M1 GR + Cosmology - 2 - Historical sketch of GR

Nick Kaiser

September 17, 2020

Contents

List of Figures

1 Newtonian gravity

We will now review the essentials of Newtonian gravity.

1.1 Newtonian reference frames

- Newtonian dynamics in general (and Newtonian gravity in particular) plays out in a space-time arena where there is an *absolute time* t measured by universal clocks and positions $\mathbf{r}$ measured by rulers
- there is, however, no absolute system of spatial coordinates as there is for time

1.1.1 Inertial reference frames

- In the absence of gravity, Newton's laws of motion obey the *Galilean principle of relativity:*
  - these laws are the same in any one of a set of *inertial frames* that are in *uniform linear motion* with respect to one another
  - there is no one inertial frame that is special, and with respect to which one can determine one's absolute state of motion
  - Newton's laws embody this principle – they are obeyed in all inertial frames
  - examples are:
    - a freely moving object moves at constant velocity
    - total momentum is conserved in collisions between objects
- differently *accelerating* frames, however, are distinguishable and the above laws do not generally hold
- neither does the principle extend to relative uniform rotation – this being an example of acceleration
  - in this sense rotation is absolute – the inertial frames are non-rotating
- observers – attached to a rigid frame of rulers – can establish if that frame is inertial (i.e. if they are inertial – or freely falling – observers). If they are not:
  - they will sense strain in their bodies
  - test-particles they release will accelerate

1.1.2 Newtonian frames with gravity

- with the addition of gravity we still have a family of relatively linearly moving frames, but the observers that are fixed in such a frame are no longer inertial
  - observers at fixed position must be accelerated so as not to fall under gravity
  - the frame must somehow be kept rigid to keep the observers in place (assuming they have mass)
  - so they need to be either supported by a rigid framework – perhaps anchored to masses at large distance – or kept in place by rockets

1.2 The inverse-square gravitational force

- according to Newton, the gravitational force is an inverse square attraction
- for a collection of point-like particles of mass m_i and positions $\mathbf{r}_i$
  - where i is an index that labels the particles
- the force on the i^{th} particle is
  - $\boxed{\mathbf{F}_i = Gm_i \sum_{j \neq i} m_j \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|^3}}$
- which is *linear* in the sense that the force is the linear sum of the forces from the other particles

1.2.1 Passive, active and inertial mass

- within the general framework of inverse square laws, we can, in principle, define 3 types of mass
- we could write the force $\mathbf{F}_{21}$ on a particle 2 due to particle 1 as
- $\mathbf{F}_{21} = G m_{\mathrm{a}1} m_{\mathrm{p}2} \frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3}$
  - where subscripts a and p denote the active and passive masses
  - so $m_{\mathrm{a}1}$ determines how much field particle 1 produces
  - and $m_{\mathrm{p}2}$ determines how much force that field induces
- and the acceleration of the 2nd particle as
- $\mathbf{a}_2 = \mathbf{F}_{21}/m_{\mathrm{i}2}$
  - where $m_{\mathrm{i}2}$ is the 2nd particle's *inertial mass*
- in Newton's theory of gravity, the passive and active masses for all particles are proportional to one another (and may be taken to be identical)
  - this guarantees that Newton's 3rd law is obeyed (i.e. $\mathbf{F}_{12} = \mathbf{F}_{21}$)
  - this is like in low-velocity electrodynamics where the field $\mathbf{E}$ that a particle creates and the force on a particle from a given field are both proportional to the charge
  - gravity is more restricted in that all 'charges' are positive, so forces are always attractive
- gravity is quite different to electrodynamics, however, in that, according to Galileo, *all particles accelerate identically in a given gravitational field* independent of their composition
  - so the ratio of passive to inertial mass is the same for all objects,
  - and can be taken to be equal
  - this is called the *Galilean principle of equivalence*
- both gravity and electrodynamics obey the Galilean principle of relativity but only gravity obeys the Galilean principle of equivalence
- another feature of Newtonian gravity is that, while it involves a field $\mathbf{g}$ (or the gravitational potential field ϕ) it is *acausal* (unlike Maxwell's electromagnetism) in that changes in the configuration of particles is communicated *instantaneously* to the other particles
  - we say that there is *instantaneous action at a distance* in Newtonian gravity
  - and the field here is determined by the instantaneous positions of the masses elsewhere
  - Newton did not like this - as he wrote to Bentley: *"That Gravity should be innate, inherent and essential to Matter, so that one body may act upon another at a distance thro' a Vacuum, without the Mediation of any thing else, by and through which their Action and Force may be conveyed from one to another, is to me so great an Absurdity that I believe no Man who has in philosophical Matters a competent Faculty of thinking can ever fall into it. Gravity must be caused by an Agent acting constantly according to certain laws; but whether this Agent be material or immaterial, I have left to the Consideration of my readers."*
  - and it was in regard to this that he said *"hypotheses non fingo"*

1.3 The kinetic energy T

- The energy of a gravitating system has two components: the *kinetic energy* (KE) and the *potential* or *binding* energy (PE)
- of these, the kinetic energy is straightforward:
  - it is the sum over particles of their individual kinetic energies

– it is often denoted by T:

– $\boxed{T = \sum_i m_i |\mathbf{v}_i|^2/2}$

- the KE thus defined depends on the inertial frame
  - however, it is the sum two terms: one giving the energy of the particles relative to the centre of mass frame
  - which *is* frame independent
  - and the other being the KE of the system as a whole

1.4 The gravitational binding energy

- the potential energy is more complex as there are various ways to express it
  - one is as the sum of forces dotted with positions
  - another is as a pairwise sum of shared potential energies (like particles connected by springs)
  - which can be expressed as an integral of the density times a gravitational potential field
  - and which can also be expressed, if one likes, purely as an integral involving the gravity alone

1.4.1 The gravitational binding energy is $U = \sum_i \mathbf{r}_i \cdot \mathbf{F}_i$

- one very useful expression for the gravitational binding energy
  - it is what appears in the virial theorem
- and which we will denote by U, is
  - $\boxed{U = \sum_i \mathbf{r}_i \cdot \mathbf{F}_i}$
- the proof of this is given in the caption of figure 1
- it follows from asking how much energy would be released if we were to assemble a collection of particles into a static configuration by bringing them in from infinity
- the total energy for a non-static configuration is then obtained simply by adding the kinetic energy

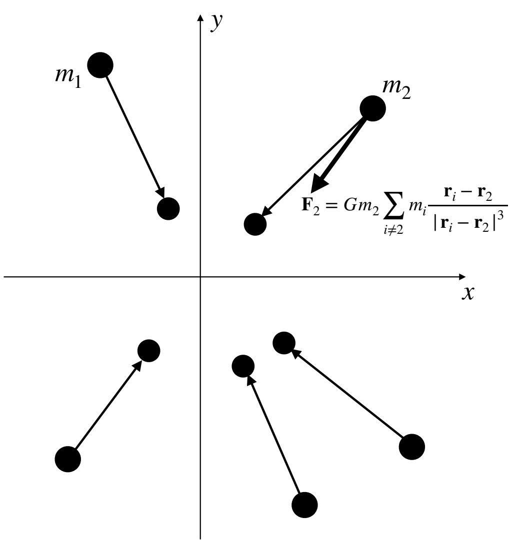

Figure 1: One expression for the gravitational binding energy U of a collection of particles is as the sum over particles of their position $\mathbf{r}$ – relative to the origin of spatial coordinates – dotted with the gravitational force from all the other particles: $U = \sum_i \mathbf{r}_i \cdot \mathbf{F}_i$. To prove this we consider a succession of different configurations where the particles have positions equal to some scale factor $a \geq 1$ times their final positions and we ask how much energy is released as the configuration contracts. This is $dW = \sum_i \mathbf{F}_i(a) \cdot d\mathbf{r}_i(a)$. But $\mathbf{F}_i(a) = \mathbf{F}_i(1)/a^2$ and $\mathbf{r}_i(a) = a\mathbf{r}_i(1)$ and hence $d\mathbf{r}_i = \mathbf{r}_i(1)da$ so $dW = (\sum_i \mathbf{r}_i(1) \cdot \mathbf{F}_i(1)) \times da/a^2$. If we start at $a = \infty$, the energy released is $W = \int_\infty^1 dW = (\sum_i \mathbf{r}_i(1) \cdot \mathbf{F}_i(1)) \times [-1/a]_\infty^1 = -\sum_i \mathbf{r}_i(1) \cdot \mathbf{F}_i(1)$. Conservation of energy requires – as we are considering initial and final configurations with no kinetic energy – that $U = -W = \sum_i \mathbf{r}_i \cdot \mathbf{F}_i$.

