easily transformed. With the further assumption that the right side of equation (4a) is form invariant (see Canuto *et al.* 1977, for details), we get

$$G_{\mu\nu}^{(A)} + f_{\mu\nu} = -8\pi G_A T_{\mu\nu}^{(A)} \quad (4b)$$

with

$$f_{\mu\nu} = 2 \frac{\beta_{\mu;\nu}}{\beta} - 4 \frac{\beta_\mu \beta_\nu}{\beta^2} - g_{\mu\nu}^{(A)} \left(2 \frac{\beta_{;\rho}^\rho}{\beta} - \frac{\beta^\rho \beta_\rho}{\beta^2} \right)$$

and

$$\frac{d^2 x^\mu}{d\lambda_A^2} + \Gamma_{\nu\rho}^\mu \frac{dx^\nu}{d\lambda_A} \frac{dx^\rho}{d\lambda_A} + \frac{\beta_\nu}{\beta} \left(\frac{dx^\mu}{d\lambda_A} \frac{dx^\nu}{d\lambda_A} - g^{\mu\nu} \frac{dx^\rho}{d\lambda_A} \frac{dx_\rho}{d\lambda_A} \right) = 0 \quad (5b)$$

with

$$\beta_\nu \equiv \beta_{,\nu}$$

We note that $G_{\mu\nu}^{(E)}$ and $G_{\mu\nu}^{(A)}$ are the Einstein tensors constructed from $g_{\mu\nu}^{(E)}$ and $g_{\mu\nu}^{(A)}$ respectively. Covariant differentiation in equation (4b) is defined with respect to $g_{\mu\nu}^{(A)}$. Likewise, the Γ 's in equations (5a) and (5b) are the Christoffel symbols constructed from $g_{\mu\nu}^{(E)}$ and $g_{\mu\nu}^{(A)}$ respectively.

It can be shown from equations (4b), using equation (3) that

$$T^{\mu\nu}{}_{;\nu} = (\pi_g - 2) T^{\mu\nu} (\ln \beta)_{,\nu} + (\ln \beta)^\mu T^\nu{}_\nu. \quad (8)$$

(Henceforth we drop the index A for atomic units). Hence the energy momentum conservation law in atomic units must be modified as a consequence of our assumptions. In the same spirit the modified baryon number conservation equation reads

$$(nu^\mu)_{;\mu} + (1 - \pi_g) n \frac{\dot{\beta}}{\beta} = 0 \quad (9)$$

where n is the baryon number density. The generalization given by equation (9) is compatible with equation (8) if we assume a perfect fluid form for the energy momentum tensor with $p = 0$. However, the validity of equation (9) is independent of this assumption.

If one writes ($p = \Gamma\rho$)

$$T^{\mu\nu} = (\rho + p) u^\mu u^\nu - pg^{\mu\nu} \quad (10)$$

and integrate equations (8) and (9) over a volume element V , it can be shown that equation (8) yields

$$\rho V^{(1+\Gamma)} \sim G^{-1} \beta^{-(1+3\Gamma)} \quad (11)$$

and equation (9) yields

$$nV \sim G^{-1} \beta^{-1} \quad (12)$$

where Γ has been assumed constant, and we have again made use of equation (3). Equations (11) and (12) give the allowed variation of total energy (in atomic units) and particle number within a co-moving volume. Obviously, the results of the standard theory are recovered if we set G and β equal to constants.

With the aid of equations (11) and (12), we can indicate how cosmological considerations such as Dirac's LNH can be used for the determination of β . (A more detailed discussion and a brief review of LNH can be found in Canuto *et al.* 1977.) We assume homogeneous cosmological models so that only functional dependence with respect to cosmic time t (in atomic units) needs to be considered.

(1) First gauge. If we assume (Dirac 1974)

$$G \sim \frac{1}{t}, \quad N \equiv nV \sim t^2, \quad (13)$$

equations (11) and (12) imply that

$$\beta \sim \frac{1}{t}, \quad \pi_g = -1. \quad (14)$$

(b) Second gauge (Dirac 1938). If we assume

$$G \sim \frac{1}{t}, \quad N \equiv nV \sim t^0 = \text{constant.} \quad (15)$$

equations (12), (15) and (3) then yield

$$\beta \sim t, \quad \pi_g = +1. \quad (16)$$

(c) From considerations of the spectrum of the background radiation (Canuto & Hsieh 1978), one may prefer to impose the auxiliary condition that radiation be adiabatically conserved. Thus with $\Gamma = 1/3$, we find from equation (11) that

