Quantization of gravitation

Jorma Jormakka
jorma.o.jormakka@gmail.com

Abstract: The article shows that the scalar theory of gravitation is the correct theory of gravitation and proposes that this scalar gravitational field is the massless scalar Goldstone boson field of the Higgs symmetry breaking method in the electroweak theory.

1. Introduction

Before Einstein presented his General Relativity Theory (GRT) Gunnar Nordström tried to formulate a relativistic scalar theory of gravitation. His starting points were two equations from Newton's gravitation theory. The first was the equation of motion:

(1) $\frac{d}{dt}u_a = -\partial_a \phi$, which follows directly by combining $ma = F$ and $F = -\nabla\phi$, $u(t,\bar{r})$ is the velocity of a test mass. The second was the field equation

(2) $$\Delta\phi = 4\pi G\rho .$$

This field equation can be explained by setting the gravitational potential as

(3) $$\phi = \phi(r) = -\frac{G\rho}{r},$$

where the mass density ρ is a constant. Then, as ϕ depends only on r

(4) $$\Delta\phi \equiv \nabla\cdot\nabla\phi = r^2\frac{d}{dr}\left[r^2\frac{d}{dr}\phi\right] = \frac{d}{dr}\left[r^2\frac{d}{dr}\left(-\frac{G\rho}{r}\right)\right] = \frac{d}{dr}[G\rho] = 0 .$$

The equation (2) is a generalization of (4). It may not be so obvious that if we let the mass density depend on r, $\rho = \rho(r)$, the mass density ρ in (2) and ρ in (3) are not the same function. Let

(5) $$\phi = \phi(r) = -\frac{G\rho_1(r)}{r},$$

then

(6) $$\Delta\phi = r^2\frac{d}{dr}\left[r^2\frac{d}{dr}\phi\right] = \frac{d}{dr}\left[r^2\frac{d}{dr}\left(-\frac{G\rho_1(r)}{r}\right)\right] = -G\frac{1}{r}\rho_1''(r)$$

where ρ_1'' is the second derivative of $\rho_1(r)$. Thus, (2) implies that

(7) $$\rho(r) = 4\pi r\rho_1(r) .$$

Using the same letter in (2) and (3) can cause some confusion. Nordström tried to find a relativistic form generalizing the equation of motion (1) and the field equation (2). Abrams had already suggested

(8) $$\dot{u}_a = -\partial_a\phi$$

for the equation of motion, but it did not work. Here the dot above the velocity component to the direction of the coordinate a means derivation with respect to the proper time, the time of the moving test mass. Nordström made two proposals. The first was

(9) $\quad \dot{u}_a = -\partial_a \phi - \dot{\phi} u_a \qquad$ and $\Box \phi = -4\pi G\rho$.

The D'Alembertian in Cartesian coordinates has the signature (+,-,-,-)

(10) $\quad \Box = \partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2$.

The second proposal was

(11) $\quad \phi \dot{u}_a = -\partial_a \phi - \dot{\phi} u_a \qquad$ and $\qquad \phi^{-1} \Box \phi = -4\pi G\rho$ or $\phi \Box \phi = -4\pi G\rho$

depending on the source. The Wikipedia [1] mentions the second version. In 1915 Einstein and Fokker published paper stating that Nordström's equation of motion in (11) comes from the geodesic Lagrangian of a curved Lorentzian manifold with $g_{ab} = \eta_{ab}\phi^2$. This is strange, because it does not. The geodesic equation is

(12) $\quad \ddot{x}^\mu + \Gamma^\mu_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta = 0$.

Here $\Gamma^\mu_{\alpha\beta}$ are the Christoffel symbols for the metric $g_{\alpha\beta}$, x^α is a coordinate and $\dot{x}^\alpha$ is the velocity in the inertial frame of the moving test mass. The geodesic can be derived by minimizing an action, but Einstein apparently stated that the geodesic Lagrangian is

(13) $\quad L = g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta$, which with the chosen metric is $L = \eta_{ab}\phi^2 \dot{x}^a \dot{x}^b$.