1.4.2 The gravitational binding energy is $U = -\frac{1}{2}\sum_i \sum_{j \neq i} G m_i m_j / |\mathbf{r}_i - \mathbf{r}_j|$

- A more common expression of the binding energy is as a pairwise sum of the potential for a pair of particles $-Gm_1m_2/|\mathbf{r}_1 - \mathbf{r}_2|$
  - but with a factor 2 to account for the fact that the energy is 'shared' between the particles in the pair
- we can justify this starting from $U = \sum_i \mathbf{r}_i \cdot \mathbf{F}_i$ as follows:
  - switching labels on the particles $i \leftrightarrow j$ has no effect, so
  - $U = \sum\limits_i \sum\limits_{j \neq i} G m_i m_j \mathbf{r}_i \cdot (\mathbf{r}_j - \mathbf{r}_i)/|\mathbf{r}_j - \mathbf{r}_i|^3 = \sum\limits_j \sum\limits_{i \neq j} G m_i m_j \mathbf{r}_j \cdot (\mathbf{r}_i - \mathbf{r}_j)/|\mathbf{r}_j - \mathbf{r}_i|^3$
  - but $\sum\limits_j \sum\limits_{i \neq j} = \sum\limits_i \sum\limits_{j \neq i}$, since both are just summing over the non-diagonal squares on the 'chess-board', and using $(\mathbf{r}_i - \mathbf{r}_j) = -(\mathbf{r}_j - \mathbf{r}_i)$ the latter expression is
  - $U = -\sum\limits_i \sum\limits_{j \neq i} G m_i m_j \mathbf{r}_j \cdot (\mathbf{r}_j - \mathbf{r}_i)/|\mathbf{r}_j - \mathbf{r}_i|^3$
  - and averaging this with the original expression gives
  - $U = \frac{1}{2} \sum\limits_i \sum\limits_{j \neq i} G m_i m_j \frac{(\mathbf{r}_i - \mathbf{r}_j) \cdot (\mathbf{r}_j - \mathbf{r}_i)}{|\mathbf{r}_j - \mathbf{r}_i|^3}$ or
- $\boxed{U = -\frac{1}{2} \sum\limits_i \sum\limits_{j \neq i} \frac{G m_i m_j}{|\mathbf{r}_j - \mathbf{r}_i|}}$
  - This is the same as the potential energy for a set of particles connected by springs with potential energy – as a function of length r – $G m_i m_j / r$
  - the factor 1/2 coming in as the energy in each spring is 'shared' between the two particles that it connects
  - or, equivalently, because the double sum $\sum_i \sum_{j \neq i}$ counts each pair of particles, and therefore each spring, twice
- this form for the binding energy makes transparent the fact that the gravitational binding energy is always negative

1.4.3 The gravitational potential and acceleration fields $\phi(\mathbf{r})$ and $\mathbf{g}(\mathbf{r})$

- we can also express the binding energy as
  - $\boxed{U = \frac{1}{2} \sum\limits_i m_i \phi(\mathbf{r}_i)}$
- where the *gravitational potential* is defined to be
  - $\boxed{\phi(\mathbf{r}) = -\sum\limits_j G m_j / |\mathbf{r}_j - \mathbf{r}|}$
- and whose gradient is minus the *gravitational acceleration* (or just the *gravity*) vector
  - $\boxed{\mathbf{g}(\mathbf{r}) = -\boldsymbol{\nabla} \phi(\mathbf{r})}$

1.4.4 Potential and gravity for a continuous density field $\rho(\mathbf{r})$

- We can make the transition to a continuous mass distribution by writing the *mass density* for a set of particles as $\rho(\mathbf{r}) = \sum_i m_i \delta^{(3)}(\mathbf{r} - \mathbf{r}_i)$
  - with $\delta^{(3)}(\mathbf{r})$ the 3-dimensional *Dirac* δ-function
    - which can be thought of as e.g. the the limit of a small normalised Gaussian:

* $\delta^{(3)}(\mathbf{r}) = \lim_{\sigma\to 0} (2\pi\sigma^2)^{-3/2} \exp\left(-|\mathbf{r}|^2/2\sigma^2\right)$
  * though there are many other possibilities, of which perhaps the simplest is the limit of a small normalised 'box-car'
    * whose 1D version is: $\delta(r) = \begin{cases} \epsilon^{-1} & \text{if } |r| < \epsilon/2 \\ 0 & \text{otherwise} \end{cases}$
    * and whose 3D version is just the product $\delta^{(3)}(\mathbf{r}) = \delta(x)\delta(y)\delta(z)$
  - and which has the property that $\int d^3r' f(\mathbf{r}')\delta^{(3)}(\mathbf{r}' - \mathbf{r})f(\mathbf{r}') = f(\mathbf{r})$

- and then considering a continuous mass distribution to be the same as a very finely distributed set of point masses in the limit that $m \to 0$

- so we have
  - $\boxed{U = \frac{1}{2} \int d^3r \rho(\mathbf{r})\phi(\mathbf{r})}$
  - with
  - $\boxed{\phi(\mathbf{r}) = -\int d^3r' G\rho(\mathbf{r}')/|\mathbf{r}' - \mathbf{r}|}$

- and $\mathbf{g}(\mathbf{r}) = -\boldsymbol{\nabla}\phi(\mathbf{r})$ as before.

1.5 Poisson's equation and Gauss's law

- the expression above for $\phi(\mathbf{r})$ is a solution of *Poisson's equation:*
  - $\boxed{\nabla^2\phi = 4\pi G\rho}$
- with *boundary conditions* $\phi \to 0$ as $r \to \infty$
- the validity of Poisson's equation can be established as follows
  - first, consider a single point mass m at $\mathbf{r} = 0$
  - the potential is $\phi = -Gm/|\mathbf{r}|$
  - and the gravitational acceleration is $\mathbf{g} = -Gm\mathbf{r}/|\mathbf{r}|^3$
  - from which it follows that $\boldsymbol{\nabla} \cdot \mathbf{g} = 0$ for $\mathbf{r} \neq 0$ (see figure 2)

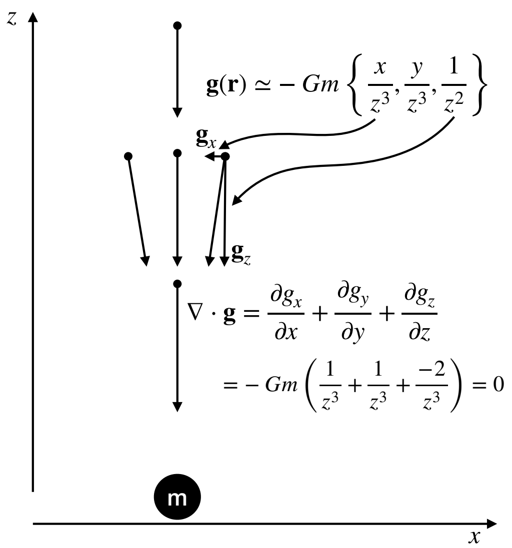

Figure 2: Proof that $\boldsymbol{\nabla} \cdot \mathbf{g} = 0$ outside a point mass. Consider a point $\mathbf{r}_0 = \{0, 0, z\}$ along the z-axis above a point mass m. In the vicinity of that point, the gravity is, to first order in the transverse displacements x and y given by $\mathbf{g}(\mathbf{r}) \simeq -GM\{x/z^3, y/z^3, 1/z^2\}$, whose divergence vanishes. A vector field $\mathbf{f} \propto \hat{\mathbf{r}}/|\mathbf{r}|^2$ like $\mathbf{g}$ is the simplest, and perhaps archetypical, example of a *divergence free flow:* If we think of a fluid with flux density (product of density $\rho(\mathbf{r})$ and velocity $\mathbf{v}(\mathbf{r})$): $\mathbf{f}(\mathbf{r})$ then if we consider concentric spheres, the rate at which fluid is crossing the surface is proportional to the area ($\propto |\mathbf{r}|^2$) times the flux density $\mathbf{f}$ ($\propto 1/|\mathbf{r}|^2$) and is the same for all spheres. Provided we keep supplying fluid at $\mathbf{r} = 0$ (perhaps think of a garden sprinkler, in which case $\rho \propto 1/r^2$ and $\mathbf{v} =$ constant) then there will be no build-up or decrease of density of fluid. This is expressed in the *continuity equation* $\partial\rho/\partial t = -\boldsymbol{\nabla} \cdot \mathbf{f}$.

- but owing to linearity of the force law that means that anywhere that $\rho = 0$ the divergence of the gravity must vanish: the matter elsewhere having no effect

- and hence $\nabla^2\phi = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\phi = -\boldsymbol{\nabla}\cdot\mathbf{g} = 0$ also
- so $\nabla^2\phi$ must be proportional to the density, and the constant of proportionality ($4\pi G$) is readily established considering a small uniform density sphere
- applying the *divergence theorem* gives *Gauss's law:* (see figure 3)
  - $$\boxed{\int d\mathbf{A}\cdot\mathbf{g} = -4\pi G\int d^3r\rho}$$
  - so the integral of the outward component of the gravity over the boundary of a volume is $4\pi G$ times the mass enclosed.
- solving Poisson's equation for some given mass density field gives the potential, and hence the gravity
  - though this is arguably of somewhat limited utility given that we can simply write down the potential – and be confident that the result has the proper boundary conditions at infinity
  - it is very useful for establishing useful relations between other quantities as we shall now illustrate

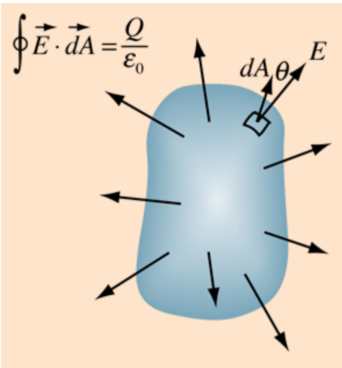

Figure 3: This illustrates (the integral form of) Gauss's law in electromagnetism. It is equivalent the the Maxwell equation $\boldsymbol{\nabla}\cdot\mathbf{E} = \rho/\epsilon_0$ by virtue of the divergence theorem. The gravitational version follows from Poisson's equation in precisely the same manner.

