28 July 1997

PHYSICS LETTERS A

ELSEVIER

Physics Letters A 232 (1997) 165–170

Consistently implementing the field self-energy in Newtonian gravity

Domenico Giulini [1]

Fakultät für Physik, Universität Freiburg, Hermann-Herder Strasse 3, D-79104 Freiburg, Germany

Received 9 May 1996; revised manuscript received 1 May 1997; accepted for publication 6 May 1997
Communicated by P.R. Holland

Abstract

We consider alterations to Newtonian gravity satisfying the principle that all energies contribute to the active gravitational mass. We discuss earlier attempts and point out their inconsistency with this principle. A consistent prescription is derived and discussed. © 1997 Elsevier Science B.V.

1. Introduction

The issue addressed in this Letter arises if one wishes to model the self-coupling of the gravitational field within Newtonian gravity. Simple non-linear alterations of Newton's field equation are often employed as simplified models for general relativity. The purpose of this Letter is to show how this can be done and to point out certain flaws in the usually accepted prescription, as for example given in Refs. [1,2].

To be more precise, we recall that the Newtonian gravitational field, φ, and the density of (ponderable) matter, ρ, obey

$$\Delta\varphi = 4\pi G\rho, \tag{1.1}$$

where G is Newton's gravitational constant. The force per unit volume is given by

$$f = -\rho\nabla\varphi. \tag{1.2}$$

Together these equations imply that in order to build up a field φ from $\varphi = 0$ one has to invest the work

$$A = -\frac{1}{8\pi G}\int_{\mathbb{R}^3} \|\nabla\varphi\|^2 \, dV. \tag{1.3}$$

If we add the assumption that all energy acts as active gravitational mass, according to $E = mc^2$, and also think of the integrand in (1.3) as representing energy density, we might be tempted to consider the modified equation

$$\Delta\varphi = 4\pi G\left(\rho - \frac{1}{8\pi Gc^2}\|\nabla\varphi\|^2\right) \tag{1.4}$$

with the aim to incorporate into Newtonian gravity the following heuristic

Principle (**P**). All energy acts as source for the gravitational field.

Note that a priori we do not know what "energy" is in the theory to be formulated, so that **P** should be read as requirement of self-consistency. In the next section

[1] E-mail: giulini@sun2.ruf.uni-freiburg.de.

0375-9601/97/$17.00 © 1997 Elsevier Science B.V. All rights reserved.
PII S0375-9601(97)00369-1

we follow a standard practice by introducing the notion of "energy" through the mechanism of coupling the gravitational field to mass distributions and calculating the work that the field performs on them. Using energy conservation then leads to an expression for the field energy.

A field equation satisfying **P** must be non-linear. One might wonder whether (1.4) gives a Newtonian model that satisfies **P**. If it were true that it shared this qualitative feature with general relativity one might profitably employ this single scalar equation to study certain qualitative features of general relativity in a mathematically simpler environment. In fact, (1.4) is often proposed in pedagogical discussions to precisely this end [1–3]. For example, in Refs. [2,3] the authors suggest that some useful lessons concerning the energy-regulating power of the gravitational field can be learned from model theories of charged particles based on (1.4). In passing we remark that (1.4) can be written in a linear form by making the field redefinition $\psi := \exp(\varphi/2c^2)$,

$$\Delta\psi = \frac{2\pi G}{c^2}\rho\psi, \tag{1.5}$$

where the boundary conditions $\varphi(r \to \infty) = 0$ translate to $\psi(r \to \infty) = 1$. In the following we shall for simplicity always assume ρ to have compact support $B \subset \mathbb{R}^3$.

In Section 3 we discuss what is wrong with a theory based on (1.2), (1.4) and suggest a different and consistent theory in Section 4. Section 5 briefly discusses some properties of spherically symmetric solutions to the latter. In Section 2 we summarize some facts from Newtonian gravity. We employ the standard summation convention for repeated indices in up-down positions and use the Euclidean metric δ_{ab} to raise and lower indices. ∇_a denotes the partial derivative with respect to x^a. Three-component vectors are also written in boldface italic, $\boldsymbol{\xi}$, with $\boldsymbol{\xi} \cdot \boldsymbol{\eta} = \xi_a \eta^a$ denoting the scalar product.