We can easily see that (12) does not come from (13). The Euler-Lagrange equations that apply here are

(14) $\quad \frac{d}{ds}\frac{\partial}{\partial \dot{x}^\mu} L(s, x(s), \dot{x}(s)) - \frac{\partial}{\partial x^\mu} L(s, x(s), \dot{x}(s)) = 0$

as $\frac{\partial}{\partial \dot{x}^\mu} L(s, x(s), \dot{x}(s)) = \eta_{\mu\mu}\phi^2 2\dot{x}^\mu \quad$ and $\frac{\partial}{\partial x^\mu} L(s, x(s), \dot{x}(s)) = \eta_{\alpha\beta}(2\phi \partial_\mu \phi)\dot{x}^\alpha \dot{x}^\beta$

equation (14) takes the form

(15) $\quad \phi^{-2} \frac{d}{ds}\left(\phi^2 \dot{x}^\mu\right) + (-\eta^{\mu\mu}\eta_{ab}\phi^{-1}\partial_\mu \phi)\dot{x}^\alpha \dot{x}^\beta = 0 \quad$ i.e.,

$$\ddot{x}^\mu + (-\eta^{\mu\mu}\eta_{ab}\phi^{-1}\partial_\mu \phi)\dot{x}^\alpha \dot{x}^\beta = -2(\phi^{-1}\dot{\phi})\dot{x}^\mu .$$

Assuming that ϕ does not depend on the proper time s we almost have the geodesic equation (12), but not quite and it is not the sign only: though the terms

$$\Gamma^\mu_{\mu\mu} = \eta^{\mu\mu}\eta_{\mu\mu}\phi^{-1}\partial_\mu \phi \text{ and } \Gamma^\mu_{\alpha\alpha} = \eta^{\mu\mu}\eta_{\alpha\alpha}\phi^{-1}\partial_\mu \phi$$

are correct, there are nonzero terms in (12) that do not appear in (15)

$$\Gamma^\mu_{\alpha\mu} = \eta^{\mu\mu}\eta_{\alpha\alpha}\phi^{-1}\partial_\alpha \phi \neq \eta^{\mu\mu}\eta_{\alpha\alpha}\phi^{-1}\partial_\mu \phi \text{ if } \mu \neq \alpha .$$

Thus, the Lagrangian (13) is not quite correct and we cannot use it to derive the geodesic (12). Additionally, (15) is not the equation of motion in (11) or (9).

Einstein was correct in stating that the equation of motion Nordström was looking for is the equation of a geodesic of a Lorentzian manifold with the metric $g_{ab} = \eta_{ab}\phi^2$. As Nordström was trying to create a scalar field theory, the metric must be $g_{ab} = \eta_{ab}\phi^2$. It is because a scalar field is a conformal mapping and at each point all sides dx_0, dx_1, dx_2, dx_3 of a small cube are multiplied by the same number $\phi(x)$. The correct equation of motion follows easily the equivalence principle: all masses fall with the same speed (accelerate equally much, or a particle falling in a gravitation field does not have acceleration in the inertial frame of reference), as Galileo Galilei had noticed long ago. A mathematical consequence of the equivalence principle is that test masses move along geodesics.

According to [1] Einstein is said that Nordström's second field equation, in (11), comes from the Lagrangian

(16) $$L = \frac{1}{8\pi}\eta^{ab}\phi_{,a}\phi_{,b} - \rho\phi .$$

It does not. The Euler-Lagrange equation in this case is

(17) $$\partial_{\mu}\left(\frac{\partial}{\partial\phi_{,\mu}}L(\phi,\phi_{,a})\right) - \frac{\partial}{\partial\phi}L(\phi,\phi_{,a}) = 0$$

$$\frac{\partial}{\partial\phi_{,\mu}}L(\phi,\phi_{,a}) = \frac{1}{4\pi}\eta^{\mu\mu}\phi_{,\mu} \quad \text{and} \quad \frac{\partial}{\partial\phi}L(\phi,\phi_{,a}) = -\rho .$$

Thus, equation (17) gives

(18) $$\frac{1}{4\pi}\eta^{\mu\mu}\phi_{;\mu} + \rho = 0 \quad \text{i.e., } \Box\phi = -4\pi\rho .$$

This is Nordström's first field equation. Einstein correctly told that the metric $g_{ab} = \eta_{ab}\phi^2$ yields

(19) $$R = -6\phi^{-3}\Box\phi \ ,$$

where R is the Ricci scalar curvature. Consequently, (19), the field equation of Nordström's second theory, can be obtained form the Lagrangian

(20) $$L = \eta^{ab}\phi_{,a}\phi_{,b} - \frac{R}{12}\phi^4 .$$

This is interesting as a corresponding scalar quantum field is renormalizable. Einstein also gave the following (heuristic) expression to Nordström's second field equation

(21) $$\phi\Box\phi = -4\pi G T_{matter}$$

where T is the trace of the energy-stress tensor. By (19)

