Advanced Gravity

Antonio Padilla[1]

School of Physics and Astronomy, University of Nottingham, Nottingham NG72RD, UK

Abstract

These lecture notes cover the Advanced Gravity module (F34AGR) forming part of the Gravity, Particles and Fields Masters course. General Relativity is based on the geometry of four dimensional spacetime, the curvature of which is governed by the Einstein's equations. This theory is extremely well tested and represents our best description of the gravitational force. This module develops the ideas behind General Relativity to an advanced level. The derivation of certain solutions to the Einstein equations will be presented, including black hole and cosmological solutions. Gravity in the weak-field limit will be derived from the full theory, demonstrating how one should understand the gravitational interaction in terms of graviton exchange. The module will then move on to advanced topics, including modified gravity models (e.g. models with extra dimensions) that are at the forefront of current research.

[1] `antonio.padilla@nottingham.ac.uk`

Useful resources

- D'Inverno, *"Introducing Einstein's Relativity"*
- Carroll, *"Spacetime and Geometry"*
- Hartle, *"Gravity: an introduction to Einstein's General Relativity"*
- Zee, *"Quantum field theory in a nut-shell"*.

For more advanced stuff:

- Wald, *"General Relativity"*
- Misner, Thorne, & Wheeler, "Gravitation"
- Parker & Toms, *"Quantum fields in curved spacetime"*

Contents

"And then I looked up at the Sun
And I could see
Oh, the way that gravity pulls on you and me"

Coldplay, *Gravity*

1 What is gravity?

"*The laws in this city are clearly racist. All laws are racist. The law of gravity is racist.*" So said Mayor Marion Barry of Washington D.C. We cannot comment on the laws in America's capital, but we can say that Mayor Barry was wrong about gravity. Gravity does not care about how we look, or how we are put together. I suspect that Mayor Barry did not understand the principle of equivalence.

To most people, gravity is what keeps your feet on the ground. It is the force that caused the apple to fall on Newton's head. It is the force that keeps the earth in orbit around the Sun. It is even the force that controls the expansion of the entire Universe! Newton described gravity using his famous inverse square law, a law from which we can derive those planetary orbits with great accuracy. We can even get cosmology from those laws!

A particle physicist would take a more sophisticated view. To him, gravity, like any other force, comes from the exchange of "virtual" particles. The particle that mediates this force is often referred to as the graviton, and most likely has spin two, and vanishing mass. A string theorist would suggest that this graviton is not a particle, but a closed string, like a loop of finite size.

In your *Gravity* course you learned that *Gravity is Geometry.* This course is largely aimed at providing us with the necessary mathematics to add meat to the bones of that statement.

1.1 The Equivalence Principle

The statement that *Gravity is Geometry* stems from the Equivalence Principle, or perhaps more precisely, the Einstein Equivalance Principle (EEP). The reason we are being pedantic about this is that the Equivalence Principle comes in many forms, with some subtle differences:

Weak Equivalence Principle (WEP) The trajectory of an uncharged, freely falling test particles is independent of its mass or composition.

Einstein Equivalence Principle (EEP) The WEP holds, and furthermore, local non-gravitational experiments performed in a freely falling laboratory yield the same results as those performed in Minkowski spacetime, regardless of the position and velocity of the laboratory.

Strong Equivalence Principle (SEP) The WEP holds for massive gravitating objects as well as test particles, and furthermore, local non-gravitational experiments performed in a freely falling laboratory yield the same results as those performed in Minkowski spacetime, regardless of the position and velocity of the laboratory.

If we are to discuss what we been by each of these, we first need to remind ourselves what is meant by a freely falling test particle, or a freely falling lab. A freely falling particle is one whose motion is subject only to gravity. In other words, it only feels acceleration due to gravity, and not any other force. When you stand on the surface of the earth, you are *not* freely falling because your feet are subject to those electromagnetic forces that stop you from plummeting towards the earth's core. In contrast, if we neglect the effects of friction, a man crazy enough to jump from the top of the physics building would indeed be freely falling, as the only force affecting his motion is gravity. Note that a laboratory on a space station orbiting the earth is also freely falling since its motion is entirely governed by the earth's gravitational field. Sometimes you will find that people use *inertial* rather than *freely falling*. They mean the same thing: *subject only to gravity*.

The WEP is the generalisation of Galileo's famous claim, that a moving body falls with uniform acceleration, independently of its mass or composition. In other words, when a canon ball and a wooden ball are released simultaneously form the top of the leaning tower of Pisa, they hit the ground at the same time. It is sometimes described as *the Universality of Free Fall*. The EEP tells us that provided we restrict our attention to a small enough region of spacetime, we can eliminate the effects of gravity to the point where our (non-gravitational) experiments become indistinguishable from those performed in Minkowski spacetime. The SEP merely extends the scope of the EEP, taking into account the trajectory through space of self gravitating objects.

The WEP is easily incorporated into our geometrical picture by demanding that all test particles move along geodesics – the shortest paths through spacetime. In flat spacetime, the shortest paths are just straight lines, but in curved spacetime the shortest paths can be curved. In direct analogy, because the surface of the earth is curved, the shortest path, or geodesic, connecting London to New York, is not a straight line, but curved along the great circle.

The EEP holds because Minkowski spacetime is a good approximation within a small enough neighbourhood of any point in curved spacetime. Again, in direct analogy, recall that the ancients believed the earth was flat. This is precisely because its surface *looks* flat when viewed over small enough regions. The EEP also allows us to extend the statement about test particles to massless species such as light, which are now required to move along *null* geodesics (see section 1.2).

At this stage it is worth reminding ourselves of the Machian element to our ideas about gravity. Mach's principle states that the distribution of all matter relative to a given observer determines his/her inertial frame. In our geometrical view of gravity, this is equivalent to saying that the distribution of all matter determines the curvature of spacetime. In General Relativity this is almost true. Einstein's equation relates the energy and momentum of matter to the spacetime geometry, but it does leave room for gravitational waves, which are freely propagating and technically violate Mach's principle. Some alternative gravity theories, such as Brans-Dicke, permit additional fields to play a role in determining the geometry, but the spirit of Mach's principle remains intact up to freely propagating waves.

Finally, we return to the SEP. It essentially extends the EEP to apply to bodies with significant gravitational self interaction. In the geometrical picture, it requires gravity to be *entirely* geometrical, in the sense that all we have is matter and geometry. The former completely determines the latter, up to gravitational waves, so there are no additional fields playing any role. This picks out General Relativity (plus cosmological constant) uniquely in four dimensions.

1.2 Some details from the *Gravity* course

Let us pause to recap a few details from the *Gravity* course. There we learnt that curved spacetime is described using a metric

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu \tag{1.1}$$

where $\mu = 0, 1, 2, 3$. This tells us how to measure spacetime distance in the curved geometry. In this course, we will adopt the convention that the signature of the metric is $(- + ++)$. For example, this would mean that Minkowski spacetime would be written as $ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$. You will find that some old fashioned people prefer to use the convention $(+ - --)$, in which Minkowski would be written as $ds^2 = dt^2 - dx^2 - dy^2 - dz^2$. Its just convention. Like we said, we will adopt the "mostly +" convention in this course. Note that we will also work in units where the speed of light, $c = 1$, unless explicitly stated.

Provided that the metric is non-degenerate, that is $\det g_{\mu\nu} \neq 0$, we can always define the inverse metric $g^{\mu\nu}$ satisfying $g^{\mu\nu}g_{\nu\alpha} = \delta^\mu_\alpha$. The metric and inverse metric are useful for turning a *covariant* vector V_μ into a *contravariant* vector $V^\mu = g^{\mu\nu}V_\nu$, and vice-versa, $V_\mu = g_{\mu\nu}V^\nu$. We can also use them to define the Levi-Civita connection,

$$\Gamma^\mu{}_{\alpha\beta} = \frac{1}{2}g^{\mu\nu}\left(g_{\nu\alpha,\beta} + g_{\nu\beta,\alpha} - g_{\alpha\beta,\nu}\right) \tag{1.2}$$

where we have used the following shorthand: $g_{\mu\nu,\alpha} = \frac{\partial}{\partial x^\alpha}g_{\mu\nu}$.

The geodesic paths $x^\mu = x^\mu(\lambda)$, parametrised by some parameter, λ, can be found by minimising the spacetime distance

$$\int \sqrt{\left|g_{\mu\nu}(x(\lambda))\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right|}d\lambda \tag{1.3}$$

This yields the following differential equation

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu{}_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = \alpha(\lambda)\frac{dx^\mu}{d\lambda} \tag{1.4}$$

where $\alpha(\lambda) = \frac{d}{d\lambda}\ln\sqrt{\left|g_{\mu\nu}(x(\lambda))\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right|}$. If λ is an *affine* parameter then we have

$$g_{\mu\nu}(x(\lambda))\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = \text{constant}$$

and so

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu{}_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \tag{1.5}$$

One can always find an affine parameter for a geodesic. We should also recall that in spacetime, geodesics can be categorised as timelike, null or spacelike. This depends on the properties of the tangent vector to the geodesic, $u^\mu = \frac{dx^\mu}{d\lambda}$. Given our signature convention, we have that $u_\mu u^\mu$ is negative, zero, or positive for timelike, null and spacelike geodesics respectively.

In General Relativity, ordinary massive test particles follow timelike geodesics, since their action is given by

$$S = -m \int \sqrt{-g_{\mu\nu}(x(\lambda)) \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda \tag{1.6}$$

Using the EEP we extend this philosophy to massless particles like the photon, and require them to follow null geodesics. If they existed, which they do not, we would also say that nonsensical particles with imaginary mass (tachyons) follow spacelike geodesics.

A useful short-cut for calculating geodesics is as follows. We simply assume an affine parametrisation and set

$$M^2 = -g_{\mu\nu}(x(\lambda)) \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \tag{1.7}$$

with the constant M^2 being positive, zero, or negative for timelike, null and spacelike geodesics respectively.