$$G\beta^2 = \text{constant.} \quad (17)$$

Together with equations (13) and (3) we obtain

$$\beta \sim t^{1/2}, \quad \pi_g = 2. \quad (18)$$

The common assumption among the above three cases is the Dirac hypothesis that the gravitational constant decreases as the inverse of the cosmological epoch. The hypothesis of matter creation which is incorporated only in case (a), can be modified. Instead of specifying the time variation of the particle number within a co-moving volume, one can stipulate that the number of particles within the observer horizon increases as the square of the epoch:

$$N_H = \frac{\rho_m}{m} (ct)^3 \sim t^2. \quad (19)$$

We shall now show that for certain cosmological models, the assumption (19) leads back to the hypothesis on the variation of G .

For that purpose, we note first that equation (11) with $\Gamma = 0$ gives the variation of mass density

$$\rho_m \sim R^{-3} G^{-1} \beta^{-1} \quad (20)$$

where we have replaced the volume by R^3 , R being the scale factor in the Robertson–Walker metric. Next we note that equations (4a), (4b) are conformal transformation of each other and so must be their solutions. Hence, using equation (7) it can be shown that

$$R_E(t_E) = \beta(t) R(t) \quad (21)$$

where $R_E(t_E)$ is the scale factor of the R–W metric in Einstein units satisfying equation (4a). It is well known that in the matter-dominated era, for $k = 0$, we have

$$R_E(t_E) \sim t_E^{2/3}. \quad (22)$$

Using equation (1), equation (21) can be rewritten as

$$R(t) \sim \beta^{-1}(t) \left\{ \int_0^t \beta(t') dt' \right\}^{2/3}. \quad (23)$$

With equations (20) and (23), equation (19) becomes

$$N_H \sim \beta^2 t^3 G^{-1} \left(\int_0^t \beta(t') dt' \right)^{-2} \sim t^2$$

which, for simple power laws $\beta \sim t^p$, ($p \neq 1$) yields

$$N_H \sim \frac{t}{G} \sim t^2,$$

i.e. $G \sim t^{-1}$, the hypothesis on the variation of G (Dirac 1938).

3 Observational determination of β

From standard theory, it is easy to show that

$$G_E M_E = \text{constant} \quad (24)$$

$$\frac{\dot{n}_E}{n_E} = 2 \frac{(G_E M_E)'}{G_E M_E} \quad (25)$$

$$\frac{\dot{R}_E}{R_E} = - \frac{(G_E M_E)'}{G_E M_E} \quad (26)$$

where R_E is the radius of the planetary orbit and $n_E = 2\pi/T_E$, with T_E denoting the period of revolution. Without a clear distinction between different dynamical clocks, it has been tempting for observers to interpret the variation of (\dot{n}/n) and (\dot{R}/R) , over and above that due to tidal effects, in terms of equations (25) and (26). In fact, assuming a constant total mass, it is often stated that

$$\frac{\dot{n}_E}{n_E} = 2 \frac{\dot{G}_E}{G_E} = -2 \frac{\dot{R}_E}{R_E} \quad (27)$$

Observational results are often presented in terms of \dot{G}/G , using equation (27) (see, e.g. Reasenberg & Shapiro 1977, 1978). The latter is the result of what Dirac has referred to as a 'primitive theory' of varying G which stipulates the same Newtonian equations of motion with G being a function of time. Such a stipulation by itself is not necessarily wrong unless one also imposes certain conservation laws which are not compatible with the assumed dynamical equations. This is most easily understood in the framework of Einstein's theory of gravitation where for a given Lagrangian, both the dynamical equations and the conservation laws are specified. Considering the Newtonian equations as a classical limit of Einstein's theory, equations (25) and (26) must be used in conjunction with the constraint (24), i.e. $G_E M_E = \text{constant}$. Hence, the only information one can get from these two equations is that

$$\frac{dn_E}{dt_E} = 0, \quad \frac{dR_E}{dt_E} = 0. \quad (28)$$

Thus imposition of the standard conservation law, namely, constant total mass, cannot be compatible with variable G .

