MOTIVATIONS FOR A GENERAL THEORY OF RELATIVITY

SPECIAL AND GENERAL RELATIVITY

V.H. BELVADI*

―――――

IN THE PREVIOUS part of this course we discussed the special theory of relativity. It was an excellent description of the relativistic nature of observations made across a rest frame and a frame in uniform motion—both as seen by an observer inside a hypothetical laboratory[1]. We also reconciled electrodynamics with this theory to get several covariant formulations.

A couple of other observations we made were that, first, the speed of light was a constant across frames—an invariant in other words—and it was the speed limit of the universe; and, second, that the principle of relativity demands that observers across all frames of reference agree as to an event; that is to say, the laws of physics must be the same across frames of reference.

We did notice a couple of drawbacks of our theory, though. Special relativity was concerned solely with rest frames (co-moving with our lab frame) and a 'primed' frame which was in motion with a uniform velocity. The entire theory breaks down if the frames are accelerated. Additionally, special relativity is formulated entirely in flat space and says nothing as to the curvature of space. The world around us is filled with accelerating systems and spaces that are not always flat indicating that the theory we constructed up to this point is likely no more than a special case of a much larger, more general theory of relativity.

1 SOME FOUNDING IDEAS

Besides these drawbacks of special relativity there were certain other observations that led to the formulation of a more accurate, universally valid theory of relativity that we now call the general theory of relativity. We will take a look at a few such ideas in this section.

―――――

For more visit vhbelvadi.com/teaching.

*vh@belvadi.com

[1] This is why we sometimes also call the rest frame the 'lab frame'.

1.1 *Shortcomings of Newtonian gravity*

The person who formed the earliest scientific theory of gravity was Isaac Newton. His work seemed to explain a lot of observations, from objects falling here on earth to the orbits of planets in out solar system. Newton's theory proposed the following formula to describe the gravitational attractive force $\mathbf{F}$ between two masses m and M separated by a distance r:

There was never any talk of repulsion due to gravity. It is a purely attractive force.

$$\mathbf{F} = -\frac{Gm\mathrm{M}}{r^2}\hat{\mathbf{r}} = -\frac{Gm\mathrm{M}\mathbf{r}}{r^3}$$

If this force gives the effect on mass M then we have

$$\mathbf{F} = \mathrm{M}\mathbf{g} = -\frac{Gm\mathrm{M}\mathbf{r}}{r^3}$$

$$\mathbf{g} = -\frac{Gm\mathbf{r}}{r^3}$$

with the x-component alone given by

$$g_x = Gm\frac{r_x}{\left(r_x^2 + r_y^2 + r_z^2\right)^{3/2}} = Gm\frac{\partial}{\partial r_x}\frac{1}{\left(r_x^2 + r_y^2 + r_z^2\right)^{1/2}} = \frac{\partial}{\partial r_x}\left(\frac{Gm}{r}\right)$$

which means, in three dimensions, we have

$$\mathbf{g} = -\boldsymbol{\nabla}\left(-Gm/r\right)$$

This is the gravitational acceleration for any mass m at a distance of $\mathbf{r}$ from that mass. Since any conservative force can be expressed as the gradient of a scalar potential and we have above the gradient of the scalar $-Gm/r$ we realise that this is nothing but a 'gravitational' potential:

$$\mathrm{U} = -\frac{Gm}{r}$$

Now Gauss's law for gravity says $\boldsymbol{\nabla}\cdot\mathbf{g} = -4\pi G\rho$ where ρ is the mass density. It is worth remembering that ρ a function of position $\mathbf{r}$. Substituting $\mathbf{g}$ into this relation gives us

$$\boldsymbol{\nabla}\cdot\mathbf{g} = \nabla^2\mathrm{U} = -4\pi G\rho$$

The nature of U should remind us of the electrostatic potential V which obeys $\nabla^2 \mathrm{V} = -\rho/\epsilon_0$. In other words, like V the gravitational potential U takes values all over space depending on ρ (and therefore mass m) and the constant G.