Finally, we note that throughout this course we will make use of the following shorthand denoting symmetrization and anti-symmetrization of indices,

$$Q_{(\mu\nu)} = \frac{1}{2} Q_{\mu\nu} + \frac{1}{2} Q_{\nu\mu}, \qquad Q_{[\mu\nu]} = \frac{1}{2} Q_{\mu\nu} - \frac{1}{2} Q_{\nu\mu}$$

Exercises

1. Suppose we have a theory with a metric, $g_{\mu\nu}$ and a scalar field, ϕ, such that the action for a test particle of mass, m, is given by

$$S = -m \int e^{4\phi(x(\lambda))} \sqrt{-g_{\mu\nu}(x(\lambda)) \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda \tag{1.8}$$

   Which of the following holds: WEP, EEP, and/or SEP? Explain your answer.

2. Suppose we have a theory with two metrics, $g_{\mu\nu}$, and $h_{\mu\nu}$, such that the action for a test particle of mass, m, is given by

$$S = -m \int \left[\sqrt{-g_{\mu\nu}(x(\lambda)) \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} + \sqrt{-h_{\mu\nu}(x(\lambda)) \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} \right] d\lambda \tag{1.9}$$

   Now, one can always choose coordinates so that either $g_{\mu\nu}$ or $h_{\mu\nu}$ is locally Minkowski. However, in general, there exists no choice of coordinates for which both metrics are locally Minkowski at the same time. Therefore, which of the following holds: WEP, EEP, and/or SEP? Explain your answer.

3. The Schwarzschild geometry is given by

$$ds^2 = -V(r)dt^2 + \frac{dr^2}{V(r)} + r^2(d\theta^2 + \sin^2\theta d\phi^2), \qquad V(r) = 1 - \frac{r_H}{r} \tag{1.10}$$

Show that the radial null geodesics are given by

$$\left(\frac{r}{r_H} - 1\right) e^{r/r_H} = Ae^{t/r_H}, \qquad A = \text{constant} \tag{1.11}$$

for $r > r_H$.

2 General Relativity: the successes

Nearly 100 years after its inception, General Relativity has time and again stood up to experimental scrutiny. By now there are many different tests of GR, but three in particular are regarded as the classics.

2.1 The perihelion precession of Mercury

According to Kepler's first law, the motion of each planet in the solar system is given by an ellipse with the Sun at its focus. Viewed on short enough timescales, these ellipses appear fixed, but a more careful examination reveals this not to be the case. The point of closest approach (the *perihelion*) is seen to precess (rotate) around the Sun. This precession can usually be accounted for in Newtonian gravity by corrections due to the gravitational fields of the other planets in the Solar System. However, most notably in the case of the planet Mercury, this is not enough, leading Urbain Le Verrier to speculate on the existence of a new unseen dark planet that he dubbed "Vulcan". However, if we include the corrections from General Relativity we find that Mercury's perihelion precession can be reproduced to excellent accuracy without any reference to the mythical "Vulcan". Being closest to the Sun, Mercury is more sensitive to GR corrections than the other planets. Given that average radius of Mercury's orbit is of the order 56,000,000 km and the Schwarzschild radius of the Sun is of the order 3km, so the GR corrections are of order 1 part in 10^7. The anomalous precession of Mercury is 42.98 ± 0.04 arcsecs/century, which is emphatically consistent with the GR prediction of 42.98 arcsecs/century.

For details on how to derive the GR correction to the perihelion precession of a planet, see *D'Inverno*, Section 15.3.

2.2 The bending of light

Light bending per se is not unique to a geometric description of gravity. If we embrace the WEP and model the photon as having an infinitesimally small mass we obtain light bending even in Newtonian gravity. To see this consider a photon passing a heavy object of mass, M. By extending the WEP we conclude that the photon feels a radial acceleration $a = GM/r^2$ pointing towards the massive object, and thus its trajectory is deflected by some angle, θ_N. Eddington was the first to find experimental evidence of starlight bent by the gravitational field of the Sun in 1919. However, for a photon grazing the edge of the Sun, the Newtonian prediction , $\theta_N = 0.875$ arcsecs is not consistent with observation. Modern tests give a deflection angle of $\theta = (0.99992 \pm 0.00023) \times 1.75$. In General Relativity, one can compute the path of a lightray by calculating the null geodesics in the curved geometry around a heavy object. The same photon trajectory is found to be deflected by an angle $\theta_{GR} = 2\theta_N = 1.75$ arcsecs, and is in excellent agreement with observation.

For details on how to compute the bending of light around a heavy object in General Relativity, see *D'Inverno*, Section 15.4.

2.3 Gravitational redshift

Any theory that satisfies the Einstein Equivalence Principle will experience gravitational redshift. This is the redshifting of light as it climbs out of a gravitational potential well. Conversely, light is blueshifted as it falls down the potential well, which is why the accretion disks of black holes appear so energetic.

To understand this effect, consider the following thought experiment. Imagine a giant water wheel running between the earth at the top and the Sun at the bottom. This water wheel is modified slightly so that on one half of the wheel the atoms in the water buckets are excited, but in the other half they lie in their ground state. By mass-energy equivalence, the buckets of excited atoms are "heavier" and will fall towards the Sun, causing the wheel to rotate. At some point we reach a state of equilibrium with the lowermost bucket of excited atoms arriving at the Sun (bucket A), and the uppermost bucket of ground state atoms arriving at the earth (bucket B). At this point, there is a device on the Sun that returns the excited atoms in bucket A to their ground state and extracts the energy difference as radiation. This radiation is then beamed back to earth where it is used to excite the atoms in bucket B. Because bucket A is now a little lighter, and bucket B a little heavier, the wheel rotates some more, and the process is repeated. We are left with a state of perpertual motion, which would seem to be in violation of conservation of energy.

The resolution of this puzzle lies in the fact that the radiation beamed back to earth is redshifted as it climbs out of the gravitational potential well sourced by the Sun. Redshifting causes the frequency, ν of the photon to decrease, and by the Planck-Einstein equation, $E = h\nu$, this leads to a reduction in its energy. In other words, the photon must do work to climb out of the potential well, and this drop in energy manifests itself through a drop in the frequency.

Gravitational redshift was famously verified in 1959 by the Pound-Rebka experiment between the top and bottom of the Jefferson tower at Harvard University.

Further details on gravitational redshifting can be found in *D'Inverno*, Section 15.5.

Exercises

1. In Newtonian gravity, a particle in the gravitational field of the Sun follows a trajectory given by

$$r = \frac{r_{min}(1+\epsilon)}{1+\epsilon\cos\theta}, \qquad \epsilon = \sqrt{1 + 2\mathcal{E}\frac{h}{GM_\odot}} \tag{2.1}$$

where (r, θ) represent polar coordinates in the plane of motion, with the Sun at the origin and r_{min} is the point of closest approach. We assume a photon has energy per unit mass, $\mathcal{E} = \frac{1}{2}c^2$, and angular momentum per unit mass, $h = r_{min}c$. Show that a light ray is deflected through an angle,

$$\delta = \left(\theta_{in} - \frac{\pi}{2}\right) + \left(\frac{3\pi}{2} - \theta_{out}\right) = 2\sin^{-1}(1/\epsilon) \approx \frac{2GM_{\odot}/c^2}{r_{min}} \tag{2.2}$$

assuming $r_{min} \gg 2GM_{\odot}/c^2$. [Hint: see Fig. 2.1.]

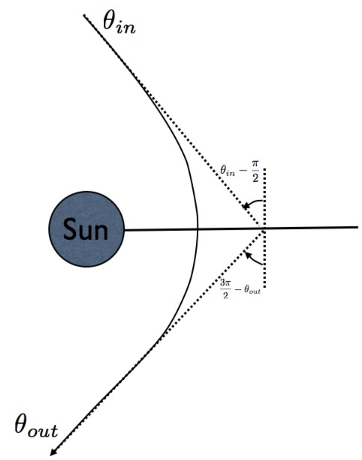


Figure 2.1: Path of a light ray bent by the gravitational field of the Sun.

2. In General Relativity, a light ray in the gravitational field of the Sun follows a trajectory approximately given by

$$r = \frac{r_{min}}{\cos\theta + \frac{2GM_{\odot}/c^2}{r_{min}}\left(1 - \frac{1}{2}\cos^2\theta\right)}, \tag{2.3}$$

Show that it is deflected through an angle,

$$\delta \approx \frac{4GM_{\odot}/c^2}{r_{min}} \tag{2.4}$$

assuming $r_{min} \gg 2GM_{\odot}/c^2$. [Hint: for small deflection angles you may neglect $\cos^2\theta_{in}$ and $\cos^2\theta_{out}$.]

3. Browse Clifford Will's *Living Review* article to familiarise yourself with some of the latest results on experimental tests of GR:

Clifford M. Will, *"The Confrontation between General Relativity and Experiment"*, Living Rev. Relativity 9, (2006), 3. URL: http://www.livingreviews.org/lrr-2006-3

3 General Relativity: the failures

Lets not get carried away. GR is good, but its not that good. Indeed, it predicts its own demise at short distances, whilst at large distances it faces serious challenges from the dark side.

3.1 Non-renormalisability

As you will learn in your black hole course, when stars run out of fuel, they begin to collapse under their own weight. For the heaviest stars, there is no process that can permanently halt the collapse, ensuring that the star forms a black hole, with a *singularity* at its centre. This singularity corresponds to a region of infinite spacetime curvature, where the classical equations of General Relativity break down. Indeed Hawking and Penrose proved that, given reasonable energy conditions on matter, such singularities are inevitable in General Relativity. To resolve them, one must resort to quantum theory beyond the Planck scale $M_{pl} = \sqrt{\hbar c/8\pi G_N}$, where G_N is Newton's constant. The trouble is that GR is not a renormalisable theory – loop corrections at higher and higher order generate a never-ending set of counter-terms. This essentially follows from the fact that the graviton propagator goes like $1/k^2$, and that the gravitational coupling, G_N, has negative mass dimension, $[G_N] = -2$. For more details on what makes a theory non-renormalisable, see Chapter III.3 of *Zee*.

