General Relativitiy
Einstein attempts at a theory of gravity

0.1 Variable light velocity

One of the big lacunna in Einstein's theory of special relativity was Newton's theory of gravity. That theory could be encapsulated in the equations for the gravitiational potential

$$m\frac{d^2\vec{x}}{dt^2} = -m\vec{\nabla}\phi(\vec{x}) \tag{1}$$

$$\nabla^2\phi = 4\pi G\rho(\vec{x}) \tag{2}$$

where $\rho)\vec{x}$ is the Newtonian potential. This theory clearly does not treat time and space in a similar manner. It is also a kinematic theory– ie, the equations of motion of the particle, because mM cancels out on the two sides of the first equaiton, means that motion of the particle does not depend on the structure of the particle. It depends only on its position and changes in its postion. This "Gallilean" invariance (since it had been Gallileo who noted that the rate of fall of an object was independent on the composition of the object) of gravity was one of the key features that Einstein grabbed onto.

Until 1908, busy with other consequences not only of special relativity, and quantum mechanics (the particle nature of light) and statistical physics, he did not have a way of understaing how to attack the problem. He came up with the elevator analogy (a person in and elevator would not be able to tell the difference between the elevator being near the surface of body, like the earth, of being in an elevator which was being accelerated upwards), suggested that somehow the essense of gravity had something to do with changes of frame, from unaccelerated to accelerated frames. In 1908 Minkowski had noted that all of special relativity could be encompassed in the notion that, instead of regarding time and space separately, they should be combined into a notion of space-time, which time being just another direction in the 4 dimensional space-time and that distances should be generalised into distances in space-time with a metric

$$ds^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2) \tag{3}$$

All of the Lorentz transformation could then simply be regarded as the transformation of of the coordinates t, x, y, z which left the form that distance function the same. Just as rotations leave the form of spatial distances the same, Lorentz transformations leave this 4 dimensional form of 4 distances the same.

Dismissing this as silly mathematization of the theory without physics, Einstein rapidly changed his mind to seeing this metric (the distance function) as being a possible road to making a theory of gravity.

His first attempt was– maybe the velocity of light in special relativity is not constant everywhere, but could change from place to place. Ie, one could have a metric

$$c_0^2 ds^2 = c^2(\vec{x})dt^2 - dx^2 - dy^2 - dx^2 \tag{4}$$

Then a straight line, a geodesic as people like Riemann had considerd it, would be the motion of the particle without extenal forces acting on it. Regarding gravity as not being a force (since forces affect particles via F/m and depend on the mass of the particle (and why should the force, an outside agent, know anything about the particle). However, straight lines, as Newton had taught us, do not depend on outside agents, and are the motion in the absense of forces. The geodesic equation of this theory, if we make the simplifying assumption that the vecity of light depends only on $\vec{x}$ is

$$\frac{d}{d\tau}\left(\frac{2c^2(\vec{x})}{c_0^2}\frac{dt}{d\tau}\right) = 0 \tag{5}$$

$$\frac{2}{c_0^2}\frac{d^2\vec{x}}{d\tau^2} + \nabla\frac{c^2(\vec{x})}{c_0^2}\left(\frac{dt}{d\tau}\right)^2 = 0 \tag{6}$$

where τ is the proper time of the particle. If the velocity of the particle is small ($\frac{dx}{d\tau} << c_0$) then $\frac{dt}{d\tau}$ is very close to 1, and the second equation becomes just Newton's equation of motion is $\vec{\nabla}c^2 = \phi$ the Newtonian potential. Thus if we take $c^2(\vec{x}) = c_0^2 - 2\phi(\vec{x})$, and take $s \approx t$ then we have Newton's equation of motion. Furthermore, The first equation gives us that

$$\frac{dt}{d\tau} = \frac{E}{c^2} \tag{7}$$

where E is some constant very close to 1 (say $1 - epsilon$) , and if $\frac{\phi}{c_0^2} << 1$

then the equation

$$c_0^2 = c^2(\frac{dt}{\tau})^2 - \frac{d\vec{x}}{ds} \cdot \frac{d\vec{x}}{ds} \quad (8)$$

$$= \frac{E^2}{c^2} - \vec{u}^2 \approx \frac{E^2}{c_0 - 2\phi} \quad (9)$$

we get

$$\epsilon \approx \frac{1}{2}(v^2 + 2\phi(x)) \quad (10)$$

which is the conservation of energy for the Newtonian system.

