International Journal of Theoretical Physics, Vol. 21, No. 8/9, 1982

# Meaning of a Varying Gravitational Constant[1]

**S. -H. Hsieh**

*Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260*

*Received October 10, 1981*

The subject of this work has the distinction of having held Professor Dirac's interest and attention for over 40 years. It started in 1937 with his observation of the numerical coincidences of several large dimensionless numbers, which led him to propose the large numbers hypothesis (LNH). There have been several variations in the statement of the hypothesis. We shall summarize it for the purpose of this paper as follows:

From the constants arising in Nature, one can form the following dimensionless numbers:

$$N_1 = \frac{e^2}{Gm_e m_p} \cong 10^{40}$$

$$N_2 = \frac{m_e c^3}{H_0 c^2} \cong 10^{40}$$

$$N_3 = \frac{4\pi}{3}\frac{\rho_o}{m_p}\left(\frac{c}{H}\right)^3 \cong 10^{80}$$

where $e$ is the charge of the electron, $G$ is the gravitational constant, $m_e$, $m_p$ are the electron and proton mass, respectively, $c$ is the velocity of light, $H_0$ is the Hubble constant, and $\rho_o$ is the observed mass density of the universe. The LNH amounts to asserting that

$$\log N_1 = \log N_2 = \tfrac{1}{2}\log N_3 \tag{1}$$

is a law of nature. Noting that $N_2$ is a measure of the age $t$ of the universe in

[1]Presented at the Dirac Symposium, Loyola University, New Orleans, May 1981.

an atomic time scale, equation (1) implies that both $N_1$ and $N_3$ are epoch dependent.

Stringent limit has been set for the variation of atomic constants (see, e.g., Dyson, 1978). Hence accepting equation (1) leads to

$$G \sim \frac{1}{t} \tag{2}$$

More importantly, we note that $N_1$ is the ratio of electrostatic to gravitational forces between an electron and a proton. This ratio is the key to our understanding of $G$ variation and serves as the cornerstone of our theoretical framework. The number $N_3$ can be interpreted as the number of baryons within our horizon. Depending on detailed theoretical models, the relation

$$N_3 \sim t^2 \tag{3}$$

may or may not necessitate the postulate of spontaneous matter creation.

The scale covariant theory (SCT) of *gravitation* is an attempt toward the understanding of equations (2) and (3). In this paper we shall concentrate on presenting the theoretical interpretation of equation (2) and remark only that the formulation of the SCT is sufficiently general to admit the possibility of matter creation. The work reviewed below resulted from the collaboration between the author and V. M. Canuto, along with P. J. Adams, J. Lodenquai, J. R. Owen, and E. Tsiang at one time or another.

Given the suggestion that $G$ changes with cosmological time, one must first inquire what exactly is meant by a varying $G$. A meaningful answer should possess at least the following two ingredients: a theoretical framework with clearly stated basic hypotheses and a clear explanation of how such a variation is to be observed. For example, $G$ is a constant by definition in the Newtonian theory of gravitation. Having disproven the Newtonian theory experimentally does not imply that $G$ is varying. A similar remark would apply to Einsteins' theory of gravitation should it be proven wrong. These obvious statements are designed to invoke the equally obvious converse: Even if $G$ is observed to vary, it need not imply that the Einstein theory of *gravitation* is proven wrong.

$G$ has dimensions and therefore its measured values depend on what units have been used. In space-time theories, the fundamental unit is a length. This unit length is given by a well-specified measuring instrument and procedure. However, the measuring instrument, being a physical system must itself be governed by dynamical laws. We can thus classify units by the governing dynamics of the measuring instrument. For example, in a hydrogen maser or other atomic clocks, photons are emitted and absorbed by atoms. The governing physical law is electrodynamics. Hence the atomic clock is an electrodynamic clock. For a gravitational clock, one can use the Marzke–Wheeler construction (1964), with light signals sent back and forth between two nearby geodesics. A more naive model is given by two orbiting gravitating bodies, such as the earth–sun system. Thus the *year* is a gravitational unit of time. The earth–sun distance, commonly called the *astronomical unit*, is a gravitational unit of length.