1.6 The potential energy in terms of the gravity

- just as in electro-statics, where one can consider the energy to be the sum over charges of the potential
  - i.e. considering the energy to be associated, and localised, with the charges
- or as an integral $\frac{\epsilon_0}{2}\int d^3r|\mathbf{E}|^2$
  - i.e. considering the energy to be associated and localised with the field
- we can express the binding energy entirely in terms of the gravity
  - if we perform the integral $I = \int d^3r|\boldsymbol{\nabla}\phi|^2$ by parts we get, with sensible boundary conditions, $I = -\int d^3r\phi\nabla^2\phi$
  - which with Poisson's equation is $I = -4\pi G\int d^3r\rho\phi = -8\pi GU$
- $$\boxed{U = -\frac{1}{8\pi G}\int d^3r|\mathbf{g}|^2}$$
- so the energy density associated with the gravitational field is $\epsilon = -|\mathbf{g}|^2/8\pi G$
- but, as with electrostatics, we cannot 'double-count':
  - either we consider the energy to be associated with the masses
  - or with the field
  - but not both
- we will see that in GR there is a similar expression for the energy density of gravitational waves

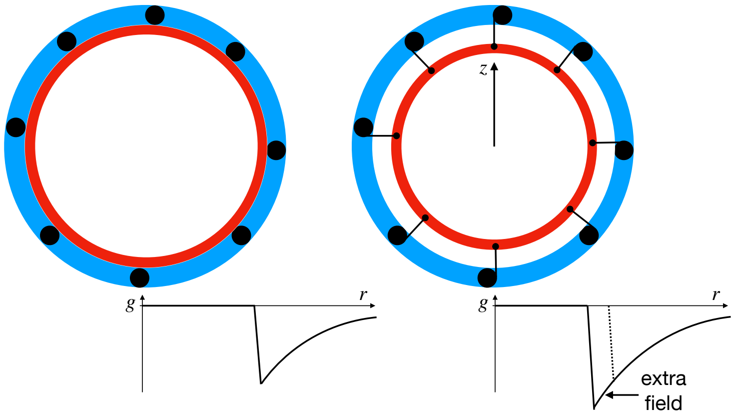

Figure 4: Inside a light rigid shell (blue) is suspended (against its own self-gravity) a massive shell (red). This is lowered by winches which gain energy. This creates new **g**-field in the region that was previously inside the shell (and therefore field-free). The energy gained is $|\mathbf{g}|^2/8\pi G$ times the volume it occupies: hence the energy density of the newly created field is $\epsilon = -|\mathbf{g}|^2/8\pi G$. The cable at the top is carrying +ve z-momentum downwards: a negative flux of z-momentum. For momentum to be conserved, there must be a positive flux of z-momentum in the **g**-field.

1.7 The Newtonian gravitational stress tensor

This section is not absolutely essential for what follows.

- In electromagnetism, we have the Maxwell stress tensor
  - $\mathbf{T} = -\epsilon_0(\mathbf{EE} - \mathbf{I}|\mathbf{E}|^2) - \mu_0^{-1}(\mathbf{BB} - \mathbf{I}|\mathbf{B}|^2)$
  - where $\mathbf{I}$ is the identity matrix
- or, in component form,
  - $T_{ij} = -\epsilon_0(E_i E_j - \frac{1}{2}E^2\delta_{ij}) - \mu_0^{-1}(B_i B_j - \frac{1}{2}B^2\delta_{ij})$
- and which is the *momentum flux density* of the field
  - momentum being a vector quantity, its flux density is necessarily a tensor
  - this describes the transport of momentum by EM fields
  - such as the flux of momentum in the field between two capacitor plates that are being prevented from coming together by springs: it provides the continuity of momentum needed as the plates are neither gaining nor losing momentum
- in Newtonian gravity there is a precisely analogous flux of momentum that one can associate with the *gravitational* field. It is given by
  - $T_{ij} = (8\pi G)^{-1}(g_i g_j - \frac{1}{2}g^2\delta_{ij})$
  - so just like the stress of an electric or magnetic field, but with opposite sign
  - so there is pressure in a direction parallel to $\mathbf{g}$ and tension in the transverse directions
- Maxwell was the first to show that, at the surface of the earth this is positive
  - so if we sit at the N-pole (i.e. +ve z) there is a flux of z-momentum upwards
  - the caption to figure 4 describes how this flux balances the negative momentum flux density in the cables supporting the mass shell
- and it's value is 32,000 tons per square inch!
  - this deterred Maxwell from developing a Lorentz invariant gauge-field theory for gravity
  - he didn't believe that the underlying mechanism of space could be strong enough to withstand such a stress
- Some problems to consider:
  - Q: Show that the 3 components of the divergence of the flux density: $\partial T_{ij}/\partial x_j$ (outside the Sun for instance) vanishes in empty space

    * but that it is non-zero inside the planets in the solar system
    * and that it properly accounts for the changing mechanical momentum of those objects
  - Q: The strength of gravity falls off as $1/r^2$ so the components of T_{ij} fall off as $1/r^4$
    * thus it would seem that, in empty space, the flux of momentum across spheres around a point mass (area times flux density) is decreasing
    * is this a problem? Does this violate conservation of momentum? Where is all of that momentum going?

1.8 The tidal field tensor: $\mathbf{\nabla g} = -\mathbf{\nabla\nabla}\phi$

- the tidal field is (minus) the gradient of the gravity $\mathbf{g}$ or equivalently the second derivative of the potential $\mathbf{\nabla\nabla}\phi$
- or in component notation
  - $\boxed{-\mathbf{\nabla g} \longrightarrow \dfrac{\partial^2\phi}{\partial r_i \partial r_j} \equiv \phi_{,ij}}$
  - where $\mathbf{r} = \{r_i\} = \{x, y, z\}$
- its key features are
  1. the tidal field tensor $\phi_{,ij}$ is *symmetric* so has 6 independent components
  2. its Laplacian $\nabla^2\phi$ (i.e. the *trace* of $\phi_{,ij}$) is what appears in Poisson's equation
  3. the *relative tidal acceleration*: for pair of neighbouring freely falling (i.e. inertial) particles with separation $\mathbf{r}$
     - $\boxed{\ddot{\mathbf{r}} = -\mathbf{r} \cdot \mathbf{\nabla\nabla}\phi}$
     - the fact this this is *linear* in the separation $\mathbf{r}$ is a key characteristic of a tidal field
     - it means that if we have a cold 'dust' of test particles
       * where by 'cold' we mean that the particles are initially at rest with respect to each other
       * or perhaps more generally have some velocity field (relative to one of the particles – the 'fiducial' particle) $\dot{\mathbf{r}} = \mathbf{H} \cdot \mathbf{r}$, with some symmetric expansion rate tensor $\mathbf{H}$
     - the cloud of particles will continue to have the linear expansion law but the expansion rate tensor will evolve with time
     - the same thing happens in GR, where the equation is called the *equation of geodesic deviation*

1.9 Gravity vs. electrostatics

- superficially gravity and electricity are very similar
  - both are inverse square forces (though sign is different)
  - both obey Poisson's equation and Gauss's law
- but there is a key difference: Galileo: all objects accelerate *identically* in a gravitational field
  - like (attractive) electrostatics if all particles had the same *charge to mass ratio*
- so it is *impossible* to measure the gravity $\mathbf{g}$ from e.g. motions of particles
  - quite different from electrostatics where we can compare the acceleration of a neutral and a charged particle
  - in gravity physics there are no 'neutral' particles
  - sometimes called the *Galilean equivalence principle* (GEP)
- an illustration: uniform spherical ball of dust
  - $\mathbf{g} = -GM\mathbf{r}/r^3 = -(4/3)\pi G\rho\mathbf{r}$

        * acceleration $\propto \mathbf{r}$ so the ball starts to collapse (or slows down expansion) maintaining its uniform density
        * but freely falling observers – e.g. dust particle – cannot locally determine $\mathbf{g}$
    - the tidal field is $\phi_{,ij} = (4/3)G\rho\delta_{ij}$ and is spatially constant and *isotropic*
        * if an observers releases a set of test particles at rest in his environment they just accelerate towards him – no tidal *deformation* – the tidal force is isotropic
        * so all points within the sphere are equivalent
        * and there is no way for an observer to determine the direction to the centre
    - this is very different from an electrically charged sphere where
        * $\mathbf{E}$ *is* locally measurable
        * points towards the centre (breaks symmetry)
        * and diverges – with bad consequences – if we let radius $R \to \infty$
    - for a gravitating sphere we can let $R \to \infty$ with no bad effects
        * this is why we can "do cosmology" with gravity
        * but one cannot have a charged infinite universe

1.10 Measuring the gravitational field

- the gravity $\mathbf{g}$ *can* be measured (e.g. with a weighing scale)
    - but only with inclusion of non-gravitational forces (e.g. the surface of the earth)
- the *tide* in contrast is directly measurable from the relative motions of masses
    - and is the *only* effect of distant masses
    - if all we can do is observe motions of freely falling particles then the gravitational potential cannot be determined unambiguously
        * we can always add a global constant to ϕ – no surprise
        * and we can add a constant *gradient* $\phi \to \phi' = \phi - \mathbf{g} \cdot \mathbf{r}$
- the matter distribution determines the tide ... and the tide tells the matter how to move
    - or at least determines the *relative* motion of matter particles
- it is the tide that is *the* gravitational field

2 1905: Special Relativity (SR)

2.1 Principles and main implications

- Einstein: two principles:
  - the laws of physics are identical in relatively (linearly) moving reference frames
    - like Galilean relativity (though unlike the ancient Greeks' concept)
    - note only *linear* motion is relative – rotation is still absolute in SR
  - the speed of light is c in all such reference frames
- implications:
  - time dilation – length contraction – no simultaneity