Ie, to a reasonable approximation this gives one Newton's equations of motion.

This also produced another surprizing result. If we look at the equation for a light ray (whose proper length of its "velocity" is not 1, but is rather 0,) we get

$$0 = \frac{E^2}{c^2(x)} - |\vec{u}|^2 \quad (11)$$

where we take $E = c_0^2$. (Ie, at infinity, we choose τ to be equal to t) since i the parameterization of the geodesic can be rescaled by an arbitrary constant. Going into polar coordinates, rather than cartesian coordinates, we have

$$\frac{d\theta}{d\tau} = \frac{l}{r^2} \quad (12)$$

where l can be considered as the angular momentum per unit mass of the particle, which is also conserved. As in Newtonian theory, we can look at the energy equation as

$$0 = \frac{}{c^2(r)} - (\frac{dr}{d\theta})^2(\frac{d\theta}{ds})^2 - r^2(\frac{d\theta}{ds})^2 \quad (13)$$

$$= \frac{E^2}{c^2(r)} - l^2\left((\frac{dr}{d\theta})^2 + \frac{l^2}{r^2}\right) \quad (14)$$

$$\approx l^2\left(\frac{c_0^2}{l^2c^2(\frac{1}{u})} + (\frac{du}{d\theta})^2 + u^2\right) \quad (15)$$

Defining $u_0 = \sqrt{\frac{E^2}{l^2 c_0^2}}$ we find

$$u = u_0(sin(\theta) + \frac{GMu_0}{c_0^2})) \tag{16}$$

Since $u = 0$ corresponds to $r = \infty$, this says that the orbit of a light ray comes in from infinity at an angle of $\frac{-GMu_0}{c_0^2}$ and leaves at an angle of $\frac{\pi+GMu_0}{c_0^2}$. Were M=0, the total angle would have differed by π. This means that the deflection would be greater than π by $\frac{2GMu_0}{c_0^2}$. Ie, a light ray would have been deflected by the a massive body by an amount which depends on the mass and the inversely as the radius $\frac{1}{u_0}$ of closest approach.

The gravity could affect light had been thought of by Cavendish (unpublished) and by Soldner in the early 1800s. Both of their ideas were based on Newton's theory of light as light being massive particles. Wave theories of light had originally been developed by Huygens at the time of Newton, burried under Newton's dominance of physics, been revived by experiments in the early 1800s, and solified by Maxwell's theory in the mid 1800. The Cavendish/Soldner arguments failed for a wave theory. However, for a metric theory, the results for the deflection of light for wavelengths much less than the scale of change of the gravitational field arethe same for a wave or a particle theory.

Experiments were almost immediately set in motion after Einstein's prediction. But clouds and war meant that none gave results. See Lus C. B. Crispino& Santiago Paolantonio "The first attempts to measure light deflection by the Sun" Nature Astronomy 4 pp6-9 (2020)i https://www.nature.com/articles/s41550-019-0995-5 for an account of the attempts.

For example, an experiment was arranged in 1914 by Freundlich from the Berlin Observatory to the Crimea, where a solar eclipse was to occur. Unfortunately, shortly before the eclipse, WW1 broke out, and, since Russia and Germany were on opposite sides, Freundlich was arrested as a spy, and his telescope and equipment was confiscated. (Since the director of the Berlin observatory was not enamoured of Einstein's speculations, he was upset that not only had he lost an employee, but also had lost one of his expensive telescopes. Both were returned at the end of the war.)

While it had some real successes, Einstein was not happy with this theory. It favoured one particular idea of time. What the equations for the velocity of light were, given a certain matter distribution, was also obscure to him. Also, from $E = Mc^2$, and from the fact that Energy changed in special relativity if one went into some other frame of reference, how to write equations for how c^2 changed from frame to frame were also obscure to him.

0.2 Nordström

In 1912 Nordstrom published a theory which was the "simplest" generalisation of the Newtonian theory. Taking ϕ as the equivalent of the Newtonian potential, he postulated a field equation

$$\partial_t^2 - \nabla^2\phi = -4\pi Grho \tag{17}$$

Defining the proper length $u^i = \frac{dx^i}{ds}$ where s is the proper length along the path, (and thus $u^i u_i = 1$) the equations of motion for a non-relativistic particle are

$$\frac{d^2u^i}{ds^2 = \eta^i j\partial_i\phi - u^j\partial_j\phi u^i} \tag{18}$$

where the extra term is to ensure that

$$0 = \frac{deta_{ij}u^iu^j}{ds} = 2\eta_{ij}u^i\frac{du^j}{ds} \tag{19}$$

where η_{ij} are the components of Minkowski metric given by

$$ds^2 = c^2dt^2 - d\vec{x}\cdot d\vec{x} \tag{20}$$

The last term term in the force equation is required to satisfy the fact that the proper velocity has unit length.