On the other hand, measurements of planetary orbital parameters using atomic instruments ought not to be considered as a test of Einstein's *theory of gravitation per se*. Rather, assuming the latter's validity, such observations should be considered a test of the constancy of β . Since the orbital period and radius are both small intervals compared to the cosmo-

logical scale, the differential scaling law, equation (1), applies so that

$$n_E = \frac{1}{\beta} n; \quad R_E = \beta R \quad (29)$$

where n and R now refer to the corresponding orbital parameters measured in atomic units. With the assumed constancy of n_E and R_E , we find immediately

$$\frac{\dot{n}}{n} = \frac{\dot{\beta}}{\beta} = -\frac{\dot{R}}{R} \quad (30)$$

Thus having introduced the distinction between two dynamical clocks, a non-vanishing observational result in \dot{n} and \dot{R} can be easily understood. Even without reference to the constraint (24), primitive theory and the new theory can be distinguished in that they imply different ratios,

$$\frac{\dot{n}_E/n_E}{\dot{R}_E/R_E} = -2, \quad \frac{\dot{n}/n}{\dot{R}/R} = -1$$

which can be checked by observations.

One can derive equation (30) in a more elaborate fashion, using the equation of motion (5b). However, as can be recognized, equations (5a) and (5b) are conformal transformations of each other. Results of equation (5b) can be obtained from those of equation (5a) by a conformal transformation.

To gain more information from equation (30), we first note that for non-relativistic matter, the energy density ρ is mostly due to the rest mass density. Hence with $\Gamma = 0$, equation (11) gives

$$\rho V \sim (G\beta)^{-1}$$

or

$$GM\beta = \text{constant}, \quad (31)$$

which replaces equation (24). Clearly for $\beta = 1$, the two coincide. We now have

$$\frac{\dot{n}}{n} = -\frac{(GM)^{\cdot}}{GM} = -\frac{\dot{R}}{R} = \frac{\dot{\beta}}{\beta} \quad (32)$$

Thus, unless one makes more detailed stipulations, observational determination of the variation of β by measuring \dot{n} and \dot{R} gives no information about the separate variation of G and M . This is expected of all gravitational experiments dealing with geodesic motions. For the latter respond to the source strength which is always characterized by the combination GM . On the other hand, given hypotheses (a), (b) or (c) as described in the previous section, the variations of G and/or M are specified by cosmological considerations. These in turn determine $\dot{\beta}/\beta$ which can be checked, using equation (30), against the measured values of \dot{n}/n and \dot{R}/R . Before making more detailed comments on such a comparison, we shall consider another effect of varying β .

If the strength of the gravitational interaction does change with respect to that of atomic dynamics, the size of a macroscopic object such as a planet or a star would be expected to vary with time. To see this, we consider a model in which matter has an equation of state of the form

$$p \sim \rho^\gamma \quad (33a)$$

where p and ρ are respectively the pressure and mass density. γ is called the polytropic index.

It has been shown (Canuto *et al.* 1977) that from equation (4b), the equation of hydrostatic equilibrium in the non-relativistic limit retains the standard Newtonian form

$$\frac{dp}{dr} = -\rho \frac{GM}{r^2}. \quad (33b)$$

From equations (33a) and (33b) we get

$$GM^{2-\gamma} r^{3\gamma-4} = \text{constant},$$

so that

$$\frac{\dot{r}}{r} = \frac{1}{4-3\gamma} \frac{\dot{G}}{G} + \frac{2-\gamma}{4-3\gamma} \frac{\dot{M}}{M}. \quad (34a)$$

Making use of equation (31), we can write

$$\frac{\dot{r}}{r} = \frac{\gamma-2}{4-3\gamma\beta} \frac{\dot{\beta}}{\beta} + \frac{\gamma-1}{4-3\gamma} \frac{\dot{G}}{G}. \quad (34b)$$

It is important to note that, unlike the case of a planetary orbit, the variation of the radius of the kind of macroscopic object under consideration cannot be reduced to purely a variation of the product GM . Hence the two types of observations (32) and (34b) together give not only the value of $\dot{\beta}$, but \dot{G} and \dot{M} separately.

There has been indications (McElhinny, Taylor & Stevenson 1978) that the Earth and Mars have been slowly expanding whereas Mercury has been contracting. On the other hand, no observable variation of the size of the Moon is detected. Notwithstanding the difficulties of interpreting palaeomagnetic data, McElhinny *et al.* results cannot be directly compared with equation (34b) because the latter is merely a rough approximation for a simplistic polytropic model. Many geothermal effects which can contribute to the variation of the radius have not been included. *Nevertheless, the model does point out the fact that in situations where a balance of two types of forces is at play, one can gain information on the separate variation of G and M .* We therefore venture to suggest that with sophisticated computation of stellar structure and high-resolution measurement of the solar radius, one can perhaps have sufficient accuracy for a determination of \dot{G}/G .