Denoting the infinitesimal electrodynamical and gravitational units of time by $dt$ and $d\bar{t}$, we can write

$$d\bar{t} = \beta(x)\, dt \tag{4}$$

Since electrodynamics and gravitation are separate dynamical theories, there is no *a priori* reason for asserting the constancy of $\beta$. In principle, $\beta$ can be a general function of space-time. For practical purposes, we shall consider it to be a function of the epoch only. Once there are two units that do not have constant scaling between them, the dimensionful value of $G$ can be constant in one unit and not constant in another. In particular, we can have Einstein's theory of gravitation with its constant $G$ value and still measure a varying $G$ when atomic clocks are used.

There have been several atomic time measurements of the period $P$ of the moon orbiting about the earth. A description of the work of several independent research groups can be found in Van Flandern's recent paper (1981). We simply point out that after subtracting the gravitational perturbative (tidal) effects, Van Flandern gives

$$-\frac{\dot{P}}{P} = \frac{\dot{n}}{n} = (3.2 \pm 1.0) \times 10^{-11}/\text{year} \tag{5}$$

where $n = 2\pi/P$ is the angular velocity. Given the complexity of the data analysis, we must certainly await further confirmation by different, independent experimental tests before concluding that $G$ does indeed vary. Nevertheless, it can be asserted that at present, there exists no evidence against a variation of $G$ at the level suggested by the LNH. For an age of the universe in the range $(10. - 20.) \times 10^9$ years, equation (2) gives

$$\frac{\dot{G}}{G} = (5. - 10.) \times 10^{-11}/\text{year} \tag{6}$$

People may wonder why one should be interested in such a slow variation of $G$. If the finiteness of the lifetime of the proton is of profound theoretical interest, even though it is known to be at least 20 orders of magnitude larger than the age of the universe, we are biased to think that the variation of $G$ in a time scale comparable to age of the universe has even greater theoretical interest. In any case, if the variation is confirmed, the result is most easily interpreted in terms of the scaling between different dynamical units. With this point of view, we can set up a theoretical framework to systematically analyze the astrophysical consequences of a varying $G$.

We want to emphasize the words "framework" and "systematic." In many earlier works on varying $G$, it had been assumed that whenever the constant $G$ appeared explicitly in a standard equation, it could be simply replaced by a variable $G$. If it was not involved in a given equation, the latter was not expected to be modified even if one accepted the possibility of a changing $G$. Such an approach is what Professor Dirac has termed a *primative theory of a varying G*. We shall see that the consistency of the primative theory is questionable. In the theoretical analysis, we must proceed with caution, ascertaining that all equations used, no matter how obvious they may be in the corresponding standard theory, are either consequences of equations already stipulated in the new theory or compatible, independent assumptions.

Before describing our framework, we first summarize the basic assumptions of the scale-covariant theory of gravitation.

(1) There exist two distinct dynamical units, gravitational and electrodynamical, measuring space-time intervals $d\bar{s}$ and $ds$, respectively.
(2) $d\bar{s} = \beta(x)\,ds$.
(3) Einstein's theory of gravitation correctly describes gravitational phenomena.

These assumptions allow the description of gravitational phenomena when observations are made with atomic instruments: Assumption (2) implies

$$\bar{g}_{\mu\nu} = \beta^2 g_{\mu\nu} \tag{7}$$

the gravitational and atomic metrics are conformally related. Thus in atomic units, the equations of motion are simply obtained by conformal transformation from the corresponding equations of Einstein's theory. By so doing, we also give a physical meaning to the otherwise purely mathematical operation of conformal transformation.