(22) $$\phi\Box\phi = -\frac{1}{6}R\phi^4 = -4\pi G T_{matter} .$$

Thus $R = 24\pi T$ and $T = G\phi^{-4}T_{matter}$ gives the desired form of the trace of the energy-stress tensor. These considerations were later important in the development of the General Relativity theory. Einstein convinced Nordström to include the energy-stress tensor to the theory, which is quite good as the emphasis on the four diagonal Ricci entries R_{aa} makes it clear that there are four equations. It may initially look like there is only one equation in (19) and that there could be linear combinations of the four terms $g^{aa}R_{aa}$ which are summed to R, but this is not the case. There are four equations because in Cartesian coordinates all R_{aa} for $a > 0$ have symmetric forms. That is, in R_{aa} the coordinate a is in a special position, while R_{bb} has the same form with a changed to b. The time coordinate $a = 0$ is a bit different because of the signature (+,-,-,-), but taking this into account, all four coordinates are symmetric in form. There are four functions to solve, $\partial_a \phi$, and four equations. (Naturally the values these functions take need not be symmetric. The field ϕ takes non-symmetric values if the mass distribution is non-symmetric.) In the case of the scalar field the Ricci entries R_{ab} are automatically zeros if $a \neq b$ in orthogonal coordinates. This means that (19) has all ten equations but only the four diagonal entries need to be solved and their solutions follow directly from the solution to the sum equation (19) and using the expressions for R_{aa} for the scalar metric. I calculated these entries for Cartesian and spherical coordinates in [2]. Einstein's General Relativity has ten equations for four unknowns.

2. Why the scalar gravitation theory is the correct one?

Already in 1915 it was known to Einstein, and after he published his 1915 paper it must have been clear to his close co-worker Nordström, that the equations of motion in (9) and (11) are incorrect and that the correct equation of motion in the scalar theory is the geodesic (12). It is therefore curious that some evaluations of the scalar theory, dubbed Nordström's (incorrect) theory of gravitation, take one of the equations of motion (9) or (10), or the (incorrect) geodesic Lagrangian (13) and derive a result from it that conflicts with experiments.

We can see how the Wikipedia experts evaluate the scalar theory in [1]. The calculations in [1] make use of the general solution of ϕ in vacuum, which is derived from the field equation in vacuum:

(23) $\quad \Box \phi = 0$.

In theory one can proceed like this, but I doubt that the general solution in [1] is correct. The vacuum solution for ϕ can be obtained quite easily. First notice that

(24) $\quad \phi(\bar{r}) = -\dfrac{1}{\|\bar{r} - \bar{a}\|}$, where $\bar{a}$ is a constant vector,

is a solution of the time independent equation

(25) $\quad \Delta \phi = 0$.

Let $\bar{a}$ be a function of time. Then $\bar{a}(t)$ describes a path in the 3-space. We select this path to be a geodesic so that $\bar{a}(t)$ is a path of a test mass falling in the gravitation field. In local coordinates the test mass does not feel any gravitation field, thus $\Box \phi = 0$ in the moving coordinates. Making a linear coordinate change we get the solution in the original coordinates. In the original coordinates we also must have $\Box \phi = 0$ for this solution of $\bar{a}(t)$. This is the explanation why Einstein stated that the equation of motion (the geodesic) follows directly from the field equation for the vacuum. Yet, it is very difficult to solve the field equation for the vacuum directly in the time-dependent case: the equation for $\bar{a}(t)$ becomes nonlinear and complicated. [1] claims to have the general solution to the time dependent case from the field equations, but I am not convinced, and for that reason not convinced of the evaluation in general.

The scalar theory passes the redshift and Pound-Rebka experiment tests because it fulfills the equivalence principle by construction: the equation of motion is a geodesic. It must be noted that the equation of motion in such a calculation must be (12), not (9) or (11), and that the geodesic Lagrangian is not (13). It seems that in [1] in the analysis of the Shapiro delay and in the precession of periastria for the scalar theory the authors derive the geodesic from (13).

Earlier I have seen the following argument against the scalar theory: that light does not bend in Nordström's theory of gravitation. This may be so if the equations of motion are taken from (9) or (11), but if the equation of motion is the geodesic equation (12), then all test masses move along geodesics. Nordström's scalar theory does not consider electromagnetism, but photons can be treated as a limit case when the rest mass goes to zero. Photos have moving mass and therefore pass as test masses, so they move along a geodesic and bend in the gravitation field in the scalar theory.