2. Newtonian recollections

To see what is wrong with (1.4) it is helpful to first give a derivation of (1.3). Consider a one-parameter family of diffeomorphisms $s \mapsto \sigma_s$ such that $\sigma_{s=0} = \mathrm{id}$ and $(\mathrm{d}/\mathrm{d}s)|_{s=0}\sigma_s(\boldsymbol{x}) = \boldsymbol{\xi}(\boldsymbol{x})$. We wish to use σ_s to redistribute the matter by dragging it along this flow. Pulling back the 3-form $\rho\,\mathrm{d}V$ by the inverse diffeomorphisms we obtain $\rho_s\,\mathrm{d}V := (\sigma_s^{-1})^*(\rho\,\mathrm{d}V)$ and hence for the Lie derivative of the density ρ along $\boldsymbol{\xi}$

$$\delta\rho := \left.\frac{\mathrm{d}\rho_s}{\mathrm{d}s}\right|_{s=0} = -\nabla\cdot(\rho\boldsymbol{\xi}), \tag{2.1}$$

where here and in the following we use the variational symbol, δ, for the derivative at $s = 0$ and call it "the variation" of the quantity in question.

The variation of the work done to the system is easily determined using (1.2),

$$\delta A = -\int_{\mathbb{R}^3} \boldsymbol{\xi}\cdot\boldsymbol{f}\,\mathrm{d}V = -\int_B \varphi\nabla\cdot(\rho\boldsymbol{\xi})\,\mathrm{d}V. \tag{2.2}$$

Eqs. (2.1), (2.2) imply

$$\delta A = \int_B \varphi\delta\rho\,\mathrm{d}V. \tag{2.3}$$

This equation is independent of the field equation. If we assume adiabaticity, i.e. the validity of (1.1) throughout the motion, we can use (1.1) to eliminate $\delta\rho$ and write (2.3) solely in terms of φ. The result (1.3) then easily follows.

From (1.1), (1.2) it follows that the force per unit mass may be derived from a symmetric stress tensor, $f_a = -\nabla^b t_{ab}$, where

$$t_{ab} = \frac{1}{4\pi G}\left((\nabla_a\varphi)(\nabla_b\varphi) - \tfrac{1}{2}\delta_{ab}\|\nabla\varphi\|^2\right) \tag{2.4}$$

so that

$$\delta A = -\int_{\mathbb{R}^3} f_a\xi^a\,\mathrm{d}V = \int_{\mathbb{R}^3} \xi^a\nabla^b t_{ab}\,\mathrm{d}V$$
$$= \int_{\mathbb{R}^3} \nabla^{(a}\xi^{b)} t_{ab}\,\mathrm{d}V. \tag{2.5}$$

Here we assumed $\|\boldsymbol{\xi}(r \to \infty)\| < ar$ for some real constant a and that derivatives of φ fall off as fast as r^{-2}. Vector fields which satisfy $\nabla^{(a}\xi^{b)} = 0$ (Killing equation) generate rigid motions and are given by $\boldsymbol{\xi}(\boldsymbol{x}) = \boldsymbol{k}$ (translations) and $\boldsymbol{\xi}(\boldsymbol{x}) = \boldsymbol{k}\times\boldsymbol{x}$ (rotations), for constant $\boldsymbol{k}$. For those $\delta A = 0$, as it must be by the principle of action = reaction. Otherwise the system would self-accelerate.

Finally we define the gravitational mass as the total flux of the gravitational field φ out to infinity,

$$M_g := \lim_{r\to\infty} \frac{1}{4\pi G} \int_{S_r^2} \boldsymbol{n} \cdot \nabla\varphi \, do. \tag{2.6}$$

S_r^2 denotes a two-sphere of radius r, $\boldsymbol{n}$ its (outward pointing) normal, and do the surface element on S_r^2. The limit of integrals in (1.14) is sometimes abbreviated by $\int_{S^2_\infty}$.