Finally, we return to the observational determination of $\dot{\beta}$ using equation (30). The latest values of the observed total variation of n for the Moon are

$$\dot{n}_a = (-23.8 \pm 4) \text{ arcsec cy}^{-2} \quad (\text{Williams, Sinclair \& Yoder 1978}),$$

$$\dot{n}_a = (-24.6 \pm 1.6) \text{ arcsec cy}^{-2} \quad (\text{Calame \& Mulholland 1978}),$$

$$\dot{n}_a = (-21.5 \pm 3.2) \text{ arcsec cy}^{-2} \quad (\text{Van Flandern 1979, private communication}).$$

From these one must subtract the contribution due to tidal effects (Muller 1978).

$$\dot{n}_t = (-30.0 \pm 3) \text{ arcsec cy}^{-2}$$

so that the net variation can be expressed as

$$\frac{\dot{n}}{n} = \frac{1}{n} (\dot{n}_a - \dot{n}_t) \quad (35)$$

where $n = 1.73 \times 10^9$ arcsec/cy. These data indicate that

$$\begin{aligned} \frac{\dot{n}}{n} &= (3.6 \pm 2.9) 10^{-11}/\text{yr} \\ &= (3.1 \pm 2.0) 10^{-11}/\text{yr} \\ &= (4.8 \pm 2.4) 10^{-11}/\text{yr}. \end{aligned} \quad (36)$$

Hence from equation (32),

$$(GM)^{\cdot} > 0, \quad \dot{\beta} > 0. \quad (37)$$

Comparison with equations (14), (16) and (18) shows that of the three different gauge conditions proposed for the determination of β , only Dirac's matter creation gauge (14), is excluded by present observations.

4 Conclusions

We have advocated the interpretation of a varying G as a non-constant scaling between atomic and gravitational clocks. Consistent dynamical equations with a varying G can thus be written and observational data should be analysed in terms of these equations. The scaling function β is not dynamically determined. This is because the theory is as yet incomplete. In fact so far we have only expressed gravitational dynamics in terms of atomic units. When a theory of coupled atomic and gravitational dynamics will be available, we expect β to emerge naturally as a dynamical field variable.

On the other hand, observational constraints on β can be a valuable guide for the construction of such a theory. It is therefore useful to develop further the present theory to study its astrophysical and cosmological consequences. Part of this task has been accomplished (Canuto & Hsieh 1978; Canuto, Hsieh & Owen 1979). It is found that currently available observations on magnitudes, angular diameters and radio-data do not exclude a varying G of one part in $10^{11}/\text{yr}$.

References

- Calame, O. & Mulholland, J. D., 1978. *Science*, **199**, 977.
 Canuto, V., Adams, P. J., Hsieh, S.-H. & Tsiang, E., 1977. *Phys. Rev. D*, **16**, 1643.
 Canuto, V. & Hsieh, S.-H., 1978. *Astrophys. J.*, **224**, 302.
 Canuto, V., Hsieh, S.-H. & Owen, J. R., 1979. *Astrophys. J. Suppl.*, October issue.
 Dirac, P. A. M., 1938. *Proc. R. Soc. London A*, **165**, 199.
 Dirac, P. A. M., 1974. *Proc. R. Soc. London A*, **338**, 439.
 McCrea, W. H., 1974. *Ann. N.Y. Acad. Sci.*, **262**, 495.
 McCrea, W. H., 1978. *Observatory*, **98**, 52.
 McElhinny, M. W., Taylor, S. R. & Stevenson, D. J., 1978. *Nature*, **271**, 376.
 Muller, P. M., 1978. In *On the Measurement of Cosmological Variations of the Gravitational Constant*, p. 93, ed. Halpern, L., University Presses of Florida, Gainesville, Florida, and preprint, 1978.
 Reasenberg, R. D. & Shapiro, I. I., 1977. In *Experimental Gravitation*, Roma, Academia Nazionale dei Lincei.
 Reasenberg, R. D. & Shapiro, I. I., 1978. In *On the Measurement of Cosmological Variations of the Gravitational Constant*, p. 71, ed. Halpern, L., University Presses of Florida, Gainesville, Florida.
 Williams, J. G., Sinclair, W. S. & Yoder, C. F., 1978. *Geophys. Res. Lett.*, **5**, 943.