As examples, the field equation, the conservation equation and the equation of motion in the two units are

$$\bar{G}_{\mu\nu} = -\bar{G}\bar{T}_{\mu\nu} \tag{8a}$$

$$G_{\mu\nu} = -G(\beta)T_{\mu\nu} + g_{\mu\nu}\left(2\frac{\beta^{\lambda}_{;\lambda}}{\beta} - \frac{\beta^{\lambda}\beta_{\lambda}}{\beta^2}\right) + 4\frac{\beta_\mu\beta_\nu}{\beta^2} - 2\frac{\beta_{\mu;\nu}}{\beta} \tag{8b}$$

$$\bar{T}^{\mu\nu}_{;\nu}=0 \tag{9a}$$

$$T^{\mu\nu}_{;\nu}+(2-g)(\ln\beta)_{,\nu}T^{\mu\nu}-(\ln\beta)_{,}{}^{\mu}T_{\nu}^{\nu}=0 \tag{9b}$$

$$\bar{u}^{\mu}_{;\nu}\bar{u}^{\nu}=0 \tag{10a}$$

$$u^{\mu}_{;\nu}u^{\nu}+\frac{\beta_{\nu}}{\beta}\left(u^{\mu}u^{\nu}-g^{\mu\nu}u^{\rho}u_{\rho}\right)=0. \tag{10b}$$

The barred quantities are in gravitational units and the symbols have their conventional meanings: $\bar{G}_{\mu\nu}$ and $G_{\mu\nu}$ are the Einstein tensors, $T_{\mu\nu}$ is the energy-momentum tensor, $u_{\mu}$ is the particle velocity. In this framework, the gravitational constant is treated as a coscalar, i.e., under the transformation of units, $G$ and $\bar{G}$ are related by

$$\bar{G}=\beta^{g}G \tag{11}$$

Several remarks should be made concerning the equations above.

1. Since the Einstein equation is *not* scale invariant, it is not surprising that there are various $\beta$-dependent terms in the conformally transformed equations.

2. $\beta$ enters as a nondynamical variable. This is to be expected as we retained unchanged Einstein's gravitational dynamics. At this stage we can determine $\beta$ either from observational data or from nondynamical constraints, such as the LNH, as expressed by equations (2) and (3). The latter approach, though important for our understanding of the LNH, will not be elaborated on (Canuto et al., 1977). A particularly simple example of observational determination of $\beta$ can be given. From the assumption of scaling between the dynamical clocks, we have

$$\bar{P}=\beta P=\text{const}$$

where the second equality follows from assumption (3) above. Consequently,

$$\frac{\dot{\beta}}{\beta}+\frac{\dot{P}}{P}=0 \tag{12}$$

and

$$\frac{\dot{\beta}}{\beta}=\frac{\dot{n}}{n}=(3.2\pm 1.0)\times 10^{-11}/\text{year}$$

if one accepts Van Flandern's results. Note that only the derivative of $\beta$ is fixed by the data. The value of $\beta$ can be normalized at one space-time point. It is convenient to have $\beta_0 = \beta(t_0) = 1$, normalized at the present epoch.

3. Accepting Van Flandern's result, we know only that $\beta$ is not constant, which means that our hypotheses have nontrivial content. To determine how $G$ varies, one must further specify the theoretical model. More observational data are needed before different models can be distinguished (Canuto et al., 1979a).

Using essentially these equations, we can study the cosmological implications of a varying $G$, i.e., Dirac cosmology. We shall not discuss the details here (Canuto et al., 1979b; Canuto and Owen, 1979). Suffice it to say that cosmological data cannot be used to constrain the variation of $G$ or $\beta$ at the level indicated earlier. Instead, we point out some of the consequences of the seemingly mild assumptions, and illustrate how two of the most long-standing difficulties encountered by theories of varying $G$ can be naturally resolved in the framework of the SCT. One difficulty as suggested by Teller (1948) concerns the luminosity of the sun. If the sun had been too luminous in the past, the earth would have been very hot, contradicting existent geological and biological records. The other difficulty is associated with the background radiation which is observed to have an equilibrium distribution. It was pointed out that in theories whose conservation law differs from (9a), the background radiation considered as remnants from an earlier epoch cannot have an equilibrium form.