2.2 Space-time; events; intervals; Lorentz transformation and the light-cone structure

- space and time are merged into the *space-time continuum*
  - arguably simpler than Galilean relativity
- a *reference frame:*
  - a family of observers on a rigid, non-rotating, non-accelerating (no gravity in SR!) lattice made of rulers to measure positions
  - and carrying synchronised clocks to measure 'proper' time
  - who record the 4D space-time coordinates $\{x^\alpha\} = \{x^0, x^1, x^2, x^3\} = \{ct, \mathbf{r}\}$ of 'events'
  - such a vector is called a 'contra-variant' vector (see below)
- differently 'observers' record different *components* for an interval $\{c\Delta t, \Delta\mathbf{r}\}$ between two events
  - related by *Lorentz transformation* $x^{\alpha'} = \Lambda^{\alpha'}{}_\alpha x^\alpha$ where
    - the transformation matrix (for a boost along the x-axis) is
    - $$\Lambda^{\alpha'}{}_\alpha = \begin{bmatrix} \gamma & \gamma v/c & & \\ \gamma v/c & \gamma & & \\ & & 1 & \\ & & & 1 \end{bmatrix}$$
    - and where the *Lorentz boost factor* is $\gamma \equiv 1/\sqrt{1 - v^2/v^2}$
  - transformations of components of intervals between frames are
    - a bit like (passive) *rotations* of 3-vectors in Euclidean space in that
    - the vector has a real objective existence as a *geometric object*
    - but its *description* in terms of coordinates is frame/observer dependent
  - we write $\vec{x} \underset{\mathrm{O}}{\longrightarrow} x^\alpha$ to indicate the relationship between the frame-independent geometric object $\vec{x}$ and its (contravariant) components in a specific reference frame O
- but space-time of SR is *non-Euclidean*
  - the (frame invariant) 'length' of a vector is $\Delta s^2 \equiv -c^2\Delta t^2 + \Delta r^2$
    - or $\Delta s^2 = \eta_{\alpha\beta}\Delta x^\alpha \Delta x^\beta$
    - and where the *Minkowski metric* is $\eta_{\alpha\beta} = \mathrm{diag}\{-1, 1, 1, 1\}$
    - also used to lower indices to make 'covariant' vectors $V_\alpha = \eta_{\alpha\beta}V^\beta$
  - invariant interval Δs^2 can be negative (time-like), positive (space-like) or zero (null)
  - the associated 'light-cone' structure is *absolute* – i.e. observer-frame independent – it is an intrinsic property of Lorentzian space-time

2.3 Scalars, vectors, tensors and the laws of physics

2.3.1 Lorentz scalars, 4-vectors and 4-tensors

- *Lorentz scalars* are things that are the same in all Lorentz frames
  - an example might be the temperature of a gas $T(\vec{x})$ (as measured by observers moving with the gas)
  - many other Lorentz scalars arise as the contraction (dot-product) of 4-vectors
    * such as the invariant interval
    * $\Delta s^2 = \vec{\Delta x} \cdot \vec{\Delta x} = \Delta x_\alpha \Delta x^\alpha$
- 4-vectors are things that transform like 4-displacements
  - examples are 4-velocity $\vec{u} \equiv d\vec{x}/d\tau \longrightarrow u^\alpha = (\gamma c, \gamma \mathbf{v})$ (rate of change of position wrt proper time)
  - 4-momentum $\vec{p} \equiv m\vec{u} \longrightarrow p^\alpha = (\gamma mc, \gamma m\mathbf{v}) = (E/c, \mathbf{p})$
  - and electromagnetic potential $\vec{A} \longrightarrow A^\mu = (\phi/c, \mathbf{A})$
  - the derivative operator $\vec{\partial} \longrightarrow \partial_\mu = (c^{-1}\partial_t, \mathbf{\nabla})$
- and we also have Lorentz 4-tensors
  - like, for example, the Faraday tensor
  - $\mathbf{F} \longrightarrow F_{\mu\nu} \equiv A_{\mu,\nu} - A_{\nu,\mu}$
  - whose components transform like $x_\mu x_\nu$
- the laws of physics in flat space-time (no gravity) are expressed in terms of 4-vectors etc.
- there are two types of laws: 'equations of motion' and 'continuity' or 'conservation laws'

2.3.2 Equations of motion

- an example of a force law comes from electrodynamics:
- $dp^\alpha/d\tau = qu^\beta F_\beta{}^{,\alpha}$
  - which gives rate of change of energy ($E = cp^0$) and relativistic 3-momentum $\mathbf{p}$ of a particle of charge q (a Lorentz scalar) in an EM field
- another is the (inhomogeneous) Maxwell's equations
- $F^{\mu\nu}{}_{,\mu} = j^\nu/\mu_0$

2.3.3 Conservation of continuity laws

– Conservation of particles

- let n be the number density of particles measured by some observer
- and $n\overline{\mathbf{v}}$ be the *3-current* of particles
- then conservation of particles implies $\partial n/\partial t + \mathbf{\nabla} \cdot (n\overline{\mathbf{v}}) = 0$
- which is expressed covariantly as
  - $\boxed{n^\alpha{}_{,\alpha} = 0}$
  - where $n^\alpha = (nc, n\overline{\mathbf{v}})$.
  - and we say $\vec{n} \longrightarrow n^\alpha$ is a 'conserved 4-current' of particles
  - or that its *4-divergence* vanishes
- there are many other useful continuity laws involving things like the electric charge or entropy

– **Continuity of 4-momentum**

- perhaps the most fundamental conservation law is that of energy and momentum
  - this follows from the homogeneity – translational invariance – of space-time (and the physical laws) in SR
  - and is expressed in terms of the *stress energy 4-tensor* $\mathbf{T} \longrightarrow \mathbf{T}^{\mu\nu}$
  - examples are:
    * for a collection of particles with *phase-space density* $f(\mathbf{p}, \mathbf{r})$,
    * $\quad \mathbf{T} = \int \frac{d^3p}{p_0} f \vec{p}\vec{p}$
    * for a perfect fluid (in its rest frame)
    * $\quad \mathbf{T} \longrightarrow \mathrm{diag}(\rho, P, P, P)$
    * we also saw the stress-energy tensor for EM fields and scalar fields
- translational invariance implies 4 continuity equations
  - $\boxed{T^{\alpha\beta}{}_{,\beta} = 0}$
  - one for each of $\alpha = 0, 1, 2, 3$
- For $\alpha = 0$ this expresses continuity of *energy*
  - ∂(energy density)$/\partial$t $+ \boldsymbol{\nabla} \cdot$ (energy flux density) $= 0$
  - which is essentially the *first law of thermodynamics*
- while for $\alpha = 1, 2, 3$ the expresses the continuity of the 3-components of spatial momentum
  - ∂(momentum density)$/\partial$t $+ \boldsymbol{\nabla} \cdot$ (momentum flux density) $= 0$.
  - where *momentum flux density* is – by *definition* – the *pressure*
  - so this is essentially Newton's law $\mathbf{F} = m\mathbf{a}$
  - and which is familiar in *Jeans's equation* in stellar dynamics
- one can do the same thing for fields
  - if fields and/or particles are coupled then it is the total stress-energy tensor that is conserved
- The '00' component of $T^{\alpha\beta}$ is (c^2 times) the mass density that is the source in Poisson's equation
  - it is $\mathbf{T}$ that replaces ρ in the analogue of Poisson's equation in Einstein's gravity

3 1907-1910: Einstein's "happiest thought"

3.1 What was he thinking about?

1. Galileo's observation of the constancy of the gravitational 'charge-to-mass ratio"
   - the fact that all things fall the same way in a gravitational field: the GEP is telling us something *fundamental* about gravity
2. *Mach's principle:*
   - a non-rotating – i.e. inertial – observer sees the distant stars at rest on the sky
   - so *inertia* – a property of space-time – seems somehow determined by the state of motion of distant matter
3. *the gravitational redshift* (ca. 1910)

3.2 The gravitational redshift

There were two separate ideas:

3.2.1 The tower thought experiment

- consider a tower, a mass, and two 'machines'
  - drop a mass from the top
  - the machine at the bottom catches the falling mass, converts its energy $E = m(c^2 + v^2/2 + \ldots)$ to a photon of same energy $E = h\nu$ which it sends it to the top of the tower where
  - another machine converts the photon energy back to mass which it drops to the machine at the bottom
- without a redshift $\delta\nu/\nu = \phi/c^2$ this would allow 'perpetual motion'
  - i.e. continuous extraction of useful energy from a static gravitational field
- Einstein considered this unreasonable – so there has to be a redshift
- this prediction was confirmed by Pound and Rebka