We may ask what possibly could be found from physical experiments that could discard the scalar theory and the General Relativity. If the metric of our universe is $g_{ab} = \eta_{ab}\phi^2$ then both theories give exactly the same result, but Einstein does not accept the result. In both theories the Ricci off-diagonal entries R_{ab}, $a \neq b$, get the value zero and the diagonal entries get the values for this metric (see the values e.g. from [2]). When Einstein's equations

(26) $$R_{ab} - \frac{1}{2} R g_{ab} = k_0 g^{ab} T_{ab} + \lambda g^{ab} g_{ab}$$

are summed over the diagonal entries, the equation obtained for the metric $g_{ab} = \eta_{ab}\phi^2$ is certainly the same one as for Nordström's theory:

(27) $$\phi^{-3} \Box \phi = -\frac{1}{6} R = \frac{1}{3} k_0 T + \frac{\lambda}{3} .$$

Ignoring the cosmological constant that Einstein was not sure of, (27) is just a form of (21)

$$\phi \Box \phi = -4\pi G T_{matter} \quad \text{i.e.,} \quad \phi^{-3} \Box \phi = -4\pi T \ \text{ as } T = G\phi^{-4} T_{matter} .$$

Both of these theories have ten equations, which in the case of $g_{ab} = \eta_{ab}\phi^2$ reduce to four diagonal equations. In both theories test masses move along geodesics. In order to find a measurable error, the universe should have some other metric than $g_{ab} = \eta_{ab}\phi^2$. The difference between the scalar theory and General Relativity is only that Einstein would not have accepted the metric $g_{ab} = \eta_{ab}\phi^2$. Einstein believed that he knows what the entries of the matrix T_{ab} should be and that especially in the empty space all entries T_{ab} should be zero. For the metric $g_{ab} = \eta_{ab}\phi^2$ the Ricci entries R_{aa} are not zero.

But as it is, we live in our small piece of the physical universe and we know for quite a long time that the gravitational potential in this world is closely approximated by the Newtonian time-independent gravitational potential (3), even if Einstein did not like it. This potential is a scalar field and the metric for the scalar field is $g_{ab} = \eta_{ab}\phi^2$. The Ricci entries in Cartesian coordinates get the values (see [2])

(28) $$R_{00} = \frac{1}{r^2} \ , \ R_{jj} = -\frac{1}{r^2} + 8\frac{1}{r^2}\left(\frac{\partial r}{\partial x_j}\right)^2 - 2\frac{1}{r}\frac{\partial^2 r}{\partial x_j^2} \ \text{ for } \ j = 1,2,3 .$$

The Ricci scalar curvature is zero:

(29) $$R = g^{ab} R_{ab} = \eta^{00} \phi^{-2} \left(R_{00} - \sum_{j=1}^{3} R_{jj} \right) = G^{-2} \rho^{-2} r^2 \left(R_{00} - \sum_{j=1}^{3} R_{jj} \right)$$

$$= G^{-2}\rho^{-2}r^2\left(\frac{1}{r^2}+\frac{3}{r^2}-8\frac{1}{r^2}\sum_{j=1}^{3}\left(\frac{\partial r}{\partial x_j}\right)^2+2\frac{1}{r}\sum_{j=1}^{3}\frac{\partial^2 r}{\partial_j^2}\right)$$

$$= G^{-2}\rho^{-2}r^2\left(\frac{1}{r^2}+\frac{3}{r^2}-8\frac{1}{r^2}+2\frac{1}{r}\frac{2}{r}\right)=G^{-2}\rho^{-2}(1+3-8+4)=0\,.$$

The Ricci entries R_{aa} are not zero. It should be possible to measure the gravitation potential in the space on a satellite orbit and check whether (3) is a good approximation to the potential. I assume this has been done, as (3) is the expression for the gravitational potential used in practical physics and engineering and it cannot be much wrong. Einstein believed that all engineers are wrong and he is correct. He believed that all Ricci entries should vanish in a satellite orbit and gravitation is not described by the scalar gravitational potential (3). Instead, in Einstein's opinion, the geometry of the space is deformed: small balls are not round, light has a different speed to different directions, and that the solution to gravity in a satellite orbit is the Schwarzschild solution.