This gravitational potential structure of Newton's gravity poses a fundamental problem. If we make some changes to the mass m, and therefore to the density ρ, we realise that that change is reflected in U according to the formula above. This, by itself, is not a violation of any law of physics but let us look a little deeper into the effect this has. Specifically, observe that any change in ρ is *instantly* reflected in U not least because the gravitational potential is

completely independent of the distance **r** and the time taken.

Suppose we are 3×10^9 metres away from a mass giving off light we know from our previous discussions that this light would, traveling at c, take $10\,\text{s}$ to reach us. However, going by out relation for U above, any change in this mass would be felt instantly, even before the light from that mass reaches us. In other words causality is being broken and gravitational information is traveling faster than light in a vacuum. This violates special relativity.

1.2 *Other inconsistencies*

There were certain other inconsistencies that physicists noted that would indirectly lead to general relativity. Einstein, for example, realised that standing on a lift on earth, under the influence of **g**, was in no way different from standing in a spaceship, out in space, accelerating at **g**. In other words a locally accelerated frame of reference was in no way different from a frame exposed to a global acceleration like **g**. As we will later see spacetime and motion through it would be curved in such frames and would require a rewriting of a generalised version of the special theory of relativity.

We will not be spending too much time on qualitative descriptions of observers in accelerated frames. However [1] is an excellent, short set of notes and it is strongly recommended that you find time to read it.

Another key observation that had already been made was the orbit of Mercury. Specifically the planet's orbit did not precess to the degree that Newton's theory of gravity predicted. While all other planets did obey the precession measurements expected based on Newtonian gravity Mercury's orbit (which was expected to be $5,600''$ of an arc per one-hundred years) was observed to be $5,557''$ of an arc per one-hundred years: off by $43''$ of an arc per century. This is an incredibly small measure but it is a discrepancy nonetheless. General relativity would later predict a precession that fully agreed with observations.

1.3 *Eötvös's experiment*

Early interpretations of gravity were based on common sense and therefore misled by parameters nobody bothered to account for. A heaver object, by weight, was believed to fall faster than a lighter one. In other words an object with greater mass was thought to be influenced by gravity differently than an object of lesser mass. To state it more succinctly, inertial and gravitational masses were thought of as different quantities. To wit, the former is given by Newton's laws of inertia $\mathbf{F} = m_i\mathbf{a}$ and the latter is given by our standard description of weight as a force $\mathbf{W} = m_g\mathbf{g}$. The assumption early on was the $m_i \neq m_g$.

Although experiments like the one Galileo performed at the Leaning Tower of Pisa where he dropped two equally heavy masses of different substances and found that they fell to the ground almost simultaneously did show empirically that $m_i = m_g$ there was no sufficiently exact proof until the 1890s when the Hungarian physicist Loránd Eötvös measured their equivalence to the ninth decimal using a torsion balance or, what we now call, an Eötvös pendulum.
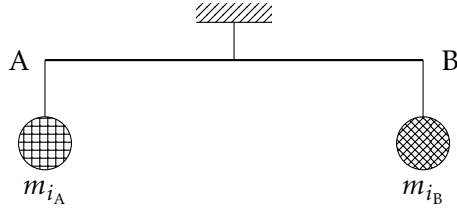
Figure 1: Schematic of Eötvös's pendulum.

As shown in fig. 1 the torsion balance consists of a pair of objects made of different materials (e.g. steel and aluminium, gold and lead etc.) but of equal weight. This way we have eliminated m_g from our equation since equal weights implies that $m_{g_A} = m_{g_B}$.

We now work on the assumption that $m_{i_A} \neq m_{i_B}$ which would be true iff $m_i \neq m_g$. The objects are suspended on the two ends of a rod AB which is itself suspended from, say, a ceiling. Eötvös used a laser, mirror and spectrometer set-up in addition to this balance (not shown in the figure) with the mirror attached to the suspension holding up rod AB and reflecting the laser onto the spectrometer. The entire balance must preferably be isolated and boxed to prevent any motion due to, say, the wind.