3.2 The dark side

General Relativity also faces a serious challenge from the dark side. The problem is that General Relativity, coupled to ordinary baryonic matter cannot account for a slew of astrophysicial and cosmological data. To get around this we fudge the issue, by postulating the existence of some invisible fields, labelled dark matter and dark energy, between them making up 96% of the cosmic energy budget!

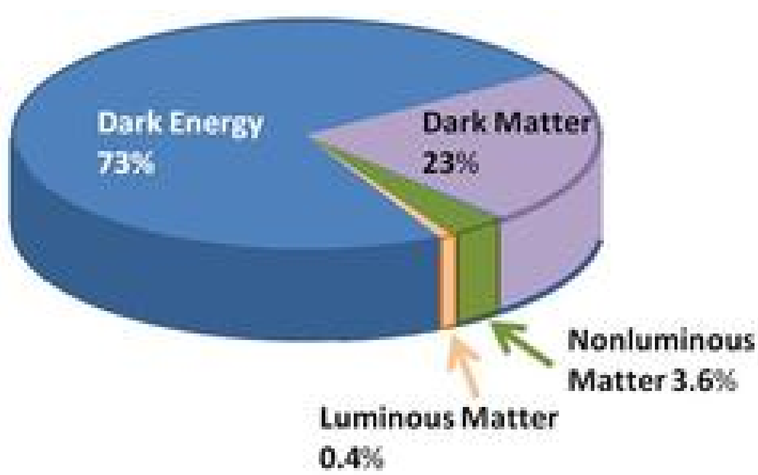


Figure 3.1: The cosmic energy budget today.

The motivation for dark matter comes from a number of sources, but most notably from the rotation curves of the outer stars in galaxies. Let us model what we *expect* this profile to be according to Newtonian gravity (which corresponds to the leading order non-relativistic limit of GR). We expect the gravitational field to be dominated by the large (and constant) baryonic mass, M_g contained in the visible core of the galaxy. Assuming circular orbits for simplicity, one would find that the velocity profile, v, of the outer stars, is governed by the following equation

$$\frac{v^2}{r} = \frac{GM_g}{r^2} \qquad \Longrightarrow \qquad v \propto \frac{1}{\sqrt{r}} \tag{3.1}$$

However, as we see in Fig. 3.2, the observed profile has $v \approx$ constant. We can recover this by assuming that the mass of the galaxy grows like $M_g(r) \propto r$, with the extra contribution coming from dark matter. The claim is that we cannot see dark matter since it barely interacts with the standard model fields. Strong evidence for dark matter also comes from gravitational lensing experiments, structure formation and measurements of the Cosmic Microwave Background (CMB). In the standard cosmological model, dubbed ΛCDM, dark matter makes up around 23% of the energy density of the Universe (see Fig. 3.1). Particle physics promises some good dark matter candidates in the form of WIMPS, Weakly Interacting Massive Particles, often corresponding to the lightest supersymmetric particle. Nonetheless, there have been attempts to modify gravity as an alternative to dark matter, most notably MOND (Modified Newtonian Dynamics) and its relativistic cousin TeVeS (Tensor-Vector-Scalar theory).

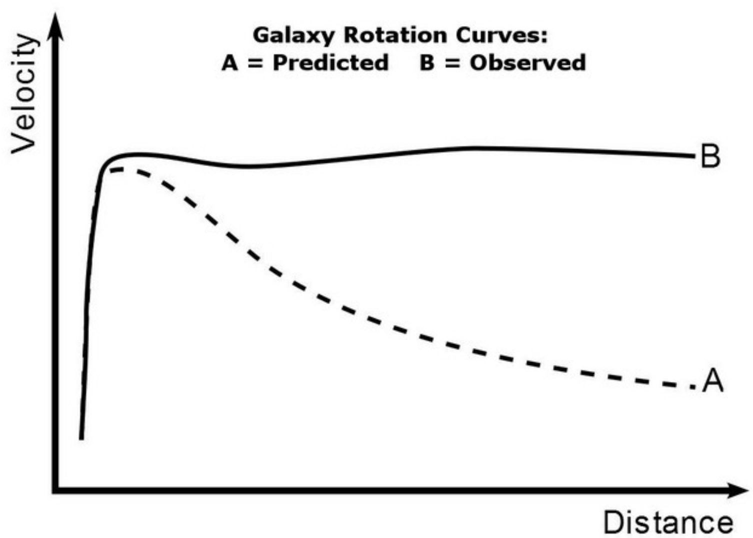


Figure 3.2: Velocity profiles for outer stars in a typical galaxy: newtonian prediction versus observation.

The motivation for dark energy comes primarily from CMB, large scale structure surveys, and, of course, supernovae observations. The latter won Perlmutter, Schmidt and Riess the nobel prize in 2011. These observations all point to one startling fact – on the largest scales the expansion of the Universe is speeding up. This is counter-intuitive: gravity is an attractive force, so surely the expansion should be slowing down. But it isn't. What is causing this repulsion on large scales. We wave our hands and say "dark energy", but the truth is that particle physics offers no good candidates for this. Some would argue that a cosmological constant, Λ, with constant energy density $\rho_\Lambda \sim M_{pl}^2 \Lambda \sim (\text{meV})^4$ can account for the observed acceleration. This is indeed true, and represents the main ingredient of the ΛCDM model, with Λ making up 73% of the cosmic energy budget (see Fig. 3.1).

However, within the realm of effective field theory, the problem with a cosmological constant is that it receives contributions from the zero point energies of each particle with mass below the cut-off. Indeed,

$$\rho_\Lambda^{obs} = \rho_\Lambda^{bare} + \frac{1}{2} \sum_{\text{particles}} \int \frac{d^3k}{(2\pi)^2} \sqrt{\vec{k}^2 + m^2} \approx \rho_\Lambda^{bare} + c_\nu m_\nu^4 + c_e m_e^4 + \ldots M_{\text{cut-off}}^4 \tag{3.2}$$

where the c_i are order one coefficients, and m_ν is the neutrino mass, $m_e \sim$ MeV is the electron mass, and so on. The electron contribution already requires us to fine tune the bare value of ρ_Λ^{bare} to one part in 10^{36}, and yet we claim to have a good understanding of physics at this scale! If we work up to a Planckian cut-off, we require a fine tuning of one part in 10^{120}. This is truly a disaster for particle physics, and is known as the cosmological constant problem. Because of it, there is strong motivation for searching for gravitational alternatives to dark energy, a field of physics which has blossomed in recent years.

Exercises

1. Verify the right-hand side of the formula (3.2).
2. The Bullet cluster is made up of two colliding galaxies, shown in Fig. 3.3. The pink bits represent the location of ordinary baryonic matter, where as the blue bits represent the location of the hypothetical dark matter, calculated using gravitational lensing experiments. Explain why this observation suggests that the dark matter problem will be solved using particle physics as opposed to a modfication of gravity such as MOND, or TeVeS.

4 Manifolds

We have said that *Gravity is Geometry*. So what is geometry? Well in General Relativity, geometry is described by a *Lorentzian Manifold* (sometimes called a *pseudo-Riemannian manifold*). The words *Lorentzian* and *pseudo* just indicate that we are interested in manifolds endowed with a metric of Lorentzian signature, $(-, + \ldots +)$, consistent with some notion of space-*time*.

Let us forget about signature for a moment and concentrate on what we mean by plain old *Riemannian* manifolds (without the *pseudo* bit). In n dimensions, the simplest example of a Riemannian manifold is n-dimensional Euclidean space $\mathbb{R}^n$. Hopefully you will already

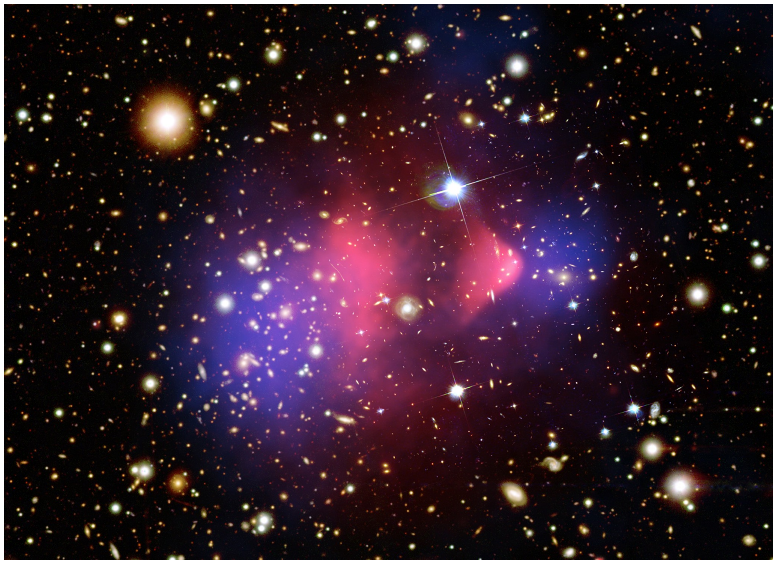


Figure 3.3: The bullet cluster.

know a lot about this space, at least in 3 dimensions. You will know that we can choose a Cartesian coordinates so that this manifold is endowed with the simplest of metrics,

$$ds^2 = dx_1^2 + dx_2^2 + \ldots + dx_n^2 \tag{4.1}$$

You may even have some idea of what is meant by a vector or a tensor, and you will probably know that we can do calculus on this space.

More generally, a Riemannian manifold is a space made up of a series of patches that are smoothly sewn together. Each patch should *look like* $\mathbb{R}^n$, when viewed on a small enough neighbourhood of any point in the manifold. Other examples of Riemannian manifolds include the n-sphere, S^n, the n-torus, T^n, as well as much more complicated objects. One of the things that characterises a manifold is its *genus*, which measures the number of holes. Manifolds with different genera (plural of genus) are shown in Fig. 4.1.