Equation (9b) shows that by accepting our assumptions we no longer have the standard energy-momentum conservation laws. However, the laws have *not* been totally abandoned; they are merely modified. To gain some insight into these modified laws, we consider pressureless matter having mass density $\rho$ and macroscopic velocity $u^\mu$ so that

$$T^{\mu\nu} = \rho u^\mu u^\nu \tag{13}$$

Assuming that $\rho$ vanishes outside a finite region, it is easy to show that (Canuto and Hsieh, 1980)

$$GM\beta = \text{const} \tag{14}$$

where $M$ is the total mass within the region. For comparison, we also consider a primitive theory in which the gravitational field equation is

$$G_{\mu\nu} = -G(t)T_{\mu\nu} \tag{15}$$

Then, by virtue of the Bianchi identity applied to the left-hand side of equation (15),

$$(GT^{\mu\nu})_{;\nu} = 0 \tag{16}$$

and thus leading to

$$GM = \text{const} \tag{17}$$

The astrophysical relevance of these constraints can be understood with the following example. Teller had argued using equations of the primitive theory that the luminosity $L$ of the sun scaled with $G$ as

$$L \sim G^7 M^5 \tag{18}$$

Restricting himself to the no-matter creation theory, he set $M$ constant and concluded that $L \sim t^{-7}$ yielding an unreasonably large solar luminosity a few billion years ago. Of course, if the constraint (17) had been applied, we would get formally only $L \sim t^{-2}$. But in fact if one had used a primitive theory as given by equation (15), the gravitational constant cannot be changing if one stipulates also that the total mass remains constant. One can contend that Teller had used a version of the primitive theory in which the Newtonian force law rather than equation (15) had been stipulated. There would then be no associated conservation laws and no constraint needs to be observed. But then Teller's analysis merely shows the nonviability of a $G$-varying theory based on a nonviable gravitational theory. On the other hand, if the Newtonian force law is considered to be derivable from equation (15) in the appropriate Newtonian limit, equation (17) must then follow as a constraint. Using the SCT, arguments analogous to those of Teller's show that (Canuto and Hsieh, 1981)

$$L \sim \frac{1}{\beta \kappa} \tag{19}$$

where $\kappa$ is the opacity of the solar interior. With the kind of $t$ dependence of $\beta$ we have been interested in (because of the requirement of compatibility with the LNH),

$$\beta \sim t^{\varepsilon}, \qquad \varepsilon = \pm 1, \pm \tfrac{1}{2} \tag{20}$$

it can be seen that the form of equation (19) should hardly cause any alarm as far as violation of observational constraint is concerned.

A lesson can be extracted from the above example. If a primitive theory of the form of equation (15) is used, and if $M = \text{const}$ is implicitly assumed, consistent analysis of data must necessarily yield $G = \text{const}$. For this reason, we echo Dyson's appeal (1978) that in analyzing data for evidence (or lack of evidence) for a varying $G$, one must be acutely aware of how theory dependent the conclusion may be.

To understand why there had been claims that the equilibrium form of the observed background radiation may constrain $G$-varying theories, it is best to first review the situation in the standard theory. There are three main ingredients in the analysis: (1) *The standard conservation law applied to radiation in the universe*:

$$\rho_{\gamma_1} R_1^4 = \text{const} = \rho_{\gamma 0} R_0^4 \tag{21a}$$

where $\rho_\gamma$ is the radiation energy density and $R$ is the scale factor in the Robertson–Walker metric. The subscripts 0 and 1 refer to the values of the quantities at different epochs. For different spectral frequency ranges, we can write