There are now two forces acting on this set-up: one, a gravitational force and two, a centripetal force. The former is due to the gravitational pull of the earth and the latter is due to the rotation of the earth. Mathematically, the former is given by $\mathbf{W} = m_g \mathbf{g}$ and the latter by $\mathbf{F} = m_i \mathbf{a}_c$ where $\mathbf{a}_c$ is the centripetal acceleration of the earth. Since we have equalised the former by setting $m_{g_A} = m_{g_B}$ we need only worry about the latter, i.e. we are now concerned only with the effect the centripetal force due to the earth's rotation has on the two different materials.

One of two things can happen at this point. Either the effects are felt differently indicating that $m_{i_A} \neq m_{i_B}$ or the two materials experience the same effect indicating that $m_{i_A} = m_{i_B}$. In the former case if the centripetal force experienced by the objects is different this amounts to a torque that would turn the balance about the suspension of AB. In the latter case nothing would happen; specifically, if the objects experienced the same centripetal force they would not experience any torque and therefore the mass would not spin about its main suspension.

If the set-up spun there would be a change in the angle of the light reflecting off the mirror and this angle could be measured by a spectrometer. Eötvös, however, in performing this experiment, did not notice a considerable rotation and managed to show, subsequently, that the gravitational and inertial masses of the two objects A and B were one and the same.

1.4 *The principle of weak equivalence*

The mathematical results of Eötvös's experiment are summed up in the principal of weak equivalence, a principle that states, essentially, that the inertial and gravitational masses of an object are equivalent to each other. We will now go through a simple mathematical treatment of the same.

We know that the force on mass m_{g_A} due to some mass M is given by

$$\mathbf{F}_A = \frac{m_{g_A} M}{r_A^2}$$

where we have used **geometrised units** where $c = G = 1$ and we will continue to use this throughout general relativity. Likewise the inertial mass too obeys Newton's familiar law

$$\mathbf{F}_A = \mathbf{a}_{c_A} m_{i_A}$$

If we now write similar laws for a second mass using the subscript B and observe under the condition that, as Eötvös's experiment showed us, $\mathbf{a}_{C_A} = \mathbf{a}_{C_B}$ we have

$$\mathbf{a}_{C_A} = \frac{\mathbf{F}_A}{m_{i_A}} = \frac{m_{g_A} M}{m_{i_A} r_A^2} = \frac{m_{g_B} M}{m_{i_B} r_B^2} = \frac{\mathbf{F}_B}{m_{i_B}} = \mathbf{a}_{C_B}$$

When we observe the masses m_A and m_B at the same distance from M, i.e. when $r_A = r_B$ we have

$$\frac{m_{g_A}}{m_{i_A}} = \frac{m_{g_B}}{m_{i_B}}$$

The precise mathematical variation between m_g and m_i is given by the difference between their ratios compared to their mean. To put it differently, we are interested in the difference between the gravitational to inertial mass ratios of two bodies divided by the mean of their gravitation to inertial mass ratios:

$$\eta(A, B) = \frac{2\left\{\left(\frac{m_g}{m_i}\right)_A - \left(\frac{m_g}{m_i}\right)_B\right\}}{\left(\frac{m_g}{m_i}\right)_A + \left(\frac{m_g}{m_i}\right)_B} \qquad \text{(Eötvös parameter)}$$

The value $\eta(A, B)$ is known as the Eötvös parameter and is always given for a pair of materials A and B. Ideally, since $m_g = m_i$ this parameter must be zero. However, due to construction and other practical differences, it is possible to notice an incredibly tiny variation between the two. Whereas Eötvös himself measured this equivalence up to the ninth decimal, i.e. to an accuracy of 2×10^{-9} we have been able to make considerably more accurate measurements in later years. A particularly notable attempt will be [2] which is intended to test the principle to an accuracy of 10^{-18} and is currently undergoing a Phase A study.

The principle of weak equivalence therefore boils down to one of three[2] possible statements essentially equating gravitational and inertial masses:

> Free fall (in a uniform, static field) and uniform acceleration (at **g**) are indistinguishable from each other.

This simply tells us that there is nothing special about **g** itself and that it is the same as manually accelerating at the same rate. Consequently this means

---

[2]This is not strictly true in that there are other forms of stating this principle but we will condense it into three fairly distinct possibilities that share a common underlying theme.

the inertial and gravitational masses are also equivalent.