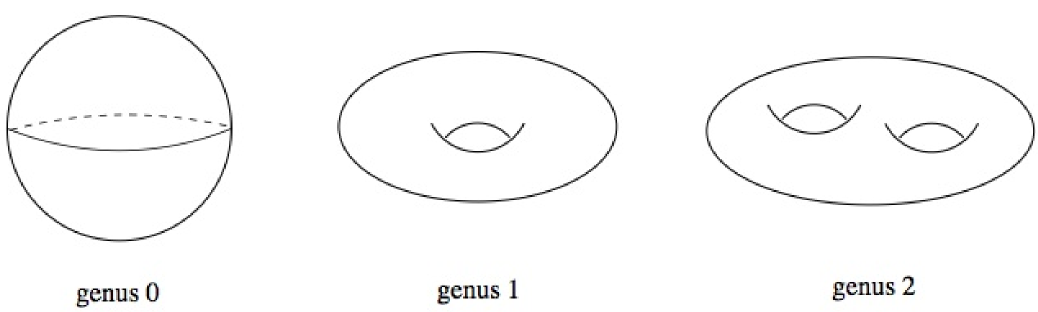


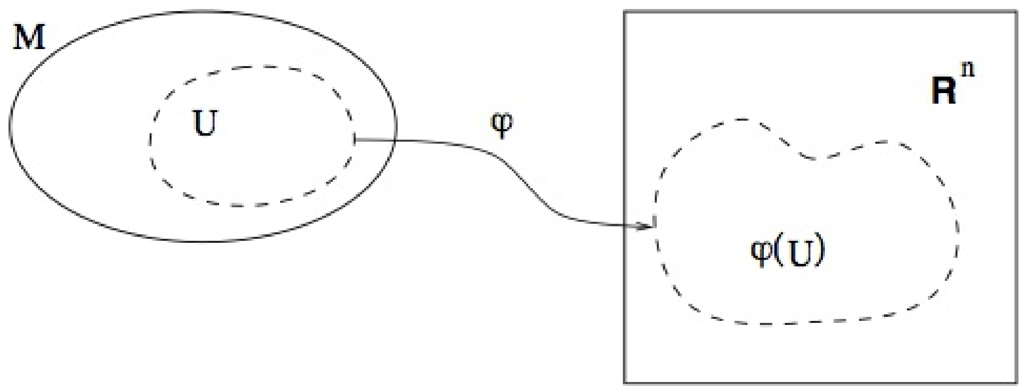
Figure 4.1: Some simple manifolds – the 2-sphere, the 2-torus and the double 2-torus.

Let us now give a more formal definition of a manifold.

Manifold A manifold is a topological space, M, together with an *atlas.*

An atlas is a set of charts (or coordinate systems), serving each patch of the manifold satisfying certain conditions. It is often written as $\{(U_\alpha, \varphi_\alpha)\}$, where the U_α denote the patches, and the φ_α denote the charts. The relevant conditions are the following:

1. M is the union of all the patches, U_α.
2. On a given patch, U_α, the chart, φ_α is a continuously differentiable one to one map from U_α onto a subset of $\mathbb{R}^n$
3. If two patches overlap, $U_\alpha \cap U_\beta \neq \emptyset$, then the map $\varphi_\alpha \circ \varphi_\beta^{-1}$ is a continuously differentiable function on (a subset of) $\mathbb{R}^n$.



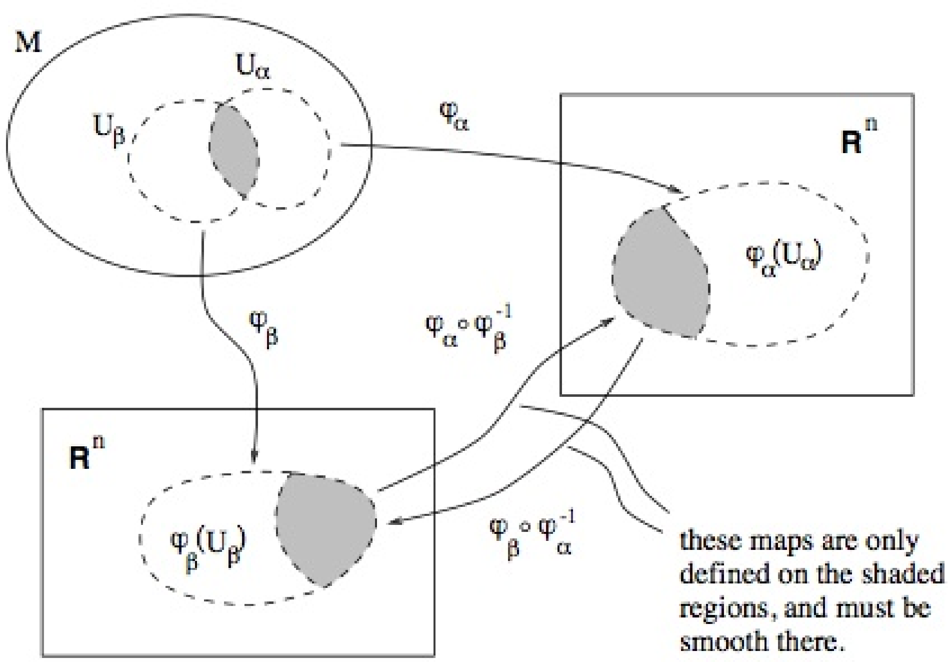


Fig. 3 contains a pictorial representations of this definition.

The surface of the earth is a topological space, whereas lines of longitude and latitude represent a chart. Together they form a manifold. To be honest, this is a lot of unnecessary hullabaloo! Basically, for the purpose of this course and plenty more beyond, all you need to know is that a manifold is a space along with a set of coordinates mapping out that space, and that it has some nice properties such as being smooth, and locally flat. But remember, in general relativity, we are interested in spacetime, rather than space, so we deal with pseudo-Riemannian manifolds rather than Riemannian manifolds. Let us rephrase our earlier sentence in this context: a pseudo-Riemannian manifold is a *spacetime* along with a set of coordinates mapping out that *spacetime*, and it has some nice properties such as being smooth, and locally *Minkowski*.

4.1 Vectors

Now let's move on to what is really important, starting with vectors. At school you learnt that a vector was something with magnitude *and* direction, such as velocity and force. Another example of a vector is the tangent to a curve at a given point. Given a set of coordinates x^i, such a curve is given by the path $x^i = X^i(\lambda)$, and the components of the tangent vector is given by $v^i = \frac{dX^i}{d\lambda}$. To picture this consider a football following some curved path from the boot to the goal. At any instance, the tangent vector to that path corresponds to the ball's velocity. This example demonstrates how the vector can depend on position, so what we should really be talking about are vector *fields*, with position dependent components, $v^i = v^i(x)$.

In special relativity you will have become familiar with the idea of a *four*-vector, with components, v^μ, $\mu = 0, 1, 2, 3$. The notion of a vector, can, of course, be generalised to any manifold (both Riemannian and Lorentzian) of any dimensionality. Again, we can think in terms of the tangent vectors to paths through the manifold. Given a set of coordinates x^μ, the tangent vector to the path $x^\mu = X^\mu(\lambda)$ has components $v^\mu = \frac{dX^\mu}{d\lambda}$. The vector itself should be written in terms of its components and a set of basis vectors adapted to the coordinate system,

$$v = v^\mu e_\mu = v^0 e_0 + v^1 \underline{e}_1 + v^2 \underline{e}_2 + v^3 \underline{e}_3 + \dots \tag{4.2}$$

We can think of each $\underline{e}_\mu$ as being the basis vector parallel to the x^μ axis. Intuitively, physics should not depend on our choice of coordinates, so what characterises a vector is the fact that the quantity v is independent of that choice. What does this mean, then, for the components of the vector? How do they transform under a change of coordinates?

Let us consider a change of coordinates from $x^\mu \to \hat{x}^\mu(x)$. As a result, our basis vectors change $\underline{e}_\mu \to \hat{\underline{e}}_\mu$, and so then do the components of a vector field, $v^\mu(x) \to \hat{v}^\mu(\hat{x})$. But remember, any vector can be thought of as a tangent vector to a path. This path is given by $x^\mu = X^\mu(\lambda)$ in one coordinate system, and $\hat{x}^\mu = \hat{X}^\mu(\lambda) = \hat{x}^\mu(X(\lambda))$ in the other. The components of the tangent vector in the transformed coordinate system are given by

$$\hat{v}^\mu(\hat{X}) = \frac{d\hat{X}^\mu}{d\lambda} = \frac{\partial \hat{x}^\mu}{\partial x^\nu}\Big|_{x=X(\lambda)} \frac{dX^\nu}{d\lambda} = \frac{\partial \hat{x}^\mu}{\partial x^\nu}\Big|_{x=X(\lambda)} v^\nu(X) \tag{4.3}$$

Since this equation holds for any choice of path, we can define the following transformation law for vectors in general:

$$\textit{Under } x^\mu \to \hat{x}^\mu(x)\textit{, the components of a vector } v \textit{ transform as } v^\mu(x) \to \hat{v}^\mu(\hat{x}) = \frac{\partial \hat{x}^\mu}{\partial x^\nu} v^\nu(x). \tag{4.4}$$

This transformation law essentially defines a vector. It also suggests a better way to think of the basis vectors $\underline{e}_\mu$. The point is that one can easily infer[1] the relation

$$\underline{e}_\mu = \frac{\partial \hat{x}^\nu}{\partial x^\mu} \hat{\underline{e}}_\mu \tag{4.5}$$

This should be compared with the Chain Rule, or the Liebniz rule, applied as follows on a function $F : M \to \mathbb{R}$

$$\frac{\partial}{\partial x^\mu} F(\hat{x}) = \frac{\partial \hat{x}^\nu}{\partial x^\mu} \frac{\partial}{\partial \hat{x}^\nu} F(\hat{x}) \tag{4.6}$$

This suggests that we should identify the basis vectors with the following differential operators,

$$\underline{e}_\mu = \frac{\partial}{\partial x^\mu}, \qquad \hat{\underline{e}}_\mu = \frac{\partial}{\partial \hat{x}^\mu} \tag{4.7}$$

In other words, we should really think of a vector as a differential operator. It corresponds to a directional derivative that acts on real valued functions on the manifold. Indeed, sometimes you will see people write

$$v \circ F = v^\mu \frac{\partial F}{\partial x^\mu}$$

If we identify v with the tangent to a curve, $v^\mu = \frac{dX^\mu}{d\lambda}$, then $\frac{dF}{d\lambda} = v \circ F$.