I made a simple evaluation in [2] following the calculations that supposedly show that the scalar theory contradicts measurements, and I did not find experiments where the scalar theory fails. Instead, the Schwarzschild solution for the General Relativity in the vacuum failed the Shapiro delay test. This is so because the ball in the Schwarzschild solution gets increasingly deformed and as a consequence the speed of light is different along different paths. A small ball (or a small cube, if you prefer) must keep its form if the speed of light is to be constant. The speed of light is the relation of dx^i and dx^0 for $i > 0$. The Ricci elements for the Schwarzschild solution are calculated in [2] and one can compare (28)-(29) to the way Ricci elements get zero values in the Schwarzschild solution.

On the Earth, normal Newtonian physics, with small corrections such as the redshift, works quite well. The space around the Earth is largely empty and R is zero in the space, or close to it. It is naturally possible that somewhere in the large universe there is a black hole where small balls of the geometry are not round and the solution is the one calculated by Schwarzschild. If so, then there is some place in the universe where the scalar theory does not hold and we must use Einstein's equations and allow for a more general metric. But this cannot be the case in the experiments of redshift, Shapiro delay and the movement of planets in our solar system that were used in the evaluation in [1].

I do not think there are experiments that can be made in practice and can demonstrate that the metric is not $g_{ab} = \eta_{ab}\phi^2$ and that the potential (3) is a poor approximation in the stationary vacuum case. Evaluations that have been made in order to show the scalar theory incorrect are based on the use of the wrong equations of motion in (9) and (11), the wrong Lagrangian in (13), and a wrong general solution to the field equations.

There are other reasons why the General Relativity is not correct. The General Relativity Theory cannot be formulated as a renormalizable quantum field theory, while the scalar theory is renormalizable (as mentioned e.g. in [3]). Especially GR cannot be turned into a renormalizable quantum gauge field theory, and there is reason to expect that all quantum field theories should be gauge field theories. Dietmar Ebert in [4] p. 53-57 there is one proposal how to turn GR into a (local) gauge field theory, but the result cannot be renormalized. There is a considerable similarity between the concepts of GR and gauge field theories. For a scalar field these similarities may be realized into new insight to the problem.

The book of supersymmetry and supergravity by Julius Wess and Jonathan Bagger ([5] p.4) starts by referring to the Coleman-Mandula theorem, which gives rather restrictive conditions to symmetries of the S-matrix. They release one condition and get the most general supersymmetric algebra on p. 8. They manage to formulate renormalizable supergravity models, but in Kähler, i.e., complex, manifolds. The real dimension of a 4-dimensional Kähler manifold is eight and they use supersymmetry. For these reasons supergravity is not a proper quantization of General Relativity. Yet, interestingly, their supergravity model includes scalar (chiral) fields, Kähler potentials.

Quantum gauge field theory even without supersymmetry is closer to the scalar theory of gravitation than to the General Relativity. In gauge field theory masses are created by spontaneous symmetry breaking induced by a scalar field. As the theory of gravitation is closely associated with masses, we should expect to find the mass creating scalar field having something to do with gravitation.

Measurements do not so far support the assumption that our universe has supersymmetry. A similar negative conclusion seems to be true with the superstrings theory. I have personally not studied superstrings (apart from briefly scanning a popular science book [5]) and cannot make my own judgment, but research in superstrings is already some 30 years old and has not led to a solution to quantum gravity. As for the scalar gravitation theory, it is well known (mentioned e.g. in [6]) that the theory is renormalizable and satisfies the strong equivalence principle. I found a further argument against the General Relativity in a most unexpected source: Martin Gardner [7] Chap. 17 tells how a number of physicists headed by J. A. Wheeler wrote a letter addressing ESP research mentioning that there are only three or four experiments that give support to General Relativity.

We can add to this very short list of experiments the recent finding of graviton, but it also is not a decisive proof of GR. What was found (unless the results are noise, as has been suggested) is an elementary particle with zero mass, zero charge, spin 2 and weak coupling with hadrons. In Einstein's theory such a particle is a graviton. However, it may be a gluon. A gluon coming from a distant star can have very high momentum and for that reason have a small coupling constant with fermions. The property of the color force, called asymptotic freedom - color confinement (see [8] p. 147) does not actually say that gluons are confined to a small distance. It is a relation with renormalization mass and the coupling constant and it says that gluons with high momentum interact weakly with fermions.

4. Cosmological implications

The main implication of the scalar theory of gravitation to cosmology is that the field equation is relatively simple, especially simple it is for the vacuum, but small amount of mass can be treated as a perturbation. It should be possible to derive results much easier in the scalar theory than in the General Relativity theory.