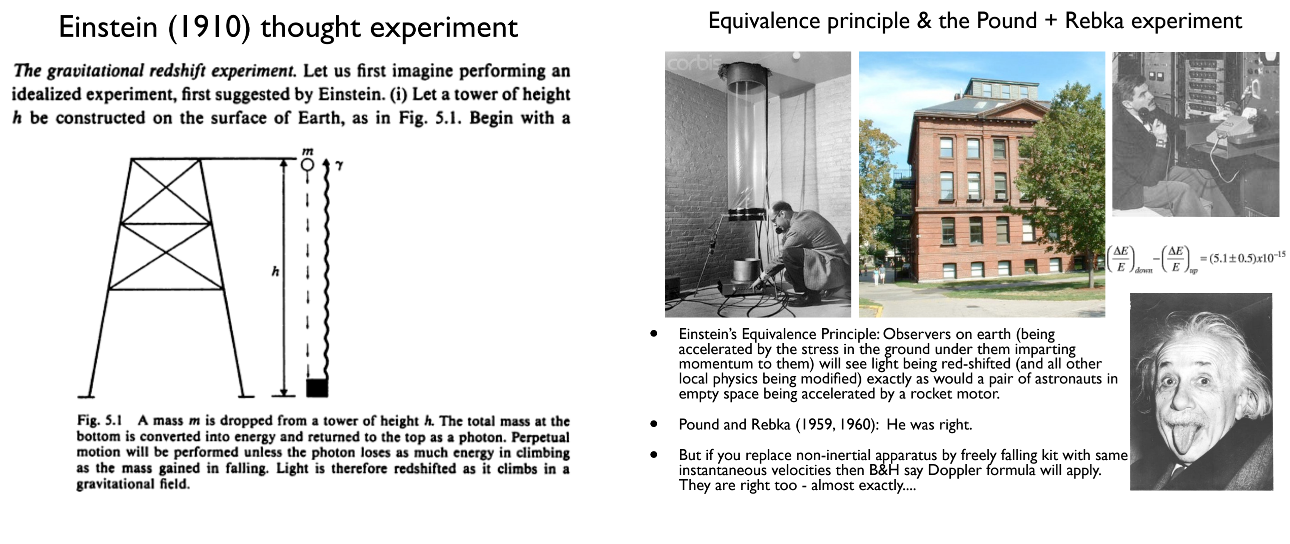

Figure 5: Einstein's tower thought experiment.

3.2.2 The 'rocket' (or elevator) thought experiment

- consider a rocket in empty space
  - let it have vertical acceleration a and length L contain 2 observers A and B
  - A (at the tail) sends a photon of frequency ν to B, at the nose
  - when the photon set off, A and B had the same velocity
  - it takes the photon a time $\Delta t = L/c$ to make the trip
    - we will assume that $a \times (L/c) \ll c$
  - in which time B's velocity has changed by $\Delta v = a\Delta t$
  - so he (or she) sees a 1st order Doppler redshift $\delta\nu/\nu = aL/c^2$
- this is the same as deduced from the tower experiment if $\phi = aL$
  - or, equivalently, if $a = \phi/L$
- but ϕ/L is just the potential gradient – or the gravity $\mathbf{g}$
- so as far as the redshift is concerned gravity and acceleration have equivalent effects
  - the rocket argument is essentially *prerelativisitic* – it has nothing really to do with SR
  - though the tower experiment invokes the equivalence of mass and energy $E = mc^2$

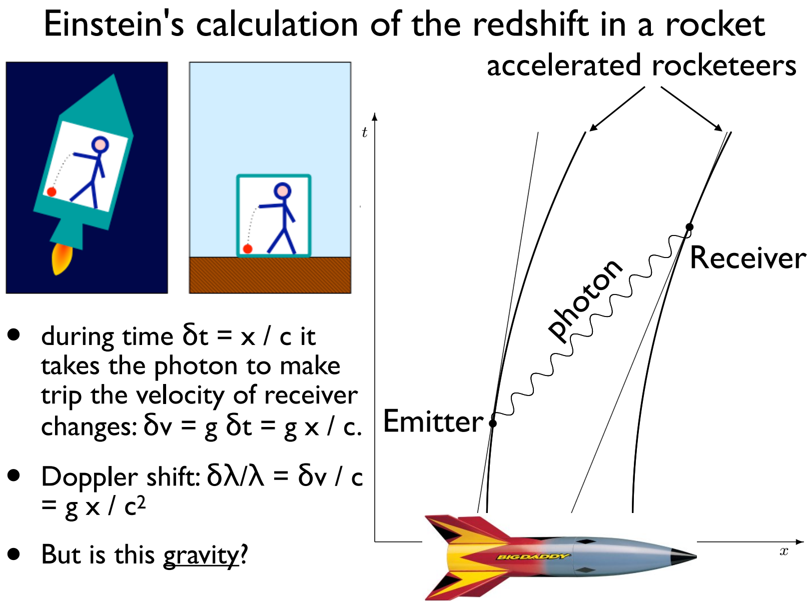

Figure 6: Einstein's rocket thought experiment.

3.3 Einstein's principle of equivalence

- Einstein realised that other phenomena that we associate with terrestrial gravity
  - things we release fall to the ground etc
  - there is stress in our ankles
  - a jar containing gas will have a pressure gradient
- also behave identically to physics in a rocket with $\mathbf{a} = \mathbf{g}$
- which leads to another expression of Einstein's principle of equivalence:
  - *acceleration and gravity are equivalent*

    * this is given many popular books and web-sites on physics
    * but is really rather dangerous and misleading

- what's wrong with this statement of the equivalence principle?
  - for one thing, if we take to heart the idea that the gravitational field is the *tidal field* then one realises quite quickly that things considered here are *not* gravitational effects
  - in a rocket in space there is no tidal field (any tidal field from the rocket structure itself can be made negligible without affecting the redshift)
    * firing the rocket motor does not great any tidal field
    * nor, in GR, as we shall see, does if create any space-time curvature – the relativistic equivalent of the tide
  - and the gravitational redshift in the tower experiment has no sensible relation with the tidal field (to measure the tidal field with something like the Pound and Rebka experiment you'd have to do it at different heights and measure the change in the effect – that would be a much more difficult experiment)
  - another – rather unrelated – problem is that an application of this to light bending will give the wrong answer (see below).

- a better way to state equivalence here is to say that if you are being accelerated by the surface of the Earth under your feet (rather than being freely falling and inertial) then you will see many phenomena (like falling objects, redshifting of light) that are identical to what you would see if you were being accelerated in a rocket in the absence of a gravitational field

- the key thing is that i.e. the presence of the gravitational field around the Earth doesn't add anything *extra* to the effect from the acceleration

- it would be possible to imagine some kind of coupling of gravity to the device in the tower experiment so that in a gravitational field the energy of the photon emitted would not be equal to the energy of the mass received. Einstein assumed that there is no such coupling and Pound and Rebka proved him right.

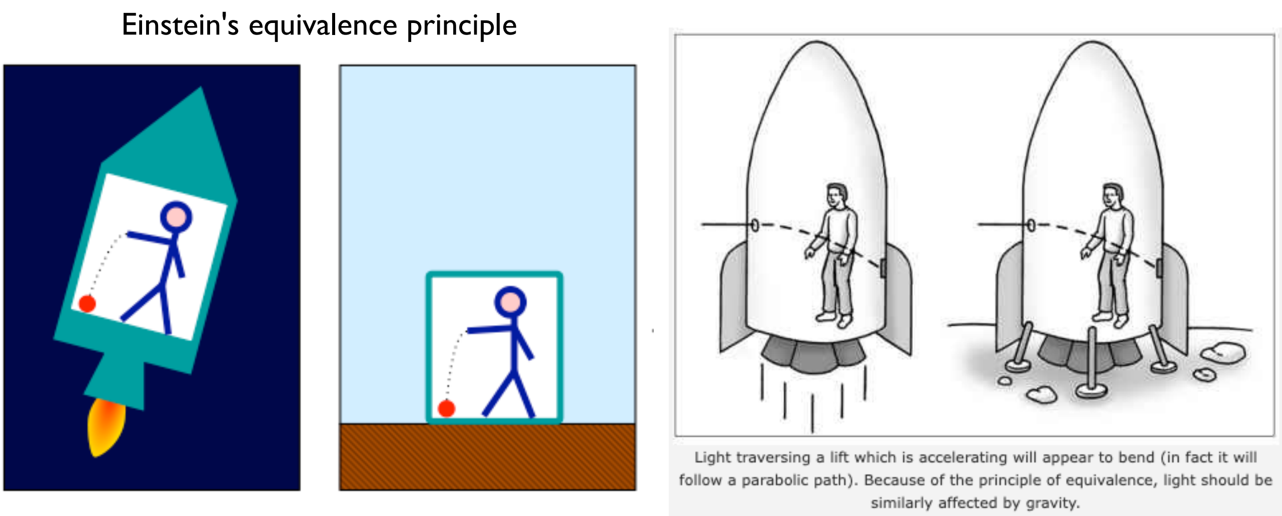

Figure 7: Einstein's principle of equivalence.