Einstein objected because this was a linear equation in ϕ and one would expect the gravitational field itself to have gravitational energy. Furthermore, under a Lorentz transformation, ρ did not change, whereas the energy or energy density would change in general. Ie ρ was a scalar, which could only be the proper mass of the source.

Nordström therefore altered his equation to make it non-linear which Einstein showed was equivalent to the a metric theory where

$$ds^2 = e^\psi(c^2dt^2 - d\vec{x}\cdot d\vec{x}) \tag{21}$$

with the equation

$$= 4\pi GT \tag{22}$$

where T was the trace of the energy momentum tensor (to be described later)

Einstein was excited by this theory and spent some time on it. However Einstein quickly realised that this theory had troubles. The electromagnetic field had an energy momentum tensor whose trace was zero. Ie, electromagnetism had no graviational field in this theory. It obeyed Newton's equations of motion for a graviational field with ψ the Newtonian potential, if the velocities were slow, for the same reason as the c^2 theory did (since only the metric term multiplying dt^2 in the equation of motion of a free particle gave non-trivial contributions to the effective force. In addition, calculating the deflection of light in this theory gave zero, which he felt from his "elevator" argument was wrong. (This reason was nonsense, since it satisfies the elevator argument under the conditions of the elevator argument. In this theory, there are two sources for the deflection. One is the from the temporal part of the metric, which is the same as in the c^2 theory, and the other is that the spatial straight lines – due to the spatial part of the metric also contribute to an equal but opposite deflection of the light. The elevator experiment is a local experiment in space, and could not mimic the spatial bending effect due to the change in the spatial part of the metric.

Having abandoned Nortstöm's idea, he went back to looking for field equations for the metric which would have the full energy-momentum tensor as a source. He then fell into a number of traps, some of his own making and stubborness, some rather deep intellectual ones which have ensnared many mathematicians and physicists since his time.

0.3 Entwurf

Already in 1912, he thought of looking at the curvature tensor R_{ij} for his equaitons. R_{ij} was a tensor, and thus defined independently of the coordinate system. In components it was something that had at most second derivatives of the metric, like all of the other field equations he knew in physics. So he first looked at the equations in the absense of any matter source. He chose $R_{ij} = 0$. $R_{ijkl} = 0$ has as its only solution flat Minkowski spacetime. $R = 0$

gives far too few equations (There are 10 components to the metric, and one equation be far too few to find solutions)

He found that there were 4 equations which had second time derivatives of different components of the metric, not just one that he would have expected, which seemed to say that in the Newtonian limit there should be four gravitational potentials not just one. (the extra three could be eliminated by an appropriate choice of coordinates, as he discovered later).

Finally he came up with the "hole" argument which convinced him that no covariant theory could ever be an appropriate theory of gravity. The argument goes something like the following:

Consider a solution of the equations of motion by finding some given intial data (for example specifying the metric and its first time derivative) on some spatial surface. One should be able to solve the equations for a unique soltion. But now, to the future of that spatial surface you look at a limited volume in spacetime (the "hole"), which is all to the future of that intial timelike surface. Inside that volume you change the coordinates. The metric components will change within that volume. And because the field equations are generally covariant, the tensor components of say R_{AB} will still be 0. The equations will be satisfied for these new metric components. The metric will exatly the same initial data. Thus your initial data will not have a unique set of solutions. There will be many solutions for the same intial conditions. Physics will be indeterminate. The future is not simply the developement of the past but there will be many futures consistant with the same present. Aaaaargh.

So it was back to square 1. He could not use tensor equations. They would all run into this problem, since they would also always allow coordinate transformations in a limited region and thus would run into the hole argument. He and Besso then worked to find new equations, and did. This was his Entwurf theory, which he then spent about 2 years on. It was a theory which had only a limited number of coordinate transformations. He calculated the perihelion advance of Mercury, and found about a 18" advance per century, instead of the measured value of 43" per century. (yes he already had hopes that a new theory of gravity would solve the puzzle of why Mercury's elliptical orbit had this excess of 43" per century.) However, Droste, a graduate student, published first, so they never pubished. The theory of Nordstöm's gave -7" per century .