One slightly negative implication of accepting Nordström's scalar theory is that the Schwarzschild solution is not the correct one as a vacuum solution in our space-time. The Schwarzschild solution has been widely studied in the context of black holes, see [10]. If the geometry of our part of the universe is $g_{ab} = \eta_{ab}\phi^2$ it does not necessarily exclude the possibility that there may be somewhere black holes behaving as the Schwarzschild solution. Even if they do not exist, the study of such constructions remains interesting in the mathematical sense. The authors of [10] were not interested in massless scalar fields in gravitation. Indeed, p. 324 in [10] (equation 8.61) states that a massless scalar field ϕ obeys $\Box\phi = 0$ and that there are no massless scalar fields in the nature. Neither claim is correct: a massless scalar field can have a $\lambda\phi^3$ term, and such massless scalar fields appear in spontaneous symmetry breaking as Goldstone boson fields.

Another implication, coming directly from the conformal nature of scalar fields, is that the problem of inflation in cosmology can be easily explained. In the early universe all matter was in a small space and the gravitational field potential ϕ was very high. Consequently, the space element was very large (each dx^i gets multiplies by ϕ in the conformal mapping). It follows that even if we let the speed of light be the same constant, as the space element is very large, the expansion of the universe in the earliest picoseconds seems for us much faster than light. This is because the unit element for the time was also very large. We do not need to change the speed of light, as is done in the VSL theory of João Magueijo [11]. The conformal mapping also solves the eternal question: what was before the universe was born? If at the beginning all matter was concentrated in a very small space, even to a singularity, the time unit element was huge, even infinite. There was no time before the universe started expanding. In the original singularity time did not tick.

Some very simple conclusions can be made based on ρ and ρ_1 in (6). The scalar theory gives the same solution as the Newtonian theory if ϕ is time-independent. We can conclude that ρ_1 cannot decrease because the gravitational potential of an empty space with mass in the center is (3). If there is mass in the space, then the potential falls slower than (3). Let us have ρ_1 behave as r^α. Then the mass distribution ρ behaves as $r^{\alpha-2}$. Integrating ρ over a ball of radius r gives $\int dr 4\pi r^2 r^{\alpha-2} \propto r^{\alpha+1}$ assuming that the volume grows as in the flat space. Consequently, the mass goes to infinity if r can grow to infinity. The simplest explanation for this behavior is that r cannot grow to infinity because all mass is in a limited volume. The universe has a finite size.

A manifold can have a finite size without having a border, in which case it is closed. In 1980s it was fashionable to study such topological oddities that might have cosmological relevance. In a stationary case the manifold can be taken as a closed, orientable 3-manifold. If the metric is given by a scalar potential, the mapping is conformal outside isolated singular points. We can generalize the mapping to the extent that the unit ball can deform to a limited extent (the speed of light can vary within some limits, the Schwarzschild solution does not fill this condition). In that case the mapping is quasiregular. I classified in 1988 closed, orientable 3-manifolds that admit a quasiregular map [11]. There is a number of possibilities. These 3-dimensional spaces would then have (positive) Ricci curvature. However, the most natural choice is that our universe does have a border: if the space-time is almost flat, the universe is a ball B^3 with S^2 border. Topological oddities may only appear in areas of high gravitation potential, if at all.

3. Quantum gravity

After these easy preliminary comments we can move to quantization of gravity. One of the physicists signing the letter by Wheeler mentioned by Martin Gardner in [7] was Richard D. Mattuck. His excellent book [9] can serve as an introduction to the quantum theoretical setting, but it does not treat gauge fields. I am somewhat familiar with only two books that treat quantum gauge field theories, Wess-Bagger [4] and Bailin-Love [8]. No doubt there are better books today, but I stopped the study of this field in late 1980s. Neither one of these two books is ideal as a background text for quantizing the scalar theory of gravitation. Wess-Bagger is difficult to apply in the standard model as the setting is in supersymmetry, while Bailin-Love is difficult to read due to errors, starting from the first page. (On p.1 there is the $\ln \det A = Tr \ln A$, the symmetric matrix A must be diagonalized: $\ln \det A = Tr \ln U^T AU$ and in order to get the integral converge all eigenvalues must be nonnegative. The book has many other errors, especially in the beginning.) Nevertheless, these books do contain all necessary background for making certain observations of scalar fields. I will follow the notations of [8].