As we have emphasized already, a vector is a field that depends on where you are on the manifold. Conversely, at each point on the manifold, we can consider the space of all the vectors that can be defined there. The set of all such vectors at a point $p \in M$ is known as the *tangent space*, $T_p(M)$. We can think of a vector as a map from a point on the manifold to an element of the tangent space. The set of all tangent spaces on a given manifold is known as the *tangent bundle*, $T(M) = \{T_p(M), p \in M\}$. Imagine taking a freekick at Anfield – in principle you could kick the ball in any number of ways. The space of all possible initial velocities for the ball corresponds to the tangent space at the position of the freekick. The set of all such tangent spaces with the freekick being taken from any point on the pitch is the tangent bundle.

The tangent space $T_p(M)$ is a real vector space in the usual mathematical sense. In particular we have distributivity: if $v_1, v_2 \in T_p(M)$ and $a, b \in \mathbb{R}$, then $(a+b)(v_1+v_2) = av_1 + bv_1 + av_2 + bv_2$. Like any vector space we can define a set of basis vectors. The examples given by Eq. 4.1 are called *coordinate* bases, because they are adapted to a given coordinate system defined in a neighbourhood of the point, p. However, we are perfectly

[1] This can be seen from $v = v^\mu e_\mu = \hat{v}^\mu \hat{\underline{e}}_\mu = v^\nu \frac{\partial \hat{x}^\mu}{\partial x^\nu} \hat{\underline{e}}_\mu = v^\mu \frac{\partial \hat{x}^\nu}{\partial x^\mu} \hat{\underline{e}}_\nu$.

entitled to consider non-coordinate bases which cannot be written in the form of Eq. 4.1, eg $\{E_a, a = 0, 1, 2, 3 \ldots\}$ where

$$E_a = e_a{}^{\mu} \frac{\partial}{\partial x^{\mu}} \tag{4.8}$$

Here $e_a{}^{\mu}$ is known as the *tetrad*, or *vierbein* in four dimensions. If you speak ancient Greek or modern German you can probably generalise its name to higher dimensions. I don't so I won't!

For the coordinate basis, we have $e_a{}^{\mu} = \delta^{\mu}_a$. Another commonly used basis is an *orthonormal basis* defined such that

$$g_{\mu\nu} e_a{}^{\mu} e_a{}^{\nu} = \eta_{ab} \tag{4.9}$$

where $g_{\mu\nu} dx^{\mu} dx^{\nu}$ is the metric on the manifold, and η_{ab} is the Minkowski metric in the same number of dimensions.

We can also define the components of a vector, v^a with respect to a general non-coordinate basis $v = v^a E_a$. We can generalise our transformation law Eq. 4.4 to correspond to any change of basis.

Under a change of basis $E_a \to \hat{E}_a = \Lambda_a{}^b E_b$, the components of a vector v transform as $v^a \to \hat{v}^a = v^b (\Lambda^{-1})_b{}^a$

Note that

$$(\Lambda^{-1})_a{}^c \Lambda_c{}^b = \Lambda_a{}^c (\Lambda^{-1})_c{}^b = \delta^a_b \tag{4.10}$$

For transformations between coordinate bases we have $\Lambda_{\mu}{}^{\nu} = \frac{\partial x^{\nu}}{\partial \hat{x}^{\mu}}$ and $(\Lambda^{-1})_{\nu}{}^{\mu} = \frac{\partial \hat{x}^{\mu}}{\partial x^{\nu}}$, and recover Eq. 4.4. For transformations between orthonormal bases, Λ is a Lorentz transformation and so $(\Lambda^{-1})_a{}^b = (\Lambda^T)_a{}^b = \Lambda^b{}_a$. This follows from the fact we need to the preserve Eq. 4.1 before and after.

So, now we know what is *really* meant by a vector. To summarize,

- A vector is a differential operator that acts on real valued functions of the manifold.
- A vector is a field that depends on where you are on the manifold.
- A vector is a map from a point on the manifold to an element of a vector space, known as the tangent space.
- A coordinate basis in the tangent space is one that is adapted to a particular coordinate system on the manifold at that point, satisfying Eq. 4.1
- The components of the vector transform according to a special law given by Eq. 4.4 when we change from one coordinate basis to another.
- It is possible (and sometimes convenient) to use non-coordinate bases instead.

4.2 Dual vectors (or 1-forms)

Everyone needs a sidekick. Who is Morecambe without Wise? Who is Mork without Mindy? Vectors also have a sidekick – they are called 1-forms. So what on earth are they? Well, just as vectors are often associated with tangent vectors to curves in the manifold, so 1-forms are sometimes identified with the normals to surfaces in the manifold. Not all 1 forms can be identified as such, but some can and it will provide us with a useful intuitive picture.

Let us be a little more precise. If the manifold has dimension D, a surface of dimension $D-1$ can defined by an equation $F(x) = constant$, where $F : M \to \mathbb{R}$ is some real map on the manifold, as shown in Fig. 4.2. Given a set of coordinates, x^μ, the normal, n_μ, to such a surface can be identified with the gradient $n_\mu = \frac{\partial F}{\partial x^\mu}$.

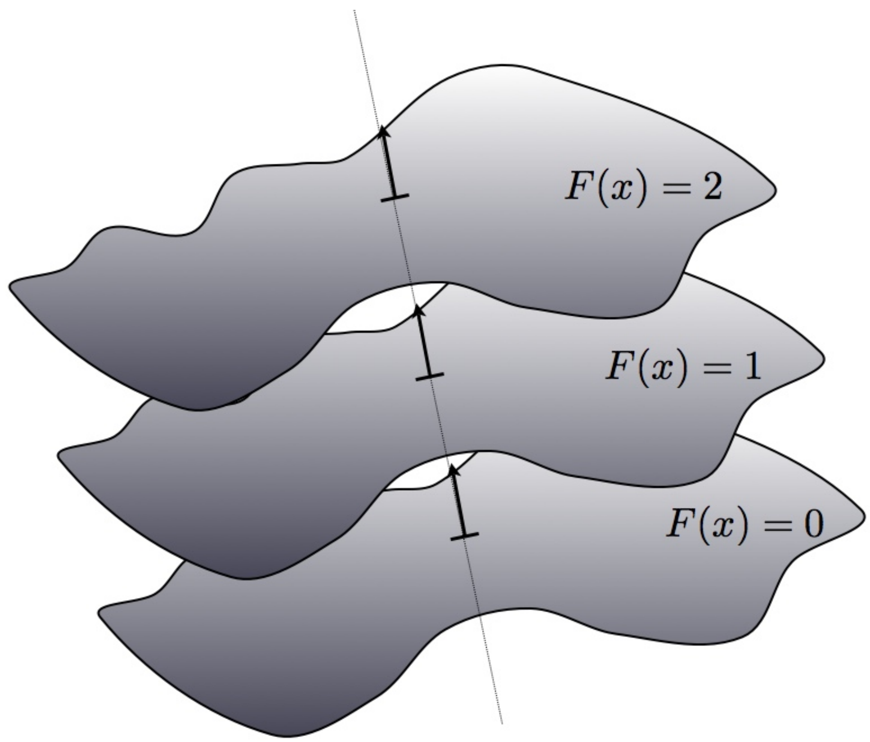


Figure 4.2: Surfaces of constant $F(x)$ and their normals.

To see why this is so, all we need to show is that for any tangent vector, v, lying in the surface, we have $v^\mu \frac{\partial F}{\partial x^\mu} = 0$. To this end, consider a curve $X^\mu(\lambda)$ lying in the surface. It follows that $F(X(\lambda)) = constant$, and so $\frac{dF}{d\lambda} = 0$. But we have already seen by simple application of the Chain rule that $\frac{dF}{d\lambda} = v^\mu \frac{\partial F}{\partial x^\mu}$, where $v^\mu = \frac{dX^\mu}{d\lambda}$, and we arrive at the expected result $v^\mu \frac{\partial F}{\partial x^\mu} = 0$. So indeed, $n_\mu = \frac{\partial F}{\partial x^\mu}$ is the normal to our surface of constant F.

One thing you will notice is that the components of the normal, n_μ have been written with indices down, in contrast to the vector, v^μ which has indices up. This will be reminiscent of contravariant and covariant vectors introduced in your *Gravity* course. Indeed, convention states that the components of 1-forms are written with indices down, and are identified with covariant vectors in a coordinate basis. In any event, as with vectors, a general 1-form,

w, can be written in terms of its components and a set of basis 1-forms adapted to the coordinate system,

$$w = w_\mu \theta^\mu = w_0\theta^0 + w_1\theta^1 + w_2\theta^2 + w_3\theta^3 + \dots \tag{4.11}$$

Here we think of each θ^μ as being the basis 1-form in the x^μ direction. But what are these θ^μ? To find out, let us return to the special case[2] where the 1-form is a normal, n_μ, and consider how F varies across the manifold,

$$\delta F = F(x+\delta x) - F(x) = \frac{\partial F}{\partial x^\mu}\delta x^\mu + \mathcal{O}(\delta x^2)$$

To leading order, the infinitesimal version of the fluctuation δF is sometimes written as

$$dF = \frac{\partial F}{\partial x^\mu} dx^\mu \tag{4.12}$$

This defines a 1-form, dF, using the gradient operator, d, which you may be familiar with from your differential geometry course. It has components $n_\mu = \frac{\partial F}{\partial x^\mu}$ and we read off the standard notation for the basis 1-forms $\theta^\mu = dx^\mu$. Clearly these are adapted to a particular coordinate system, so these correspond to our coordinate basis.