3.4 Gravitational light-bending

- Einstein extended this to argue that light must be deflected by a gravitational field since an accelerated observer in a rocket will see a light-ray coming into the rocket from outside, which will travel along a straight line as perceived by a non-accelerating observer outside, to have a curved trajectory

- and this same light bending can be understood as arising from gravitational time dilation
- it is interesting to compare this light bending with refraction by e.g. glass
  - for glass, the frequency ν is constant, but the speed of light varies, so the wavelength changes and the rays get refracted
  - for gravity, the *frequency* changes – because of the gravitational redshift – while the speed of light is everywhere the same
  - but the consequence is the same: the wavelength changes and so the direction of light – which is normal to the wave-fronts – is changed
- one can also extend this to non-relativistic particles (de Broglie waves)
- this is, unfortunately, over-simplified, and under-predicts the true effect (for light – but not for de Broglie waves) by a factor 2 for light-rays

3.5 Gravitational time dilation

- a much more profound lesson from the tower experiment, however, is this:
  - *an observer at the top of the tower will see the watch of an observer at the bottom run slow*
- there must be *gravitational time dilation* (see figure 8)
  - clocks – subject to a continuous acceleration – will drift progressively out of synch
  - so time – and hence space-time – must be warped or 'curved' near the Earth
- it is easy to see, from the gravitational redshift formula, that, relative to a clock at a great height – where the effects of gravity are negligible – a clock in a gravitational potential field $\phi(\mathbf{r})$ will run slow by a factor that is, to lowest order in ϕ given by $1 + \phi/c^2$
  - we can invoke the definition of the redshift: $1 + z = dt_{\rm obs}/dt_{\rm em}$
- so in an interval of *'coordinate time' dt* – being the proper time for the reference clock at large height – the proper time elapsed as measured by the observer where $\phi \neq 0$ is
  - $\boxed{d\tau = (1 + \phi/c^2)dt}$
  - with possibly higher order corrections, but since $\phi/c^2 \sim 10^{-8}$ on Earth these are negligible
- a feature that can be incorporated in the expression for the invariant line element $ds^2 = -c^2 d\tau^2 = dx_\alpha dx^\beta = \eta_{\alpha\beta} dx^\alpha dx^\beta$ by modifying the metric:
  - $\eta_{\alpha\beta} \quad \Longrightarrow \quad \boxed{g_{\alpha\beta} = \mathrm{diag}(-(1 + 2\phi/c^2), 1, 1, 1)}$
  - where we used $(1 + \phi/c^2)^2 = 1 + 2\phi/c^2 + \ldots$
- we have only determined here how the time-time component of the metric changes
  - but that, it turns out, is sufficient to describe motion of non-relativistic particles or matter waves in a gravitational field

3.6 The parable of the apple

- These considerations led Einstein to the radical idea that
  - the presence of matter distorts, or 'curves', space-time while
  - particles follow *geodesics*
    - paths of extremal 'length'
    - or (equivalently) paths for which the 'matter waves' constructively interfere

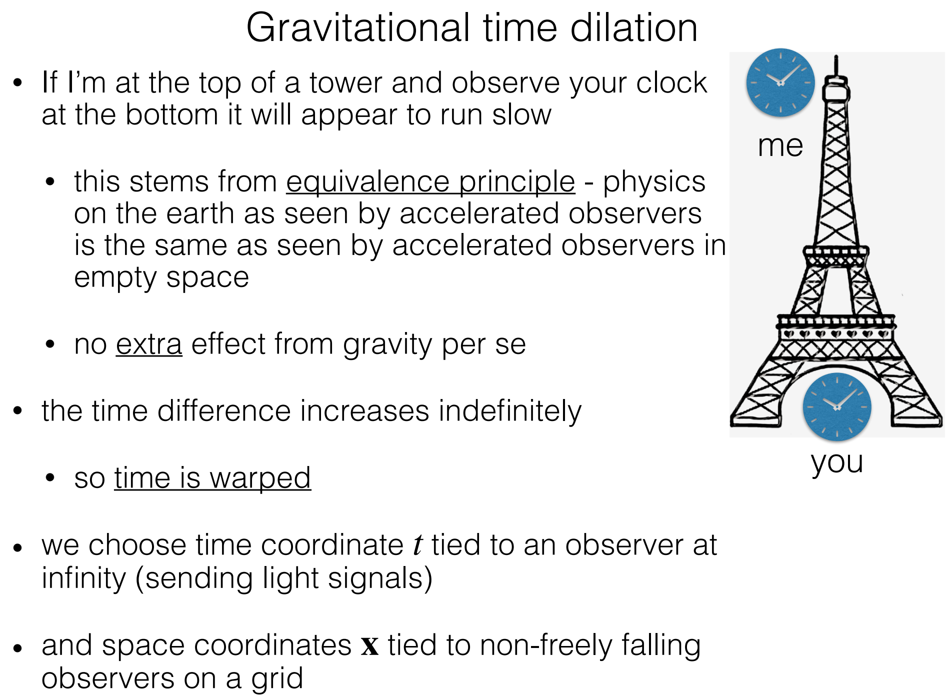

Figure 8: Gravitational time dilation.

    * the equivalent of straight lines in flat space-time
  - this is called 'the parable of the apple' in Misner, Thorne & Wheeler
  - while the geometry – implicit in the metric – is determined by the stress-energy tensor of the matter through Einsteins field equations
    * this being how *matter tells space-time how to curve*
- interestingly, though, it took a further 5 years for the final theory to emerge

3.7 Motion of particles and matter waves in a gravitational field

To get an inkling of how this works, let's see how we can use the 'weak-field' metric $g_{\alpha\beta} = \mathrm{diag}(-(1 + \phi/c^2), 1, 1, 1)$ to obtain the trajectories of particles and evolution of matter waves in a gravitational potential field

3.7.1 Particle trajectories as extremal paths

- the left panel of figure 9 shows the world line of a particle going from some position $\mathbf{x}$ at time t_1 to the same position at time t_2 in flat (Minkowski) space-time
- its world line is a straight line
  - this is a line which *extremises* $ds^2 = -c^2 d\tau^2$
    * the definition of a *geodesic*
  - it is actually a line of *maximum* elapsed proper time $\tau = \int d\tau$
  - any other line from A to B would have a shorter elapsed proper time
    * as in the 'twin paradox'
- on the right is shown a particle trajectory in weakly curved space time
  - as discussed in the caption, if we make a small displacement of the path to a region of less negative ϕ there is a 1st order increase in the proper time
  - and the extremal path has the displacement such that this 1st order rate of change is balanced by the 2nd order decrease in proper time (increase in ds^2) from the dx^2 term in the line element

- If we parameterise the path by λ then the actual path is that for which
  - $\boxed{\delta \int ds = \delta \int d\lambda \sqrt{-g_{\alpha\beta}(\vec{x}) \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda}} = 0}$
- this looks like a problem in mechanics
  - with time replaced by λ
  - and a 'Lagrangian' $L(x^\alpha, \dot{x}^\alpha) = \sqrt{-g_{\alpha\beta}(\vec{x})\dot{x}^\alpha \dot{x}^\beta}$
  - and where $\dot{x}^\alpha = dx^\alpha / d\lambda$
- the spatial generalised momentum is $p^i = \partial L / \partial \dot{x}^i = \dot{x}^i / L$
- and the Euler-Lagrange equation is
  - $d\dot{x}^i / d\lambda = \partial L / \partial x^i = \frac{1}{2} L^{-1} \dot{t}^2 \partial g_{tt} / \partial x^i$
  - since the only dependent on position in the metric is in $g_{tt} = -c^2 - 2\phi$
- this is simplest if we choose the parameterisation $\lambda = \tau$
  - this is called an *affine* parameterisation
- and the E-L equation is then simply, so linear order in ϕ,
  - $\boxed{d^2\mathbf{x}/d\tau^2 = -\boldsymbol{\nabla}\phi}$
- which is the Newtonian equation of motion

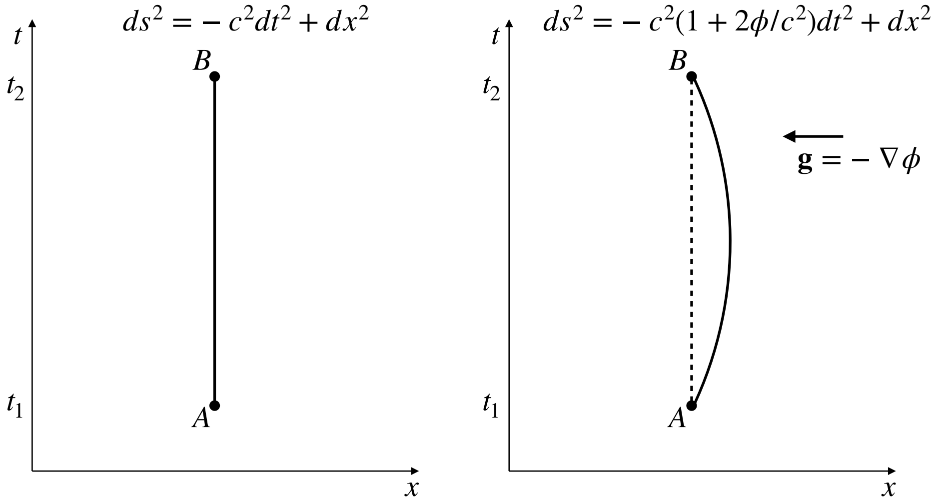

Figure 9: Extremal particle paths. In flat space-time, free particles have straight world-lines. These are paths of extremal (actually maximal) proper time. A particle that starts and ends at the same spatial location is shown at the left. In a gravitational field such a line is no longer extremal. On the right, for a negative potential $\phi < 0$ with gradient as shown, the proper time elapsed for the straight path (dashed) is less than for flat-space time, but can be increased by bending the path to positive x, with an increase that is first order in Δx. But there is also a decrease in proper time (increase in ds^2) that is 2nd order in Δx coming from the dx^2 term. If we bend the curve too far this will dominate. So there must be a curve with finite displacement – like that shown – that maximises the proper time. As shown in the text, this is the curve for which $d^2\mathbf{x}/d\lambda^2 = -\boldsymbol{\nabla}\phi$.