In quantum field theories the Lagrangian (20) would be written (see [8]) as

$$(30)\qquad L = \left(\partial_\mu \phi\right)\left(\partial^\mu \phi\right) - \frac{R}{12}\phi^4$$

and it is of the form of a Lagrangian of a renormalizable quantum scalar field

$$(31)\qquad L = \left(\partial_\mu \phi\right)\left(\partial^\mu \phi\right) - \mu^2\phi^2 + \lambda\phi^4$$

The field (31) has found an application as the Higgs field that is creating masses for elementary particles by spontaneous symmetry breaking in the electroweak gauge field theory.

The Higgs mechanism and spontaneous symmetry breaking with 't Hooft gauge for hiding the Goldstone boson field are all explained in [8] Chapter 13, so I will write very few formulae in this article, and I will make no new mathematics. I will only illustrate a possible approach.

There always appears at least one massless (complex) scalar field when gauge invariance is imposed to a Lagrangian density. Especially such a scalar field, a Goldstone boson field, appeared in the electroweak theory when the Higgs mechanism was used to break the chiral (actually parity according to [8]) symmetry. Massless scalar field was seen as undesirable in the theory. In a special gauge, the unitary gauge, the Goldstone scalar field is eliminated from the equations. In this gauge the massless scalar field does not appear in renormalization calculations and evaluation of Feynman propagators.

But the Goldstone boson field has not disappeared because of this mathematical trick, and it is not an artifact like Faddeev-Popov ghosts. The elimination of the Goldstone boson is only possible in the range where the massive Higgs scalar field is present. As mentioned in [4], the range of massive boson fields should be (roughly) inversely proportional to the mass. Therefore the Higgs massive scalar field cannot cancel the Goldstone massless scalar field in large distances. The massless scalar field must be visible in large distances. There are only two massless long range fields: the electromagnetic vector potential A_μ and the gravitation field. It is logical to expect that the massless Goldstone scalar field is the scalar gravitation field in Nordström's theory. As there are only two long range massless boson fields, we cannot expect to have more than one Higgs field. It follows that the approach of SU(N) GUT is hopeless: those theories have more gauge symmetries that need to be broken, and consequently they have more Goldstone boson fields.

There are some issues in interpreting the Goldstone boson field as the scalar field of gravitation. Firstly, we have to make the scalar field complex. This does not seem to be a problem. The phase of ϕ is cancelled in Cristoffel symbols and in the metric we can write $\phi^*\phi$ instead of φ^2. The Lagrangean density in (30 or (31) is easily written with a complex field.

The mass term $\mu^2\phi^2$ is missing from (30). The Higgs field has mass, but there is a Goldstone boson field that is massless. Thus, (30) can describe a Goldstone boson field. But there seems to be one problem. In [8] λ in (31) is positive, while R in (30) is also positive. The fact that R is positive is clear from (22). From (29) we see that R is of the order of $G^{-2}\rho^{-2}$. It seems to me that [8] has an error in this place. The choice of the coordinates in [8] is that $\Box = \partial_0^2 - \Delta$. Therefore the field ϕ is negative and it does not have a local minimum at the minimum energy place. It has a local maximum and λ must be negative. This also explains why the mass term $\mu^2\phi^2$ in [8] is negative. The number μ^2 should be positive as λ is negative and the Higgs field does not have an imaginary mass.

The issue with mass generation is briefly explained as follows. Consider the Lagrangian density of the Dirac field (describing a relativistic electron):

(32) $L = \overline{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi$ where $\overline{\psi} = \left(\psi^{*}\right)^{T}\gamma^{0}$ is the conjugate field.

This Lagrangian is gauge invariant under the global gauge transformation

(33) $\psi \to e^{-iq\lambda}\psi$, $\overline{\psi} \to e^{iq\lambda}\overline{\psi}$, where λ is a real number.

This fermion field is described by a spinor field and it has mass m. It is clearly an electron. The Lagrangian (32) is not invariant under a local gauge transformation where $\lambda = \lambda(x)$ depends on the coordinates. Local gauge invariance can be achieved by changing the Lagrangian to

(34) $L = \overline{\psi}\left(i\gamma^{\mu}D_{\mu} - m\right)\psi$ where $D_{\mu} = \partial_{\mu} + iqA_{\mu}$ and

(35) $A_{\mu} \to A_{\mu} + \partial_{\mu}\lambda$

The vector field A_{μ} is the vector potential of electrodynamics and it follows the Maxwell equations, which come from the Lagrangian

(36) $F_{\mu\nu}F^{\mu\nu}$ where $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$.