One is obviously tempted to ask what happens when we change coordinates from $x^\mu \to \hat{x}^\mu(x)$. As before we require that the general 1-form w be independent of the choice of coordinates. Under this transformation, the basis 1-forms change from $\theta^\mu \to \hat{\theta}^\mu$, where

$$\theta^\mu = dx^\mu, \qquad \hat{\theta}^\mu = d\hat{x}^\mu \tag{4.13}$$

and the components of the 1-form field transform as $w_\mu(x) \to \hat{w}_\mu(\hat{x})$. Now, using Eq. 4.12, it is easy to see that $d\hat{x}^\mu = \frac{\partial \hat{x}^\mu}{\partial x^\nu} dx^\nu$. We now infer the following transformation law for 1-forms

Under $x^\mu \to \hat{x}^\mu(x)$, the components of a 1-form w transform as $w_\mu(x) \to \hat{w}_\mu(\hat{x}) = \frac{\partial x^\nu}{\partial \hat{x}^\mu} w_\nu(x)$. (4.14)

This transformation law essentially defines what is meant by a 1-form.

In the previous section, we learnt that vectors live in the tangent space, so where do the 1-forms live? They live in something called the *cotangent* space. To appreciate what this is, we need to introduce the notion of a dual vector space. In general, every vector space, V, has a dual, V^*. This is the defined as the space of linear maps from the vector space to the real line. So, if $w \in V^*$, then $w : V \to \mathbb{R} : v \to w(v)$ is a linear map in the usual sense. Often $w(v)$ is written as an inner product $\langle w|v\rangle$.

Obviously our interest lies in the dual to the tangent space, $T_p(M)$, at each point p in the spacetime manifold, M. This is what we mean by the *cotangent* space, and it is written $T^*_p(M)$. Unsurprisingly, the set of all cotangent spaces is called the cotangent bundle, $T^*(M) = \{T^*_p(M), p \in M\}$.

[2]It is important to realise that not all 1-forms can be identified with normals.

The elements of the cotangent space, $w \in T_p^*(M)$ are what we have been calling 1-forms. They are also known as *covectors*. The basis 1-forms given by Eq. 4.13 correspond to a *coordinate* basis on the cotangent space, as they are adapted to a particular coordinate system. However, we can certainly choose a more general basis, Θ^a, so that a general 1-form can be written as $w = w_a \Theta^a$. As with vectors we can use vierbeins to relate a general basis to a coordinate basis on the cotangent space. In general we have that

$$\Theta^a = e^a{}_\mu dx^\mu \tag{4.15}$$

where the inverse vierbein $e^a{}_\mu$ is related to $e_a{}^\mu$ and satisfies the relations

$$e^a{}_\mu e_b{}^\mu = \delta^a_b, \qquad e_a{}^\mu e^a{}_\nu = \delta^\mu_\nu$$

For the coordinate basis, we have $e^a{}_\mu = \delta^a_\mu$. For the orthonormal basis we have

$$g^{\mu\nu} e^a{}_\mu e^b{}_\nu = \eta^{ab} \tag{4.16}$$

where $g^{\mu\nu}$ is the inverse metric on the manifold, and η^{ab} is the inverse Minkowski metric in the same number of dimensions.

Just like Mork and Mindy, basis vectors, E_a and basis 1-forms, Θ^a come in pairs, where one basis is described as the dual of other. This is defined as being the case whenever

$$\langle \Theta^a | E_b \rangle = \delta^a_b \tag{4.17}$$

It immediately follows that $\langle w|u \rangle = w_a v^b \langle \Theta^a | E_b \rangle = w_a v^a$. Note that the coordinate basis vectors and the coordinate basis 1-forms are, of course dual bases,

$$\left\langle dx^\mu \middle| \frac{\partial}{\partial x^\nu} \right\rangle = \delta^\mu_\nu$$

We are now ready to generalise our transformation law Eq. 4.14.

Under a change of basis $E_a \to \hat{E}_a = \Lambda_a{}^b E_b$, the dual basis transforms as $\Theta^a \to \hat{\Theta}^a = \Theta^b (\Lambda^{-1})_b{}^a$, and the components of a 1-form w transform as $w_a \to \hat{w}_a = \Lambda_a{}^b w_b$

Again, for transformations between coordinate bases we have $\Lambda_\mu{}^\nu = \frac{\partial x^\nu}{\partial \hat{x}^\mu}$ and $(\Lambda^{-1})_\nu{}^\mu = \frac{\partial \hat{x}^\mu}{\partial x^\nu}$, and so we recover 4.14. For an orthonormal basis we identify Λ with a Lorentz transformation.

In summary then,

- The cotangent space, $T_p^*(M)$ is the dual to the tangent space, $T_p(M)$ at each point on the manifold.
- Elements of the cotangent space are known as 1-forms or covectors.
- A coordinate basis in the cotangent space is one that is adapted to a particular coordinate system on the manifold at that point. Given a set of coordinates, x^μ, the coordinate basis vectors are dx^μ

- The components of the 1-form transform according to a special law given by Eq. 4.14 when we change from one coordinate basis to another.
- It is possible (and sometimes convenient) to use non-coordinate bases instead.
- In a D dimensional manifold, the normal to a surface of dimension $D-1$ is a 1-form, dF. Not all one forms can be written like this though.

4.3 Dual dual vectors (or just vectors!)

Mork has a sidekick called Mindy. But who is Mindy's sidekick? Well its Mork of course. So what about 1-forms? What is the dual of a 1-form? If we follow the Mork and Mindy logic, then presumably the answer is .... a vector. Let's see why this is correct.

To identify the dual of a 1-form, we need to ask, what is the space dual to the cotangent space? By definition, this must be the space of linear maps from the cotangent space to the real line. All we need to show is that a vector is such a map. Consider $v \in T_p(M)$ and $w \in T_p^*(M)$, then we know from the previous section that $w(v) = \langle w|v\rangle \in \mathbb{R}$. So the map which takes $w \to \langle w|v\rangle$ is certainly a map from $T_p^*(M) \to \mathbb{R}$, and since $\langle aw_1 + bw_2|v\rangle = a\langle w_1|v\rangle + b\langle w_2|v\rangle$ we know it to be linear. This linear map can obviously be identified with the vector v.

4.4 Tensors

Both vectors and 1-forms can be thought of as one dimensional arrays. Some people distinguish between the two by writing vectors as a vertical array

$$v = \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix}$$

and 1-forms as a horizontal array

$$w = (w_0, w_1, w_2, w_3)$$

Having made this identification, we might ask whether there is any geometrical quantity that can be associated with higher dimensional arrays, such as matrices

$$\begin{pmatrix} T^0{}_0 & \dots & T^0{}_3 \\ \vdots & & \vdots \\ T^3{}_0 & \dots & T^3{}_3 \end{pmatrix}$$

Well, we know that higher dimensional arrays can be obtained by taking products of lower dimensional ones. For example, the outer matrix product of two one dimensional arrays is a 2 dimensional array.

$$v \otimes w = \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix} (w_0, w_1, w_2, w_3) = \begin{pmatrix} v^0 w_0 & \dots & v^0 w_3 \\ \vdots & & \vdots \\ v^3 w_0 & \dots & v^3 w_3 \end{pmatrix}$$

Let us push this idea a little further, and imagine taking the outer products of say, k vectors, $v_1, \dots, v_k$, and l 1-forms, $w^1, \dots, w^l$. The resulting object is known as a (k, l) *tensor field*

$$T = v_1 \otimes \dots \otimes v_k \otimes w^1 \otimes \dots \otimes w^l$$

Given a set of coordinates, x^μ, we may write each vector and 1-form in terms of the coordinate vector and 1 form bases respectively, giving

$$\begin{aligned} T &= \left(v_1^{\mu_1} \frac{\partial}{\partial x^{\mu_1}} \right) \otimes \dots \otimes \left(v_k^{\mu_k} \frac{\partial}{\partial x^{\mu_k}} \right) \otimes (w^1_{\nu_1} dx^{\nu_1}) \otimes \dots \otimes (w^l_{\nu_l} dx^{\nu_l}) && (4.18) \\ &= v_1^{\mu_1} \dots v_k^{\mu_k} w^1_{\nu_1} \dots w^l_{\nu_l} \frac{\partial}{\partial x^{\mu_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{\mu_k}} \otimes dx^{\nu_1} \otimes \dots \otimes dx^{\nu_l} && (4.19) \end{aligned}$$

Here we identify the components of the tensor with k indices up and l indices down,

$$T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} = v_1^{\mu_1} \dots v_k^{\mu_k} w^1_{\nu_1} \dots w^l_{\nu_l}$$

and the coordinate tensor basis,

$$\frac{\partial}{\partial x^{\mu_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{\mu_k}} \otimes dx^{\nu_1} \otimes \dots \otimes dx^{\nu_l} \tag{4.20}$$

We can use this particular example to infer the following transformation law for tensors [3]:

Under $x^\mu \to \hat{x}^\mu(x)$, the components of a (k, l) tensor T transform as

$$T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l}(x) \to \hat{T}^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l}(\hat{x}) = \frac{\partial \hat{x}^{\mu_1}}{\partial x^{\bar{\mu}_1}} \dots \frac{\partial \hat{x}^{\mu_k}}{\partial x^{\bar{\mu}_k}} . \frac{\partial x^{\bar{\nu}_1}}{\partial \hat{x}^{\nu_1}} \dots \frac{\partial x^{\bar{\nu}_l}}{\partial \hat{x}^{\nu_l}} T^{\bar{\mu}_1 \dots \bar{\mu}_k}{}_{\bar{\nu}_1 \dots \bar{\nu}_l}(x) \tag{4.21}$$

Of course, the notion of a (k, l) tensor goes beyond the direct products we have been discussing. It applies to *anything* of the form

$$T = T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} \frac{\partial}{\partial x^{\mu_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{\mu_k}} \otimes dx^{\nu_1} \otimes \dots \otimes dx^{\nu_l} \tag{4.22}$$

provided the components satisfy the transformation law 4.21.