3. Why inconsistent?

Since in Newtonian theory $M_g = M_m := \int \rho \, dV$, M_g only depends on the amount but not on the distribution of matter and clearly **P** cannot be satisfied. Now, replacing (1.1) by (1.4), one obtains the following formula for the variation δM_g,

$$\delta M_g = \int_B \sum_{n=0}^{N-1} \frac{1}{n!} \left(\frac{\varphi}{c^2}\right)^n \delta\rho \, dV + \frac{1}{N! c^{2N}} \frac{1}{4\pi G} \int_{\mathbb{R}^3} \varphi^N \, \delta(\Delta\varphi) \, dV, \tag{3.1}$$

where we have used (1.4) N times to replace $\Delta\varphi$. For a regular matter distribution φ will be bounded, say $\varphi(\boldsymbol{x}) < K, \forall \boldsymbol{x} \in \mathbb{R}^3$. Also, the integral over $(1/4\pi G)\delta(\Delta\varphi)$ just represents δM_g so that the last term on the right-hand side is majorized by $\frac{1}{N!}(K/c^2)^N \delta M_g$. It vanishes in the limit $N \to \infty$. In this limit the sum on the right side is just the exponential function. Thus we obtain the result

$$\delta M_g = \int_B \delta\rho \, \exp(\varphi/c^2) \, dV, \tag{3.2}$$

which, recalling (2.3), deviates from $\delta A/c^2$ by all the higher-than-linear terms in the expansion of the exponential. Hence (1.2), (1.4) violates **P**. This is not really surprising, since (1.3) was derived under the assumption of (1.1), (1.2). Changing it to (1.2), (1.4) also invalidates (1.3). A correct procedure must iterate the step that led from (1.1) to (1.4). For example, the next (second) step would be to determine a modified expression for the gravitational field energy from (1.2), (1.4) and then change (1.4) accordingly. Eventually this procedure should converge to a self-consistent field equation. However, as we will see in the next section, such a self-consistent field equation can actually be guessed directly.

At the end of this section we also point out another flaw in the combination (1.2), (1.4). Using (1.5) to replace ρ in (1.2) one easily derives

$$f_a = -\exp(-\varphi/c^2)\nabla^b(\exp(\varphi/c^2)\, t_{ab}) \tag{3.3}$$

with t_{ab} from (2.4). From this expression it follows that the force density is not the divergence of a stress tensor. There are many ways to isolate the part that obstructs the right-hand side of (2.11) to be written as the divergence of a symmetric tensor. Two obvious ways are

$$f_a = -\nabla^b t_{ab} - \frac{1}{8\pi G c^2}\|\nabla\varphi\|^2\nabla_a\varphi \tag{3.4a}$$

$$= -\nabla^b\left[(1+\varphi/c^2)t_{ab}\right] + \frac{1}{8\pi G c^2}\Delta\varphi\nabla_a\varphi^2. \tag{3.4b}$$

The system (1.2), (1.4) thus potentially violates the principle action = reaction [2].

4. A consistent modification

Eq. (3.2) was derived assuming (1.4) but not (1.2). If we maintain (1.4), (.5) but define $\phi = c^2\exp(\varphi/c^2)$ (rather than φ) to be the gravitational potential, then (3.2) just expresses the validity of **P**, i.e. $c^2\delta M_g = \delta A$ with δA given by (2.3). Throughout the gravitational field is always defined as minus the gradient of the potential, so that (1.2) and (1.4), (1.5) get replaced by

$$\boldsymbol{f} = -\rho\nabla\phi, \tag{4.1}$$

$$\Delta\phi = \frac{4\pi G}{c^2}\left(\rho\phi + \frac{c^2}{8\pi G}\frac{\|\nabla\phi\|^2}{\phi}\right). \tag{4.2}$$

Note that formally (4.2) is just a rewritten version of (1.4) in terms of ϕ. What changes its physical content is that now $-\nabla\phi =$ and not $-\nabla\varphi = -c^2\nabla\phi/\phi$ is the gravitational field. To be sure, for calculations

[2] To manifestly show a violation one should prove the existence of a regular solution to (1.4) with $\varphi(r\to\infty) = 0$ for which $\int f_a\xi^a \neq 0$ for some generator $\boldsymbol{\xi}$ of a rigid motion.

it is still easier to use the linear equation (1.5) whose solutions determine solutions to (4.2) by setting $\phi = c^2\psi^2$. ϕ must satisfy the boundary conditions $\phi(r \to \infty) = c^2$. The Newtonian approximation is obtained from expanding $\phi = c^2 + \varphi + O(\varphi^2)$ and keeping only linear terms in (4.2). Note that in the expression (2.6) for M_g we must write ϕ instead of φ. But for $r \to \infty$ only the linear term in φ contributes to the surface integral so that (3.2) is still valid. This is why (3.2) indeed expresses the validity of **P** for (4.1), (4.2). To be sure, once (4.1), (4.2) are established, the equation $c^2\delta M_g = \delta A$ is most easily proven directly. For completeness we give a short direct proof in the appendix. The point of our derivation of (3.2) was that it suggested the definition of ϕ in terms of φ and hence (4.2). It is interesting to note that (4.2) is precisely the equation that Einstein already proposed before the advent of general relativity in 1912 [4].