It is easy to check that the Lagrangian term (36) is invariant in the local gauge transformation (35). Electrons have mass and the Dirac field has a mass term $m\overline{\psi}\psi$ for the electron. This mass term is invariant in the transformation (33)

(37) $\overline{\psi}\psi \to e^{iq\lambda}\overline{\psi}e^{-iq\lambda}\psi = \overline{\psi}\psi$.

If the vector field A_{μ} had mass it should have a corresponding mass term of the type $mA_{\mu}^{*}A^{\mu}$ (here the conjugate is simply a complex conjugate, not a Hermitian conjugate of matrices). However, the mass term for the vector field is not invariant

(38) $A_{\mu}^{*}A^{\mu} \to \left(A_{\mu}^{*} + \partial_{\mu}\lambda\right)\left(A^{\mu} + \partial^{\mu}\lambda\right) \neq A_{\mu}^{*}A^{\mu}$.

As a result, the electromagnetic field and its field boson, photon, is massless. In a similar way the theory of strong interactions, QCD, is a gauge theory with a local gauge symmetry. The field bosons, gluons, are massless. There is a difference with the fermions of that theory. The fermions are quarks, which do have mass. However, the gauge group of QCD is non-Abelian and the mass terms of fermions in QCD transform as

(39) $\overline{\psi}\psi \to e^{iq\lambda_a T_a}\overline{\psi}e^{-iq\lambda_a T_a}\psi$.

The difference is that instead of a real number λ there is a more complicated exponent. The terms in (39) do not commutate and we cannot move the exponents next to each other to cancel. The result is that fermion mass terms break the gauge symmetry. It follows that mass terms cannot appear in the Lagrangian: for the purposes of the color force of strong interaction, quarks are massless. We see that QED, the theory of electromagnetism, has no problem with fermion masses because the gauge group U(1) is Abelian. The theory of strong interactions, QCD, would have a problem with fermion masses as the gauge group is non-Abelian, but the color force can ignore the masses. In both of these theories boson fields (the interaction fields) are massless.

There remains the electroweak theory, which solved the problem of chirality violation in weak interactions by introducing the Higgs mechanism of spontaneous symmetry breaking.

(The mechanism is well explained in [4], [5] and [8] and there is no reason to describe it here.) This mechanism creates masses to the interaction particles. The mechanism also creates masses for fermions by spontaneous symmetry breaking. What happens is that the original massless field gets mixed with the massive Higgs scalar field. The original massless field has only two transversal modes. The mixed field has three modes: two transversal and one longitudinal. In a sense the mixed field has three dimensions and the original massless field has two dimensions. We get to a philosophical conjecture: maybe massive field must have volume and a field that does not bound volume is massless. It would support the basic view of Einstein that mass is in some sense geometry.

References:

[1] The Wikipedia page on Nordström's gravitation theory (as 23.2.2020)
https://en.wikipedia.org/wiki/Nordstr%C3%B6m%27s_theory_of_gravitation

[2] J. Jormakka: Einstein's Field Theory is Wrong and Nordström's Correct
available in the vixra archive https://vixra.org/abs/1812.0067
and in the General Science Journal webpages:
https://www.gsjournal.net/Science-Journals-Papers/Author/2321/Jorma,%20Jormakka

[3] Nathalie Deruelle, "Nordström's scalar gravitation theory and the equivalence principle," arXiv: 1104.4608, 2011.

[4] Dietmar Ebert, *Eichentheorien, Grundlage der Elementarteilchenphysics*, Akademie-Verlag, Berlin, 1989.

[5] Julius Wess, Jonathan Bagger, *Supersymmetry and Supergravity*, Princeton Univ. Press, New Jersey, 1992.

[6] Brian Greene, *The Elegant Universe, Superstrings, Hidden Dimensions and the Quest for the Ultimate Theory*, W. W. Norton &Co., 1999.

[7] Martin Gardner, *Science good, bad and bogus*, Polish ed. Pandora, 1994 (original 1981).

[8] David Bailin, Alexander Love, *Introduction to Gauge Field Theory*, IOP Publishing, 1986.

[9] Richard D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem*, 2. ed. Dover, New York, 1992 (original 1967).

[10] Kip. S. Thorne, Richard H. Price, Douglas A. Macdonald, *Black Holes the Membrane Paradigm*, Yale University, 1986.

[11] João Magueijo, *Faster than the speed of light*, Penguin books, 2003.

[12] Jorma Jormakka, *The existence of quasiregular mappings from* R^3 *to closed orientable 3-manifolds*, Ann. Acad. Sci. Fenn. Series A, 69, Finnish Academy of Science, 1988, p. 1-44.