3.7.2 Matter waves in a gravitational potential

- we can also look how a relativistic scalar field evolves in a gravitational potential
- in the absence of a potential we can have solutions of the Klein-Gordon equation (or relativistic Schrödinger equation)
  - $\Box\psi = -c^2\omega_{\rm C}^2\psi$
  - where $\omega_{\rm C}$ is the (angular) Compton frequency
- that are simply $\psi = \psi_0 \cos(\omega_{\rm C} t)$
  - such a field has no spatial derivative, and hence no momentum density,
  - so it corresponds to cold matter at rest
  - and it can be thought of as a 'clock' with period $2\pi/\omega_{\rm C}$
- if we start with this and 'switch on' the gravitational potential $\phi(\mathbf{x})$ then the field will continue to oscillate at the Compton frequency but as a function of *proper time* τ:
  - $\psi \propto \cos(\omega_{\rm C}\tau)$
- which means that in a region of negative ϕ the phase will advance less as a function of *coordinate time* t than in regions where ϕ is larger
- i.e. simply because the 'clock' runs slow (relative to coordinate time)
- this is illustrated in figure 10
- the field will develop sinusoidal variation as a function of position
  - so a (non-inertial) observer at fixed $\mathbf{x}$ will say that the field is gaining momentum
  - though a freely in-falling observer who started at rest with the field would see the field varying as $\psi \propto \cos(\omega_{\rm C} t)$ in his frame, and so would say that there was no momentum density
- thus there is a nice correspondence between the behaviour of such fields and classical particles
  - it is in this way that a field like the axion – which is a candidate for the dark matter – would behave, for most purposes, just like particle dark matter (such as WIMPS)
- Q: Particles falling into a potential well will eventually run into one another and a 'multi-streaming region' will develop. What do you think happens to a scalar field in that situation?

4 1915: Einstein's theory of gravity

4.1 The space-time manifold

- in GR, space-time is a 'manifold' – a space of 4-dimensions that is populated by events – that is 'locally flat' (in the sense of SR)
  - a 2D analogy is the surface of a potato or any other smooth object living in Euclidean 3-space
    - one can create a (Euclidean) plane that is 'tangent' to the surface at any point
    - the closer you are to that point, the closer the actual surface lies to the tangent plane
    - the deviation grows *quadratically* with distance from the point
- in general, as on a potato, we can draw coordinates in pretty much any way we like
  - subject to some basic conditions such as different events having different coordinates
- but 'local flatness' means we can lay down locally rectilinear coordinates on the tangent space-time

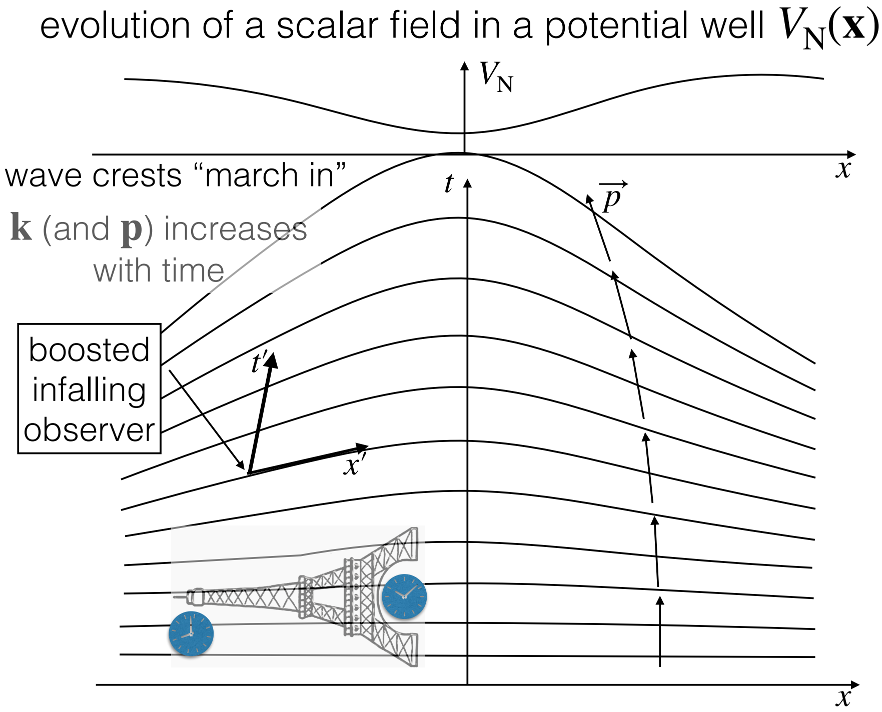

Figure 10: A scalar field in a potential well.

- physically these are the SR-like coordinates of a freely falling observer – or reference frame – with rulers and clocks to define coordinates of events and intervals etc.

- things are a bit different on the manifold of GR as compared to a potato, since the 'tangent space-time' is non-Euclidean
  - the manifold has a *light-cone* structure
  - this is intrinsic to the manifold and entirely independent of any coordinate system
  - an interesting question is whether this light-cone structure is 'orientable' in GR
    * that is to say whether one can consistently declare one direction to be the 'future' and the other to be the 'past'

4.2 The metric

- the central 'geometric entity' in GR is the *metric*
  - can be thought of as a 'machine' (or maybe a subroutine) that, given the 4D coordinate interval dx^α between two events, returns the physical interval (as measured by the freely falling observer(s))
  - i.e. $ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta$.
- this generalises the metric of SR ($ds^2 = -c^2 dt^2 + dr^2 = \eta_{\alpha\beta} dx^\alpha dx^\beta$)
  - the *components* of the metric depend on the choice of coordinate system
    * e.g. like polar or cartesian coordinates on a flat sheet of paper
  - but it has a physically real geometric existence independent of coordinate choice
- differentiating the metric provides something called the *connection*, from which we construct the all important *covariant derivative*
  - this allows us to compute how the components of particles' 4-momenta change as they move through space-time. And integrating this gives their paths.
    * we used this above for a particle moving in a weak potential

- and it provides *geodesic deviation equation* which describes how neighbouring particles get focussed or de-focussed by the gravitational field
- this is how 'space-time tells matter how to move' to borrow the phrase of John Wheeler

- and the covariant derivative allows us to transport vectors around closed loops and, just as for a 2-vector being transported on the surface of an apple, this allows us to determine the local *curvature*
  - another entity which is intrinsic to the manifold and has an absolute existence independent of coordinate system
- let's look at each of these in a bit more detail:

4.3 The covariant derivative

4.3.1 Curvilinear coordinates

- In flat space-time it is simplest to work with rectilinear (or Lorentzian) coordinates
- but we can equally well use curvilinear coordinates (see figure 11)
  - and this is forced on us in GR
  - so we *have* to do vector and tensor calculus in such coordinates
- a central idea here that a vector like $\vec{V}$ – a physically real, coordinate system independent entity – is given by a linear combination of *basis vectors* $\{\vec{e}_\alpha\}$: $\vec{V} = V^\alpha \vec{e}_\alpha$.

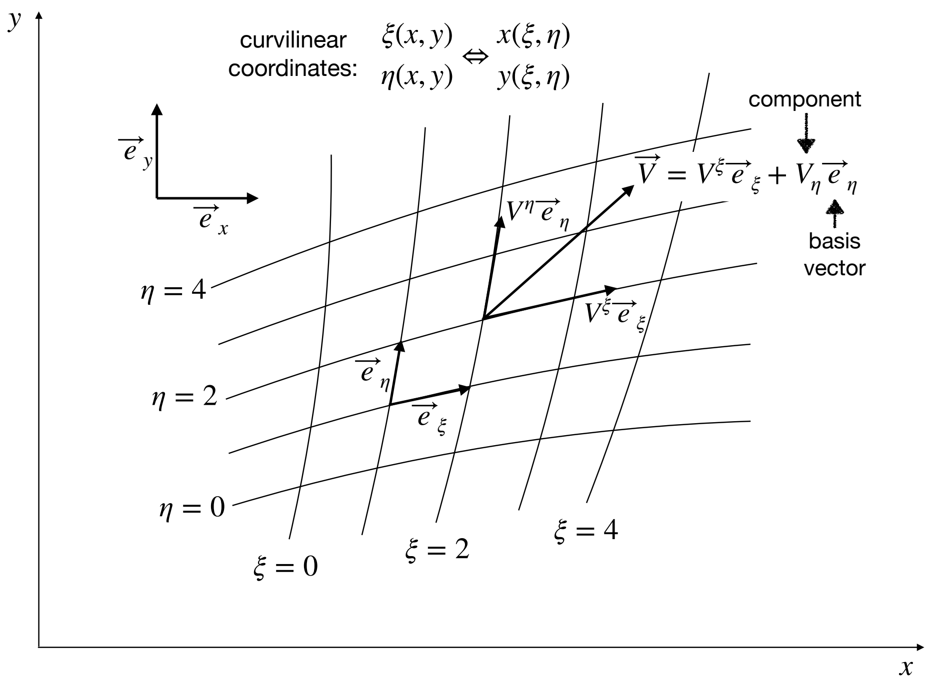

Figure 11: Curvilinear coordinates in 2 dimensions. Given 2 functions ξ, η of Euclidean x, y coordinates we can create a pair of *basis vectors* $\vec{e}_\xi$ and $\vec{e}_\eta$ and express any vector $\vec{V}$ as a sum of components times basis vectors: $\vec{V} = V^\alpha \vec{e}_\alpha$. There is considerable freedom in how to do this. The choice illustrated here uses the so-called *coordinate basis vectors*.