Eqs. (4.1), (4.2) also manifestly imply the principle action = reaction in the sense above. Indeed, we now have $f_a = -\rho\exp(\varphi/c^2)\nabla_a\varphi$. Replacing $-\rho\nabla_a\varphi$ by the right-hand side of (3.3) just cancels the exponential function outside the derivative so that the remaining divergence can be rewritten in terms of ϕ. This leads to the desired formula, $f_a = -\nabla^b t_{ab}$, with

$$t_{ab} = \frac{1}{4\pi G c^2}\left(\frac{1}{\phi}\left[(\nabla_a\phi)(\nabla_b\phi) - \tfrac{1}{2}\delta_{ab}\|\nabla\phi\|^2\right]\right). \tag{4.3}$$

We may interpret the two terms on the right-hand side of (4.2) as energy densities due to ponderable matter and the gravitational field respectively. The sum of both determines the convergence $\Delta\phi$ of the gravitational field $-\nabla\phi$. Both terms are positive since ϕ is positive. This is in contrast to (1.4), where the Newtonian gravitational field energy was negative definite, which is usually said to have its origin in the attractivity of gravity. But of course here gravity is also attractive. What is different here is that the (rest) energy of matter depends on the value of the gravitational potential at its location. This allows that a contraction of a matter distribution enhances the field energy although the total energy decreases. This is achieved by displacing the matter into regions of smaller gravitational potential and thereby sufficiently decreasing the matter part of the energy.

The total gravitational energy is given by

$$\begin{aligned} E_{\text{total}} &:= c^2 M_g = \int_B \rho\phi \, dV + \frac{c^2}{8\pi G}\int_{\mathbb{R}^3} \frac{\|\nabla\phi\|^2}{\phi}\, dV \\ &=: E_{\text{matter}} + E_{\text{field}}, \end{aligned} \tag{4.4}$$

where the expression for E_{field} can also be written in terms of an integral over B (the support of ρ) only. To see this we recall that for large distances from the source we have the expansions for ϕ and ψ

$$\frac{\phi}{c^2} = 1 - \frac{GM_g}{c^2 r} + O(r^{-2}), \tag{4.5a}$$

$$\psi = 1 - \frac{GM_g}{2c^2 r} + O(r^{-2}) \tag{4.5b}$$

so that E_{total} can also be expressed as an integral of $(c^4/2\pi G)\Delta\psi = c^2\rho\psi$ (using (1.5)) over B. Replacing E_{total} by this expression in $E_{\text{field}} = E_{\text{total}} - E_{\text{matter}}$ one obtains

$$E_{\text{field}} = c^2 \int_B \rho\sqrt{\frac{\phi}{c^2}}\left(1 - \sqrt{\frac{\phi}{c^2}}\right) dV. \tag{4.6}$$

5. Solution for homogeneous spherical star

In this section we determine the gravitational field for the externally prescribed mass distribution

$$\begin{aligned} \rho &= 3M_m/4\pi R^3, \quad &&\text{for } r < R, \\ &= 0, &&\text{for } r \geqslant R, \end{aligned} \tag{5.1}$$

where M_m is the total ("bare") mass of matter: $M_m = \int_B \rho \, dV$. It will be convenient to introduce the "matter radius" $\mathcal{R}_m$ and the "gravitational radius" $\mathcal{R}_g$,

$$\mathcal{R}_m = \frac{GM_m}{c^2}, \quad \mathcal{R}_g = \frac{GM_g}{c^2} \tag{5.2}$$

and the abbreviation

$$\omega = \sqrt{\frac{3\mathcal{R}_m}{2R}}\,\frac{1}{R}. \tag{5.3}$$

We use (1.5), set $\psi(r) = \chi(r)/r$, and obtain

$$\begin{aligned} \chi'' &= \omega^2\chi, \quad &&\text{for } r < R, \\ &= 0 &&\text{for } r \geqslant R. \end{aligned} \tag{5.4}$$