We have already come across an example of a tensor. You will sometimes see the metric written as

$$ds^2 = g = g_{\mu\nu} dx^\mu \otimes dx^\nu$$

[3] Apply (4.4) and (4.14) applied to each individual ingredient in this example

Since the metric transforms as

$$\hat{g}_{\mu\nu}(\hat{x}) = \frac{\partial x^\alpha}{\partial \hat{x}^\mu}\frac{\partial x^\beta}{\partial \hat{x}^\nu} g_{\alpha\beta}(x)$$

it clearly corresponds to a $(0,2)$ tensor. Similarly, the inverse metric

$$g^{-1} = g^{\mu\nu}\frac{\partial}{\partial x^\mu} \otimes \frac{\partial}{\partial x^\nu}$$

corresponds to a $(2,0)$ tensor.

Where do tensors live? Well, at each point $p \in M$, we can define the space of all (k,l) tensor fields evaluated at that point. This corresponds to the product space built out of the tangent and cotangent spaces,

$$T_p^{(k,l)}(M) = \overbrace{T_p(M) \times \ldots \times T_p(M)}^{k \text{ copies}} \times \overbrace{T_p^*(M) \times \ldots \times T_p^*(M)}^{l \text{ copies}}$$

In analogy with our discussion of dual vectors, and indeed, dual dual vectors, this suggests that in actual fact, a (k,l) tensor is a really multi-linear map

$$T : T_p^{(l,k)}(M) \to \mathbb{R} : (v_1, \ldots, v_l, w^1, \ldots, w^k) \to T(w^1, \ldots, w^k, v_1, \ldots, v_l)$$

The ordering of k and l here is not a typo. To see why, note that in component language this map does the following,

$$(v_1^{\nu_1}, \ldots, v_l^{\nu_l}, w^1_{\mu_1}, \ldots, w^k_{\mu_k}) \to T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} w^1_{\mu_1} \ldots w^k_{\mu_k} v_1^{\nu_1} \ldots v_l^{\nu_l}$$

Now Eq. 4.4 corresponds to the coordinate basis in $T_p^{(k,l)}(M)$. As usual, we can also consider non-coordinate bases, built out of the non-coordinate bases in $T_p(M)$ and $T_p^*(M)$,

$$E_{a_1} \otimes \ldots \otimes E_{a_k} \otimes \Theta^{b_1} \otimes \ldots \otimes \Theta^{b_l} \tag{4.23}$$

The tensors are now written as

$$T = T^{a_1\ldots a_k}{}_{b_1\ldots b_l} E_{a_1} \otimes \ldots \otimes E_{a_k} \otimes \Theta^{b_1} \otimes \ldots \otimes \Theta^{b_l} \tag{4.24}$$

From Eqs. 4.8 and 4.15, we see that one can easily switch between components in a non-coordinate and coordinate basis using the vierbeins,

$$T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} = T^{a_1\ldots a_k}{}_{b_1\ldots b_l} e_{a_1}{}^{\mu_1} \ldots e_{a_k}{}^{\mu_k} e^{b_1}{}_{\nu_1} \ldots e^{b_l}{}_{\nu_l} \tag{4.25}$$

$$T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} e^{a_1}{}_{\mu_1} \ldots e^{a_k}{}_{\mu_k} e_{b_1}{}^{\nu_1} \ldots e_{b_l}{}^{\nu_l} = T^{a_1\ldots a_k}{}_{b_1\ldots b_l} \tag{4.26}$$

You should compare these expressions with Eqs. 4.1 and 4.2

Let us end this discussion with a comment on why tensors are so cool. The point is that if a tensor equation holds in one coordinate system, it holds in all coordinate systems, because of the tensor transformation law. This can be computationally very powerful. For example, suppose you wanted to prove a tensor equation,

$$T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} = S^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}$$

This may be tricky to prove in a generic coordinate system. What you can do, however, is pick a convenient coordinate system in which the above relation is easy to prove. Because the relation is a tensor relation, you can extrapolate this proof to apply in any coordinate system! A particularly useful coordinate system is known as *Riemann normal coordinates* about a particular point. These exploit the fact that any pseudo Riemannian manifold is locally Minkowski. With Riemann normal coordinates, the metric looks Minkowski, up to second order,

$$g_{\mu\nu} = \eta_{\mu\nu} + \mathcal{O}(x_\mu x_\nu) \tag{4.27}$$

In summary then

- tensors can be identified with higher dimensional arrays
- a (k, l) tensor has k indices up and l indices down.
- a (k, l) tensor lives in the product of k tangent spaces and l cotangent spaces
- a (k, l) tensor is a multi-linear map acting on l vectors and k 1 forms
- the components of a tensor transform according to a special law given by Eq. 4.21 when we change from one coordinate basis to another.
- it is possible (and sometimes convenient) to use non-coordinate bases instead.
- tensor equations can be proven in one particular coordinate system (eg Riemann normal), but the proof automatically applies to any coordinate system.

Exercises

1. A two dimensional Euclidean vector, v, has components $\begin{pmatrix} v^x \\ v^y \end{pmatrix}$ in a Cartesian coordinate system. Show that in polar coordinates it has components

$$v^r = 2r(\cos\theta v^x + \sin\theta v^y), \qquad v^\theta = \frac{2}{r}(-\sin\theta v^x + \cos\theta v^y)$$

2. Prove that the contracted product of a vector and a 1 form, $v^\mu w_\mu$ transforms as a scalar.
3. Write down the components of the vierbein relating the orthonormal basis to the coordinate basis for the Schwarzschild metric,

$$ds^2 = -V(r)dt^2 + \frac{dr^2}{V(r)} + r^2(d\theta^2 + \sin^2\theta d\phi^2)$$

Hint: recall that for an orthonormal basis we have $ds^2 = \eta_{ab}\Theta^a \otimes \Theta^b$, *where* $\Theta^a = e^a{}_\mu dx^\mu$.

5 Tensor calculus

As we have seen, given a manifold, we can define vectors, 1-forms and tensors at each point on that manifold. If we want to know how things change from point to point, if we want to do any sort of calculus, we need to be able to compare tensors at different points. For example, suppose you want to track the changes in velocity of a ball rolling around a football pitch. At Wembley stadium, with its flat pristine pitches, all points on the pitch are more or less equivalent and one can easily compare velocities at different points. In contrast, on a bumpy pitch in Beeston things are much more complicated. How does one track the changes in velocity of a ball rolling over a series of bobbles?

Lets work with vectors for the moment, just to be definite. Given a general manifold M and a coordinate system x^μ, what is the gradient of the vector field with components, $v^\mu(x)$? Diving in head first, we make the obvious guess, and propose that it is given by a quantity with components

$$T^\mu_\nu(x) = \frac{\partial v^\mu(x)}{\partial x^\nu} \tag{5.1}$$

However, we immediately see that this cannot be right because T^μ_ν does not transform as a tensor (see Exercise 1). What goes wrong? Well to calculate the partial derivative we are required to compare the vector field at neightbouring points p and $\hat{p}$, by computing

$$\lim_{\epsilon \to 0} \frac{v^\mu|_p - v^\mu|_{\hat{p}}}{\epsilon} \tag{5.2}$$

for some appropriate parameter ϵ. The trouble is that the quantity $v^\mu|_p - v^\mu|_{\hat{p}}$ does not transform as a vector because the transformation laws 4.4 are defined independently for each point on the manifold. What we need is a way of mapping the vector at $\hat{p}$ to the point p. We will comment on two methods: *Lie dragging* and *parallel transport*. The former gives rise to the *Lie derivative* while the latter gives rise to the *covariant derivative*.

5.1 The Lie derivative

We will now outline the method of Lie dragging and describe the Lie derivative. We want to drag the vector field $v^\mu(x)$ along the integral curves of another vector field $u^\mu(x)$[4]. The result is a new vector field, $\hat{v}^\mu(x)$, that we then compare with $v^\mu(x)$. This will give us the Lie derivative of v with respect to u .

We begin with the integral curve, $x^\mu = X^\mu(\lambda)$, passing through the point p, and suppose that we have $\lambda = \lambda_p$ at p and $\lambda = \lambda_p + \delta\lambda$ at a nearby point $\hat{p}$. If we define the coordinate transformation

$$x^\mu \to \hat{x}^\mu = x^\mu + \delta\lambda u^\mu(x) \tag{5.3}$$

then it is clear that $x^\mu|_{\hat{p}} = \hat{x}^\mu|_p + \mathcal{O}(\delta\lambda^2)$. To see this, we simply note that

$$\hat{x}^\mu|_p - x^\mu|_{\hat{p}} = x^\mu|_p + \delta\lambda u^\mu(x_p) - x^\mu|_{\hat{p}} = X^\mu(\lambda_p) + \delta\lambda \frac{dX^\mu}{d\lambda}\bigg|_{\lambda_p} - X^\mu(\lambda_p + \delta\lambda) = \mathcal{O}(\delta\lambda^2) \tag{5.4}$$

[4]The integral curves of u^μ are those curves that are tangent to $u^\mu(x)$. There is only one such curve passing through each point on the manifold

In other words, the two points, p and $\hat{p}$, are mapped to the same point in $\mathbb{R}^4$ by the two different coordinate systems. Having identified the points via the coordinate transformation, we then compute the effect of transformation on the vector to extract the Lie derivative.