$$d\nu \rho_{\gamma\nu} R^4 = \text{const} = d\nu_0 \rho_{\gamma\nu_0} R_0^4 \tag{21b}$$

if there is no interaction for mixing the radiation energy at different frequencies. (2) The cosmological red shift relation:

$$\nu R = \text{const} = \nu_0 R_0 \tag{22}$$

(3) The form of the equilibrium radiation distribution:

$$\rho_{\gamma\nu} = \text{const}\, \nu^3 \mathfrak{f}(\nu/T) \tag{23}$$

Here $\mathfrak{f}$ is an universal function of its argument and $T$ is the equilibrium temperature. At earlier epochs of the universe when there were large amounts of ionized material present, radiation was coupled to matter with a scattering mean free time short compared to the expansion time scale of the universe. Thus, the radiation had an equilibrium distribution given by equation (23). After decoupling, the radiation undergoes freestreaming. Its present spectrum can be deduced from equations (21b) and (22) to be

$$\rho_{\gamma\nu_0} = \text{const}\, \nu_0^3 \mathfrak{f}(\nu_0/T_0)$$

again an equilibrium distribution which agrees with observation. Note that $T_0$ is merely a scaled temperature, having no thermodynamical meaning.

It was argued that if the conservation law is modified, initial equilibrium distribution will not evolve to the present epoch retaining its equilibrium form. Indeed, in the SCT, we have instead of (21b),

$$d\nu \beta^{2-g} \rho_{\gamma\nu} R^4 = \text{const} \tag{24}$$

If the other two ingredients remain unchanged, the theoretically predicted present distribution would have been

$$\rho_{\gamma\nu_0} = \text{const}\left(\frac{\beta_1}{\beta_0}\right)^{2-g} \nu_0^3 \mathfrak{f}\left(\frac{\nu_0}{T_0}\right)$$

where $\beta_1$ refers to the value of $\beta$ at decoupling. Comparison of the above with the observed equilibrium distribution of the form of equation (23) would require $\beta_1/\beta_0$ to be very close to unity and would strongly constrain the possible variation of $G$. However, before rushing to the conclusion that $G$ cannot be changing, one should first inquire whether equations (22) and (23) are compatible with a $G$-varying scheme. In the framework of the SCT, we have shown that equation (22) remains valid while (23) must be modified: Equilibrium radiation distribution must have the form

$$\rho_{\gamma\nu} = \text{const}\, \nu^3 \beta^{g-2} \mathfrak{f}(\nu/\beta T) \tag{25}$$

Using equation (25) along with (22) and (24), we can again assert that equilibrium radiation at decoupling evolves to the present retaining its equilibrium form. That is, observing the background radiation today, we cannot distinguish its spectral distribution from one of equilibrium radiation at some scaled temperature:

$$\rho_{\gamma\nu_0} = \text{const}\, \nu_0^3 \beta_0^{g-2} \mathfrak{f}(\nu_0/\beta_0 T_0)$$

There are two complementary ways to see that modification of the equilibrium distribution should be expected. Classical thermodynamical derivation of Wien's displacement law gives the form of the equilibrium distribution essential for our discussion. [The universal function $\mathfrak{f}$ in equations (23) and (25) may take the Boltzmann form if classical statistics is used or the Planck form if quantum statistics is considered. At the present stage of the development of the theory of varying $G$, we shall not concern ourselves with quantum statistics.] Modification of the equation of conservation in the SCT requires modification of the first law of thermodynamics, essential for the derivation of Wien's displacement law. Consequently, the equilibrium distribution must differ from the standard one if $G$ varies according to our assumptions. Alternatively, from a statistical viewpoint, the equilibrium distribution can be obtained by maximizing the probability distribution in phase space under the constraints of constant energy and particle number (if the latter's conservation holds). At least one of these constraints must be changed in the SCT. Hence we again expect modification of the equilibrium distribution. Along this line, we have recently developed the kinetic theory for classical particles, consistent with the assumptions of the SCT (Hsieh and Canuto, 1981). The corresponding Liouville equation can be solved to obtain an equilibrium distribution function which agrees with an earlier version obtained by thermodynamic arguments (Canuto and Hsieh, 1979).