He also checked as to whether or not his Entwurf equations equations would have the rotating metric as a solution. He found they did. However, he made a number of blunders in his proof (This has only come out in the past 20 years or so.) The metric has the form

$$ds^2 = dt^2(1 - \omega r^2) - 2dtd\tilde{\phi}(\omega r^2) - dr^2 - r^2(d\theta^2 + sin(\theta)^2 d\tilde{\phi}^2) \tag{23}$$

which is obtained from the usual spherical metric by replacing $\phi = \tilde{\phi} + \omega t$ He then took $g_{t\tilde{\phi}} = g_{\tilde{\phi}t} = 2\omega r^2$. (actually he did this in cartesian rather than polar coordinates, which makes the calculation messier but the idea is the same). Plugging this into the Entwurf equations, he found that it solved the equations. If he had done it properly, they would not have. Besso, whith whom he was collaborating, questioned his derivation and he ignored him.

In late 1915 he finally came back to his generally covariant ideas. He finally relooked at his "rotating" coordinates calculation and realised Besso was right. He rethought his hole argument, and realised that, yes, if you changed coordinates, you got a different metric, but this would not change the physics. Just as in Cartesian coordinates, you could add 5 to the coordinates of a particle, and the particles "position" (the value of the coordinate the particle was sitting at) would change, but that did not mean that the particle had moved. This would not change the physics. The coordinates could change. The metric components could change, but the physics would not change.

So $R_{AB} = 0$ could be the equations of motion for gravity. At first he suggested $R_{AB} = 8\pi G T_{AB}$ could be the equations, and then realised that while the right side was conserved, the left was not. And he quickly changed the equaitons to

$$G_{AB} = 8\pi G T_{AB} \tag{24}$$

in a paper he submited on Nov 25 1915. He showed that the deflection of light in this theory was twice what it was in his c^2 theory (1.7" at the edge of the sun). We now know that this was because in this theory, both the c^2 type deflection due to the change in the metric of time and the change in spatial straight lines due to the chaged spatial matric combined in the deflection of light, unlike in Nordström's theory, where they subtracted. He also quickly calculated the perihelion precession of Mercury's orbit and found that the theory predicted 43" of arc. (He with Besso had done this for the Entwurf theory, so the calculation was an duplicate of that calculation.)

He was home.

0.4 Hilbert

In Nov 1915, Einstein went to Goettingen to give a lecture on his then new theory (not yet General Relativity). In the audience was Hilbert, the greatest mathematician of the time, who had been gotten interested in the problem of Relativity and Gravity, and began thinking about the problem. He has also become interestes in some iseas of Mie as to how one might be able to incorporate electromagnetism into a geometric theory. He was also worried about the Hole argument. Ignoring that on Nov 20 he submitted a paper, which was finally published about a year later.In that paper he suggested that one could set up a theory of gravity by looking at the action He suggested that the theory should have an action of

$$I = \int R\sqrt{|detg|}d^4x \tag{25}$$

where $\sqrt{|detg|}d^4x$ is the covariant volume element, and R is the Ricci scalar. Taking the variation of this with respect to the metric components, one gets $G^{\mu\nu}$. This has led a long history of claims that Hilbert got the field equations before Einstein did. In 1997, L. Cory, a student of Juergen Renn in Germany dicovered the page proofs of Hilbert's paper in the Hilbert archive. On the published version, only the initial submission date of Nov 20 is listed. However, the page proofs show that that the original paper was substantially different from what was actually published. The action was there, but no equations were in that original paper (although a half a page is missing) before the action . In the published paper, there is a statement that the equations derived from the action are essentially those of Einstein. It is also clear that Hilbert was still very confused by the equivalent of the Hole argument. I think that claims that Hilbert discovered the equations before Einstein are overblown. (See for example the extensive paper https://arxiv.org/pdf/1201.5353.pdf Galina Weinstein "FROM THE BERLIN "ENTWURF" FIELD EQUATIONS TO THE EINSTEIN TENSOR II: November 1915 until March 1916" (2012)

Regarding the problems of the indeterminacy of the equations, If one fixes a coordinate system (giving some 4 equations that the metric compnents are supposed to obey for example, such that one cannot make any coordinate transformations) the equations do turn out to be causal. The past does determine the future.