4.3.2 Derivative of a vector field

- and given a *vector-field* – i.e. some $\vec{V}(\vec{x})$ that varies smoothly with position $\vec{x}$ – we can talk about its derivative
  - we can ask: what is the rate of change of some vector field $\vec{V}(\vec{x})$ with respect to proper time as observed by a particle with 4-velocity $\vec{U}$

- or equally the rate of change of some property of the particle itself like its 4-momentum

- in rectilinear coordinates the basis vectors point along the coordinate axes and so they are independent of position and all we need to know are the partial derivatives of the components of the vector with respect to the coordinates ${V^\alpha}_{,\beta}$

- but in curvilinear coordinates we need to worry about the variation of the basis vectors with position too

- the upshot of this is the *covariant derivative*, denoted by ${V^\alpha}_{;\beta}$

- this contains, in addition to the partial derivatives ${V^\alpha}_{,\beta}$ an additional term

  - $\boxed{{V^\alpha}_{;\beta} = {V^\alpha}_{,\beta} + {\Gamma^\alpha}_{\mu\beta} V^\mu}$
  - where ${\Gamma^\alpha}_{\mu\beta}$ is called the *connection* or the *Christoffel symbols*
  - and which tells us how the basis vectors $\{\vec{e}_\alpha\}$ vary
  - and this, it turns out, is computable from the partial derivatives of the metric components

- In addition to telling particles how to move in a gravitational field, the covariant derivative allows us to obtain equations of motion and conservations laws in curved space-time

  - we simple replace ordinary partial derivatives by covariant derivatices
  - the so called 'comma ⇒ semicolon rule'
  - which is called *'generalised covariance'*

- and it allows one to *parallel transport* vectors

  - i.e. to figure out how their components of a vector change if they remain physically constant
  - as we did, for instance, for a particle moving in a potential above

4.4 The curvature tensor

- if you transport a vector around a closed loop on a curved surface you will find that it changes

  - whereas if the surface were flat it would remain unchanged

- this is described mathematically by the 'curvature tensor' **R**

  - given, as arguments, two vectors $\vec{A}$ and $\vec{B}$ that define a little parallelogram
  - and another vector $\vec{V}$ that we are going to transport
  - the *Riemann tensor* **R** returns the change $\delta\vec{V}$
    - * this means that the description of **R** is a rank-4 tensor (it has 4-indices).

- The curvature is also useful as it tells us how the separation between a pair of initially parallel world-lines changes

  - this is 'geodesic deviation' – or tidal deflection

- so the curvature plays here exactly the same role as the tidal field in Newtonian gravity.

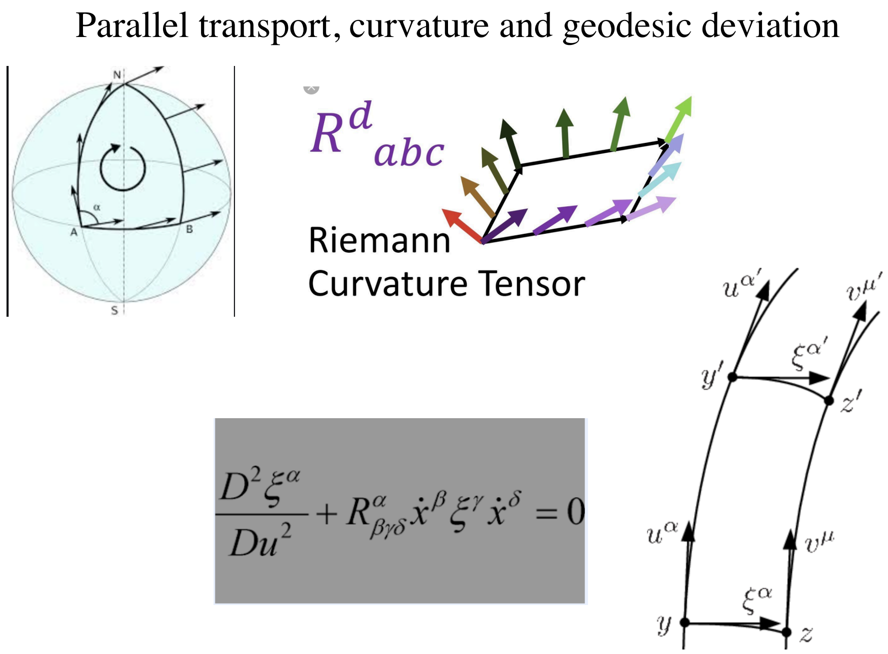

Figure 12: Riemann curvature tensor. The upper figures illustrate the definition of the Riemann tensor. The lower right figure is particularly informative: Let's say we start with a vector $\vec{u}$ (bottom left) that represents the 4-momentum of a particles. We can parallel transport this along the 'separation vector' $\vec{\xi}$ to make a vector $\vec{v}$; the 4-momentum of a neighbouring particle. These vectors are initially parallel. Now let them move freely (parallel transporting their 4-momenta). We can then parallel transport one of these 4-momenta over to the other particle's path and compare them. If they are the same, the 4-momenta have remained parallel. But this will only be the case if there is no matter present. If matter is present, or if there are gravitational fields coming from elsewhere – maybe gravitational waves or tidal field – the parallelism will have been disturbed. Once we have the rate of change of the relative momenta we can compute how the separation vector $\vec{\xi}$ will evolve. This will change, to lowest order, quadratically with distance along the path, as described by the equation in the box. This is the geodesic deviation equation; analogous to what we found for the Newtonian tidal field. This allowed Einstein to obtain his field equations by requiring that they reduce to Newtonian gravity in the appropriate limit.

4.5 The field equations of GR

- Recall that in Newtonian gravity the gravitational field is the tide ϕ_{ij}
  - a symmetric 3-tensor with two spatial indices
- but what appears in Poisson's equation, and gets equated to the 'source' – ($4\pi G$ times) the mass density – is a *contraction* of the tidal field: the Laplacian of ϕ (or the trace of ϕ_{ij})
- once one solves Poisson's equation for ϕ – invoking suitable boundary conditions at infinity – one can differentiate it to get the tide
- SR leads us to suspect that the source in GR is not ρ, but the rank-2 stress-energy tensor $\mathbf{T}$
- and, by analogy, one might expect that is is some kind of contraction of the rank-4 curvature tensor $\mathbf{R}$ that gets equated to this source
- it turns out that there is a unique rank-2 contraction of $\mathbf{R}$ called $\mathbf{G}$ (the Einstein tensor) that obeys the very same 'continuity equation' as $\mathbf{T}$
- this gives Einstein's equations:
  - $\boxed{\mathbf{G} = 8\pi\mathbf{T}}$

- the fact that $\mathbf{G}$ is essentially the only tensor one can make out of $\mathbf{R}$ that has the same properties as $\mathbf{T}$ makes this theory essentially unique

- though a modification, proposed later by Einstein, is
  - $\boxed{\mathbf{G} = 8\pi\mathbf{T} + \Lambda\mathbf{g}}$
  - where Λ is known as the *cosmological constant*

- the parallels with Newtonian gravity are very close:
  - Newton: $\phi \quad \Rightarrow \quad \mathbf{g} = -\boldsymbol{\nabla}\phi \quad \Rightarrow \quad \boldsymbol{\nabla}\boldsymbol{\nabla}\phi$
  - Einstein $\mathbf{g} \quad \Rightarrow \quad$ connection $\quad \Rightarrow \quad \mathbf{R}$

- and in both cases we use the last quantity to
  - calculate how paths of particles get 'focussed'
  - and we equate its 'contraction' to the matter density
  - A fundamental *difference*, however, is that whereas Newtonian gravity plays out in an arena where space and time are given *a priori*, in Einstein's gravity the geometry of space-time emerges as the solution of the field equations.

- What does all this give us?
  - we recover Newtonian gravity in the limit $v \ll c$
  - modification to orbits for e.g. Mercury
  - modification to light bending (the extra factor 2)
  - gravitational collapse of massive stars
    * pressure as well as density appears in $\mathbf{T}$
  - gravitational waves
  - cosmology

4.6 Some comments and questions

- Q: are gravity and acceleration really equivalent? As many popular accounts portray the equivalence principle.
  - A: No! If you are in empty space but being accelerated by a rocket motor then the curvature $\mathbf{R}$ in your neighbourhood vanishes.
  - To the extent that curvature – or the tidal field in Newtonian language – is *the* gravitational field, nothing could be further from the truth.
  - There are of course phenomena – like weightfulness if you are standing on the surface of the Earth – that are also experienced by accelerated observers. But these are better seen as indirect consequences of gravity.

- Q: Does GR explain Mach's principle (or observation)?
  - Not really. While there are ways in which the motion of matter – rather than just the density of matter – affects space-time (so called 'frame-dragging' in the vicinity of a rotating mass being the prime example), this does not really explain what so intrigued Mach and Einstein.
  - It is often said that Mach's principle played the role of 'midwife' in helping the birth of GR, but no longer provides a useful service

- Q: what, to order of magnitude, is the 'radius of curvature' of space-time in the vicinity or interior of the earth?

- Q: consider a community of blind, but intelligent, ants living on the surface of an egg. They have some pieces of string they can use to measure distances on the surface of the egg. And they have protractors to measure angles. What can they deduce – from *local* measurements only – about the curvature of the surface? In particular, how many numbers are required to express the results of their studies? Can they figure out – again, from purely local measurements – what is the direction to the pointy end?

5 Road-Map

- In the following 2 chapters we will develop the mathematical language of GR
- In the next one we consider the flat space-time of special relativity but using arbitrary curvilinear coordinates
  - this is known as *generalised covariance* (though, confusingly, this term was initially used to describe GR itself)
  - in the process we will develop the covariant derivative and show how the connection that appears therein is related to the metric
  - most of which is is largely independent of the non-Euclidean nature of SR, and we will work a lot in 2D Euclidean space
- after that we will do *differential geometry* on curved manifolds
  - we introduce the Riemann curvature
  - and obtain Einstein's field equations