The general solution which makes ϕ (and hence ψ) finite at $r = 0$ is

$$\psi(r) = K\frac{\sinh(\omega r)}{r}, \quad \text{for } r < R,$$
$$= 1 - \frac{\mathcal{R}_g}{2r}, \quad \text{for } r \geqslant R. \tag{5.5}$$

The integration constants K and $\mathcal{R}_g$ are determined by the requirement that ϕ (and hence ψ) should be continuously differentiable at $r = R$,

$$\mathcal{R}_g = 2R\left(1 - \frac{\tanh(\omega R)}{\omega R}\right), \tag{5.6}$$

$$K = \frac{1}{\omega\cosh(\omega R)}. \tag{5.7}$$

Fixing the radius R in (4.6) gives us $\mathcal{R}_g$ as a function of $\mathcal{R}_m$, i.e. the gravitational mass as a function of the bare mass. In terms of the dimensionless quantities $y = \mathcal{R}_g/R$ and $x = \mathcal{R}_m/R$ this reads

$$y = f(x) = 2\left(1 - \frac{\tanh(3x/2)^{1/2}}{(3x/2)^{1/2}}\right), \tag{5.8}$$

which for $x \geqslant 0$ maps monotonically $[0,\infty] \to [0,2]$. For small x one has $f(x) = x - \frac{3}{5}x^2 + \frac{51}{140}x^3 + O(x^4)$. The fact that $f(x) < 2 \quad \forall x \in \mathbb{R}_+$ means that the gravitational mass is bounded by a quantity depending only on the geometry (here R) of the mass distribution,

$$M_g < R\frac{2c^2}{G}. \tag{5.9}$$

Note that this is achieved with all contributions to the gravitational mass on the right-hand side of (3.3) being positive. No subtractions are taking place. Rather, high matter densities ρ are suppressed by the small potentials ϕ produced by them (i.e. "red-shifted" in general relativistic terminology). This can be seen in detail from the following expressions,

$$E_{\text{total}} = \frac{2Rc^4}{G}\left(1 - \frac{\tanh(\omega R)}{\omega R}\right) \tag{5.10a}$$

$$= M_m c^2\left(1 - \frac{3\mathcal{R}_m}{5R} + O(\mathcal{R}_m^2/R^2)\right), \tag{5.10b}$$

$$E_{\text{matter}} = \frac{Rc^4}{G}\left(\frac{\tanh(\omega R)}{\omega R} + \tanh^2(\omega R) - 1\right) \tag{5.11a}$$

$$= M_m c^2\left(1 - \frac{6\mathcal{R}_m}{5R} + O(\mathcal{R}_m^2/R^2)\right), \tag{5.11b}$$

$$E_{\text{field}} = \frac{3Rc^4}{G}\left(1 - \frac{\tanh(\omega R)}{\omega R} - \frac{1}{3}\tanh^2(\omega R)\right) \tag{5.12a}$$

$$= M_m c^2\left(\frac{3\mathcal{R}_m}{5R} + O(\mathcal{R}_m^2/R^2)\right), \tag{5.12b}$$

where the second expressions on the right-hand sides are expansions of the first in terms of $\mathcal{R}_m/R$. Also, recall that $\mathcal{R}_m c^4/G = M_m c^2$. Note the familiar $\frac{3}{5}$-term in (5.10b) for the Newtonian binding energy.

Decreasing R for fixed $\mathcal{R}_m$ we see from (5.10b)–(5.12b) that to first approximation this enhances the field energy and at the same time decreases the matter energy twice as fast, so as to decrease the total energy by the same amount as that by which the field energy increased. Clearly the total energy must decrease in accordance with the attractivity of the gravitational interaction.