Einstein and Grossman published a paper already in 2012 in which they essentially lighted on the Ricci curature tensor as perhaps being the appropriate aspect of the geometry which should be equal to some aspect of the energy momentum tensor. Unfortuately they got confused, and saw that the equations did not simply reduce to Newton's potential in the weak field limit. Apparently according to Norton, they also expected that in the weak field limit, the spatial metric should be flat space, and the equations did not say that. Einstein also got confused by the Hole argument.

The Hole argument is that the equations should be of the intial value form, in which at some time, if one specifies the variables (the metric say) and its time derivative at that time, this should determine the solution in the future. But a generally covariant field equation would state that if one had such a solution, then one could, in a finite region of spacetime to the future of that intial time, do a change of coordinates. The new metric in the new coordinates would differ from the metric in the old system, but only in that region. In particualr the initial values would be the same. Thus one would have two solutions to the field equations with the same initial data. This would be a disaster since physics would no longer describe the future as a developement of the past. There would be an element of arbitraryness in the future given the past. (Note that this theme also played itself out in Einstein's reaction to Quantum Mechanics 10 years later).

This implied to him that general covariance, that the theory should be independent of the coordinates in which one describe the theory, was wrong. However perhaps one could limit the applicable coordinate changes. After all Special Relativity was described by a coordinate freedom, but limited to special coordinate changes which were linear– the Lorentz transformations. Maybe he would have to be happy with this. He came up with a theory, which has been called the "Entwurf" theory (trial theory) which limited the coordinate transformations. For example the coordinate transformations could only be such that the determinant of the metric coefficients was a constant (-1). This led to a requirement that the trace of the energy momentum tensor had to be 0. That was true of electromagnetism, but not of other matter (eg fluids). But perhaps, as Mie hypothesised, all matter was really electromagnetic in nature.

By the mid 1915's he realised that this theory had more an more problems. The deflection of light was neither the "Newtonian"(as in the variable c^2 theory). The precession of the perihelion of Mercury was nowhere near the experimental value of 43 seconds of arc per century. And the theory seemed to be internally inconsistant.

In November 1915 he battled his way through the problems, making almost daily changes to the theory and publishing weekly changes. He gradually forced himself back to the ideas from 1913, and realised that his uneasiness with that theory could be aleviated. The determinant of the metric being unity was not a necessary condition, but could be imposed as a coordinate condition. The trace of the energy momentum tensor need not be zero. The Hole arguement could be avoided by treating the coordinate changes as changes of description rather than changes of physics (just as the change from cartesian to spherical polar coordinates did not imply any change of physics in space). And finally he realised that the equation

$$R_{\mu\nu} = 8\pi G(T_{\mu\nu} - \frac{1}{2}T_{\alpha\beta}g^{\alpha\beta}g_{\mu\nu}) \tag{26}$$

would work. This is identical to the modern equaiton which is identical to this with 1/2 the trace of both sides being subtracted from both sides

$$R_{\mu\nu} - \frac{1}{2}R_{\alpha\beta} = 4\pi G T_{\mu\nu} \tag{27}$$

At the same time, Schwartzschild was solving the linearized equations, and then the full equation, in a coordinate system $\{t, r^3, cos(\theta), \phi\}$ of the usual coordinates we now use, for which $detg = -1$. These make the equations much simpler to solve in these adapted spherical polar coordinates.

So what did he learn along the way. His c^2 theory taught him that a) describing gravity not a force, but rather as a change in the structure of space and time, with the motion of particles being along straight lines without external forces and b) that it was the time-time component of the metric that was critical to the the behaviour of these geodesics as giving the same equations as in Newtonian gravity. It was clearly insufficient since he believed it must be energy, not mass that was the source for gravity,

From Norstron's theory, he learned that it was possible to create a consistant (even though physically wrong) theory in which gravity was geometry. It was also wrong in that its source could only be the rest mass of the matter, not the energy. It also gave the "wrong" deflection of light (although no experiments existed to to show the right value), and also predicted that electromagntic fields had no gravity, and give the wrong value for the measured perihelion advance of Mercury. (Coming in the midst of his struggles with his own Entwurf theory it was also a signal that a consistant theory was possible.)

From Entwurf he was forced to learn the lesson of what coordinates really meant, and free himself from the hold that of the idea that coordinates, though crucial for doing physics, had any physical significance.