> The world lines of bodies in free fall are independent of the structure of those bodies.

Recall from our discussions on SR that the world line is a line (or curve) joining various events (points) associated with an object on a spacetime diagram. This form of stating the weak equivalence principle suggests that the world line we draw for a particle is independent on the material that an object is made of, just like in our torsion balance experiment.

> The laws of physics for a body in free fall are the same as in an accelerated reference frame.

What happens if we have light traveling upwards in freely falling lift? This effect is known as a 'gravitational redshift'. Try to read about this in your leisure time.

Arguably the most complete of the three statements this suggests that there is no way to distinguish between a frame resting in a uniform gravitational field and a frame accelerating in the absence of gravity.

1.5 *The principle of strong equivalence*

Einstein took the idea of weak equivalence and extended it to formulate what is known as **the principle of strong equivalence** or **Einstein's equivalence principle**. It is the foundation of the general theory of relativity and will be the last idea we discuss before formally beginning our examination of said theory.

Einstein was famous for his *gedankenexperiments* or *thought experiments*—a practice that was common in physics since before Einstein but were popularised largely by his work—wherein we imagine an experiment rather than perform it often since it is too complex to perform or is quite easily imagined.

Consider an observer inside a closed lift. Say there is a laser fixed to one of the inner walls of the lift that allows us to shine a beam of light that strikes the opposite wall. If the lift is now dropped from great height somewhere on the earth—that is, the lift system is now in free fall inside a uniform gravitational field—the observer experiences precisely what SR tells him: the laser beam leaves on wall and strikes another while traveling parallel to the floor of the lift.

Next say we have an stationary observer outside the lift. The question is what will he observe? Although locally (in the lift) we found that SR was valid the observer outside would see something that contradicts SR. Because observers across reference frames cannot disagree as to an event the only explanation we have is that the second observer will see the laser leave the first wall at some height and since the lift would have accelerated downwards before the beam can strike the opposite wall the beam would travel downwards along an arc proportional to the acceleration of the lift and then strike the other wall.

Since the lift itself does not interact with the motion of the beam of light it is valid to claim that while in the absence of gravity (in the first case when

the lift was in free fall) the light travels in a straight-line in the presence of gravity (in the second case) light must bend.

Einstein's equivalence principle can best be stated as follows:

> Outcomes of local, non-gravitational experiments in a freely falling laboratory are independent of the velocity of the laboratory and its location in spacetime.

While it seems similar to the weak equivalence principle the use of the word 'local' has particular significance in this context. This principle suggests not only that inertial and gravitational masses (or free fall and acceleration) are equivalent but also that all laws of physics are alike in a frame in a uniform gravitational field and in a non-inertial frame outside any such field.

Sometimes Einstein's equivalence principle is distinguished from **the strong equivalence principle** which says

> Outcomes of local experiments in a freely falling laboratory are independent of the velocity of the lab and its location in spacetime.

Clearly this statement is hardly different from that of Einstein's equivalence principle except that it extends the idea to gravitational experiments too. In this course we will not dwell on the difference between these statements and use the terms 'strong equivalence' and 'Einstein's equivalence' interchangeably.

The takeaway here is simple: one, whereas you experience a gravitational 'force' while you are stationary in a uniform gravitational field, you cannot distinguish this from the pseudo-force you would experience while you are accelerating somewhere in space far beyond the influence of such a gravitational field; two, the laws of physics in a 'local' region of spacetime are the same as in any other similarly 'local' region (so long as 'local' means sufficiently small that the gravitational field may be considered uniform) and that gravity itself must be an entirely geometric idea. That is to say, SR applies in a sufficiently small region of spacetime even in the presence of gravity; alternatively, you can transform away the effects of gravity in a sufficiently small region of spacetime.

Although there are other theories of gravity the only one we have so far that obeys the strong equivalence principle is Einstein's theory of General Relativity.

References

[1] Fratus, K. Accelerated observers. Part of a UCSB course. Available for free online at web.physics.ucsb.edu/~fratus/phys103/LN/IGR.pdf.

[2] NASA/Goddard, ESA. Test of the equivalence principle, Phase A: ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19950019764.pdf.