Coming back to (5.9) we next show that it remains valid for any spherically symmetric matter distribution. In particular, it remains valid for more realistic matter distributions (of compact support $r < R$) which are determined by a coupled system of (4.2) with some equations of state for the matter. The proof is simply this: For $r \geqslant R$ (5.4) is solved by $\chi_+(r) = r - \mathcal{R}_g/2$ and by some function $\chi_-(r)$ for $r \leqslant R$. Continuity and differentiability of ϕ at $r = R$ is equivalent to

$$\chi_-(R) = R - \tfrac{1}{2}\mathcal{R}_g, \tag{5.13}$$

$$\chi'_-(R) = 1. \tag{5.14}$$

Suppose $\chi(R) \leqslant 0$, then $\chi'' = (2\pi G/c^2)\rho\chi$ with $\rho \geqslant 0$ implies $\chi''(r) \leqslant 0$ for all $r \leqslant R$, with strict inequality if r lies in the support of ρ. Eqs. (5.13), (5.14) now imply that the curve $r \to \chi_-(r)$ lies below the curve $r \to r - \frac{1}{2}\mathcal{R}_g$ for $r \leqslant R$, which in turn implies $\chi(r = 0) < -\frac{1}{2}\mathcal{R}_g < 0$, where the last inequality just expresses the positivity of the gravitational mass. But this contradicts the regularity of the gravitational potential which requires a finite value of $\psi(r = 0)$ and thus $\chi(r = 0) = 0$. Hence we must have $\chi(R) > 0$ or, by (5.13), $\mathcal{R}_g < 2R$.

Finally we mention that in Ref. [2] solutions to (1.4) where found with $c^2\rho$ being given by the electrostatic field energy of a point charge. Rewritten in terms of ϕ these solutions obviously translate into solutions of (4.2). A very interesting observation of Ref.

[2] then was that the bare mass, measured by the flux of $(1/4\pi)\nabla\varphi$ through an infinitesimal small sphere centered at the charge, is negative and *finite*. Unfortunately this is no longer true in our theory where we have to use ϕ rather then φ in the expression for the bare mass. The simple calculation gives a negatively diverging result, the reason being the exponential increase of ϕ with φ. This implies that the finite bare mass obtained in Ref. [2] cannot be understood as a direct consequence of **P**, but so far only as an artifact of (1.4), which violates **P**. But **P** should certainly be satisfied in any model of general relativity, in particular if it is used to study questions concerning the energy regulating power of the gravitational field.

Appendix A

In this appendix we wish to give a short direct proof that (4.2) satisfies **P**, i.e. that $\delta A = c^2 \delta M_g$.

Using the generally valid Eq. (2.3), now with ϕ replacing φ, we must eliminate ρ via (3.3). This is most easily done if we set $\phi = c^2\psi^2$ and use (1.5). We obtain

$$\delta A = \frac{c^4}{2\pi G} \int_{\mathbb{R}^3} \psi^2 \, \delta \left[\frac{\Delta\psi}{\psi} \right] \, \mathrm{d}V$$

$$= \frac{c^4}{2\pi G} \int_{\mathbb{R}^3} \left[\psi\Delta(\delta\psi) - (\Delta\psi)\,\delta\psi \right] \, \mathrm{d}V \tag{A.1}$$

$$= \frac{c^4}{2\pi G} \int_{S^2_\infty} \boldsymbol{n} \cdot \left[\psi\nabla(\delta\psi) - (\nabla\psi)\,\delta\psi \right] \, \mathrm{d}o. \tag{A.2}$$

Now, the conditions for large r imply that $\nabla\psi$ falls off as fast as $1/r^2$ and $\delta\psi$ as fast as $1/r$. Hence the second term in the last bracket does not contribute. Therefore we may reverse its sign and obtain

$$\delta A = \frac{c^4}{2\pi G} \delta \int_{S^2_\infty} (\boldsymbol{n} \cdot \nabla\psi)\psi \, \mathrm{d}V$$

$$= \frac{c^2}{4\pi G} \delta \int_{S^2_\infty} \boldsymbol{n} \cdot \nabla\phi \, \mathrm{d}V = c^2 \delta M_g.$$

References

[1] R. Geroch, On the positive mass conjecture, in: Theoretical Principles in Astrophysics and Relativity, eds. N.R. Lebovitz, W.H. Reid and P.O. Vandervort (University of Chicago Press, Chicago, 1978).

[2] M. Visser, Phys. Lett. A 139 (1989) 99.

[3] T.R. Robinson, Phys. Lett. A 200 (1995) 335.

[4] A. Einstein, Zur Theorie des statischen Gravitationsfeldes, Ann. Phys. (Leipzig) 38 (1912) 443.