Thus we have seen that the basic assumptions of the SCT, forces modifications of laws which may appear, at first sight, to have nothing to do with gravitation. Upon reflection, this fact should not be surprising. General relativity, with its stipulation of a constant $\beta$ in the form of the strong equivalence principle specifies a particular form of coupling between gravitation and electrodynamics. With a variable $\beta$, we preserve only the weak equivalence principle, which is essential for a geometrical theory of gravitation. There have been many attempts in modifying general relativity. An illustrious example is the Brans-Dicke theory. But in that theory, the gravitational part is modified while the atomic (electrodynamic) part is kept intact. The SCT suggests an alternative approach. Keeping the gravitational part of general relativity intact, we are thus forced to modify the atomic and/or the coupling term.

It is appropriate to emphasize here that the scale covariant theory of *gravitation* as we have presented it, is an incomplete theory. The source of gravitation, $T_{\mu\nu}$, should only be treated on a phenominological level at this stage. Also, by professing theoretical ignorance of $\beta$ (no governing dynamical equation), we are allowing for our incomplete knowledge of the relation between gravitation and electrodynamics. To use a term that has regained its popularity, we do not yet have a unified theory. We envision that in a complete, unified theory, the variation of $\beta$ and $G$ can be determined *a priori*. Such variations can then be checked against observations and those suggested by the LNH.

It had been natural for us to conclude that the above represented our understanding of Dirac's ideas, as he inspired and motivated much of the work. But perhaps we had been immodest in having claimed understanding of Dirac's ideas. Could he not have something considerably more profound in mind? After all, what we have done is quite simple. What we suggest we should do to complete the theory may sound ambitious. But only tame, conventional ideas are involved—unified theories are the high fashion of the times. We are thus led to further speculations about the meaning of the LNH. It is a very unconventional type of hypothesis for it does not at all appear to be a fundamental principle. One's immediate inclination is to explain it in terms of some other more fundamental principles, as we have attempted to do. This is only because standard physical laws have been local, differential laws. The LNH may be a pioneering example of global laws of nature: By considering the universe as a whole, we may be led to a deeper understanding of the "fundamental" interactions. We have been fortunate to have Professor Dirac as our beacon and we hope he will continue to enlighten us.

## REFERENCES

Canuto, V., and Hsieh, S.-H. (1979). *Astrophysical Journal Supplement*, **41**, 243.
Canuto, V., and Hsieh, S.-H. (1980). *Astrophysical Journal*, **237**, 613.
Canuto, V., and Hsieh, S.-H. (1981). *Astrophysical Journal*, **248**, 801.
Canuto, V., and Owen, J. R. (1979). *Astrophysical Journal Supplement*, **41**, 301.
Canuto, V., Adams, P. J., Hsieh, S.-H. and Tsiang, E. (1977). *Physical Review, D*, **16**, 1643.
Canuto, V., Hsieh, S.-H., and Owen, J. R. (1979a). *Monthly Notices of the Royal Astronomical Society*, **188**, 829.
Canuto, V., Hsieh, S.-H., and Owen, J. R. (1979b). *Astrophysical Journal Supplement*, **41**, 263.
Dyson, F. (1978). In *Current Trends in the Theory of Fields*, AIP Conference Proceedings #48.
Hsieh, S.-H., and Canuto, V. (1981). *Astrophysical Journal*, **248**, 790.
Marzke, R. F., and Wheeler, J. A. (1964). In *Gravitation and Relativity*, Chiu, H. Y., and Hoffmann, W. F. eds., Benjamin, New York.
Teller, E. (1948). *Physical Review*, **73**, 801.
Van Flandern, T. (1981). *Astrophysical Journal*, **248**, 813.