In accordance with the transformation law 4.4,

$$v^\mu(x) \to \hat{v}^\mu(\hat{x}) \equiv \hat{v}^\mu(x+\delta\lambda u) = v^\mu(x) + \delta\lambda \frac{\partial u^\mu}{\partial x^\nu} v^\nu(x) + \mathcal{O}(\delta\lambda^2) \tag{5.5}$$

This implicitly defines a new vector field $\hat{v}^\mu(x)$. At the point p it represents the original vector having been *Lie dragged* from the point $\hat{p}$. Using $\hat{x}^\mu = x^\mu + \delta\lambda u^\mu(x)$ we now see that

$$\hat{v}^\mu(x) = v^\mu(x) + \delta\lambda \left(\frac{\partial u^\mu}{\partial x^\nu} v^\nu(x) - \frac{\partial v^\mu}{\partial x^\nu} u^\nu(x) \right) + \mathcal{O}(\delta\lambda^2) \tag{5.6}$$

We are now in a position to define the Lie derivative as

$$\mathcal{L}_u v^\mu = \lim_{\delta\lambda\to 0} \frac{v^\mu(x) - \hat{v}^\mu(x)}{\delta\lambda} = u^\nu(x)\frac{\partial v^\mu}{\partial x^\nu} - v^\nu(x)\frac{\partial u^\mu}{\partial x^\nu} \tag{5.7}$$

One can explicitly check that this transforms as a vector. The Lie derivatives of a general (k,l) tensor, $T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}$ can be computed using the same logic

$$\begin{aligned} \mathcal{L}_u T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} &= \lim_{\delta\lambda\to 0} \frac{T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}(x) - \hat{T}^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}(x)}{\delta\lambda} \\ &= u^\alpha \partial_\alpha T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} - \sum_{i=1}^k T^{\ldots\mu_{i-1}\alpha\mu_{i+1}\ldots}{}_{\nu_1\ldots\nu_l}\partial_\alpha u^{\mu_i} + \sum_{i=1}^l T^{\mu_1\ldots\mu_k}{}_{\ldots\nu_{i-1}\alpha\nu_{i+1}\ldots}\partial_{\nu_i} u^\alpha \end{aligned} \tag{5.8}$$

Again, one can explicitly check that this transforms as a tensor. Note that the derivation and definition of the Lie derivative makes no reference to the metric.

Lie derivatives tell you how a tensor changes along an integral curve, so when they happen to vanish everywhere, this is usually associated with symmetry. A particularly important example of this is when the Lie derivative of the metric vanishes with respect to some vector, ξ^μ,

$$\mathcal{L}_\xi g_{\mu\nu} = 0 \tag{5.9}$$

This is known as *Killing's equation*. When it holds, ξ^μ is known as the *Killing vector*, and the metric is understood to remain unchanged along its integral curves. A spacetime with a timelike Killing vector is known as a stationary spacetime.

Lie derivatives play an important role in perturbation theory, particularly when applying infinitesimal coordinate transformations to tensor fluctuations. Indeed, suppose we have a background (k,l) tensor $\bar{T}^{\cdots}{}_{\ldots}$, and we consider some fluctuation $T^{\cdots}{}_{\ldots} = \bar{T}^{\cdots}{}_{\ldots} + \delta T^{\cdots}{}_{\ldots}$. Under an infinitesimal coordinate transformation, $x^\mu \to x^\mu - \xi^\mu(x)$, we find that to leading order the tensor fluctuations transform as

$$\delta T^{\cdots}{}_{\ldots} \to \delta T^{\cdots}{}_{\ldots} + \mathcal{L}_\xi \bar{T}^{\cdots}{}_{\ldots} \tag{5.10}$$

In particular, this means that if $\bar{T}^{\cdots}{}_{\cdots}$ vanishes on the background, $\delta T^{\cdots}{}_{\cdots}$ is invariant under infinitesimal coordinate transformations (we typically say it is *gauge invariant*)

Finally we note that the Lie derivative obeys all the properties usually associated with a derivative operator, such as being Liebnizian (see Exercise 5). The Lie derivatives of vectors are often written as a *Lie bracket*

$$\mathcal{L}_u v = [u, v] \tag{5.11}$$

defined by the following action on a function $f : M \to \mathbb{R}$,

$$[u, v](f) = u(v(f)) - v(u(f)) \tag{5.12}$$

5.2 The covariant derivative and the connection

The covariant derivative is a different beast to the Lie derivative. This time we map the vector at $\hat{p}$ to the point p using a procedure known as parallel transport. Parallel transport involves connecting the two points along a suitable curve, and transporting the vector along that curve so it always stays instantaneously parallel (in some intuitive sense, see figure 5.1).

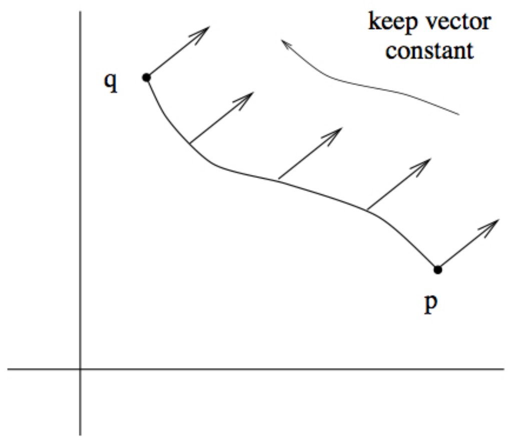


Figure 5.1: For parallel transport in flat space, the components of the vector are kept constant.

However, we will not derive the covariant derivative using parallel transport, but take another approach which uses techniques familiar to those who have ever gauged a global symmetry. Recall that the Lagrangian for a free complex scalar field

$$\mathcal{L} = -\partial_\mu \phi \partial^\mu \phi^*$$

is invariant under a global $U(1)$ symmetry, $\phi \to \phi e^{i\alpha}$, where α is constant. However, to promote this to a local $U(1)$ in which $\alpha = \alpha(x)$, we must introduce a connection A_μ, transforming as $A_\mu \to A_\mu - \frac{1}{g}\partial_\mu \alpha$, and the covariant derivative, D_μ, where $D_\mu \phi = (\partial_\mu +$

$igA_\mu)\phi$. The covariant derivative of ϕ has "nice" transformation properties, $D_\mu\phi \to e^{i\alpha}D_\mu\phi$ and the modified Lagrangian

$$\mathcal{L} = -D_\mu\phi(D^\mu\phi)^*$$

is invariant under the *local* $U(1)$.

Let's return to our discussion of geometry and the derivative of a vector, $v^\mu(x)$. We saw how the partial derivative $\frac{\partial v^\mu}{\partial x^\nu}$ did not have "nice" transformation properties, where "nice" here means "transforming like a tensor". So in analogy with the gauging process we just described for the $U(1)$ symmetry, we introduce a connection that transforms appropriately, and a covariant derivative that has the nice transformation properties we desire. To this end, let us define the connection, $\Gamma^\mu{}_{\alpha\nu}$, and the covariant derivative, ∇_μ, where

$$\nabla_\nu v^\mu = \partial_\nu v^\mu + \Gamma^\mu{}_{\alpha\nu}v^\alpha \tag{5.13}$$

Under a coordinate transformation, $x^\mu \to \hat{x}^\mu$, the connection transforms as

$$\Gamma^\mu{}_{\alpha\nu}(x) \to \hat{\Gamma}^\mu{}_{\alpha\nu}(\hat{x}) = \frac{\partial \hat{x}^\mu}{\partial x^\rho}\frac{\partial x^\sigma}{\partial \hat{x}^\alpha}\frac{\partial x^\lambda}{\partial \hat{x}^\nu}\Gamma^\rho{}_{\sigma\lambda}(x) - \frac{\partial^2 \hat{x}^\mu}{\partial x^\sigma \partial x^\lambda}\frac{\partial x^\sigma}{\partial \hat{x}^\nu}\frac{\partial x^\lambda}{\partial \hat{x}^\alpha} \tag{5.14}$$

It now follows that the covariant derivative of v^μ transforms as a tensor, as desired

$$\nabla_\nu v^\mu \to \frac{\partial x^\alpha}{\partial \hat{x}^\nu}\frac{\partial \hat{x}^\mu}{\partial x^\beta}\nabla_\alpha v^\beta \tag{5.15}$$

By applying similar methods we define the covariant derivative of a general (k, l) tensor as

$$\nabla_\beta T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} = \partial_\beta T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} + \sum_{i=1}^{k}\Gamma^{\mu_i}{}_{\alpha\beta}T^{\cdots\mu_{i-1}\alpha\mu_{i+1}\cdots}{}_{\nu_1\ldots\nu_l} - \sum_{i=1}^{l}\Gamma^\alpha{}_{\nu_i\beta}T^{\mu_1\ldots\mu_k}{}_{\cdots\nu_{i-1}\alpha\nu_{i+1}\cdots} \tag{5.16}$$

which transforms as a $(k, l+1)$ tensor.

We now pause for some notation. Just as the partial derivative is often written using commas, $\partial_\mu T^{\cdots}_{\cdots} = T^{\cdots}_{\cdots,\mu}$, so the covariant derivative is often written using semi-colons, $\nabla_\mu T^{\cdots}_{\cdots} = T^{\cdots}_{\cdots;\mu}$. Also, the covariant derivative of a tensor along a given direction, u^μ, defined by $u^\mu\nabla_\mu T^{\cdots}_{\cdots}$ is sometimes written as $\nabla_u T^{\cdots}_{\cdots}$ for short.

We should also point out a thing or two about the connection. The first is that one can always choose a coordinate system or which the connection vanishes locally. This corresponds to choosing Riemann normal coordinates, touched on in section 4.4. Calculationally it is a very useful tool, and you are invited to prove this result in Exercise 9 below.

The second thing we would like to point out that the difference of any two connections is actually a tensor. This is easily seen by glancing at the transformation law 5.14, and follows from the fact that the term by which the connection fails to be a tensor is independent of the connection itself. The best known example of such a tensor is the torsion tensor, defined by antisymmetrizing over the lower indices of the connection,

$$\mathcal{T}^\alpha{}_{\mu\nu} = \Gamma^\alpha{}_{\mu\nu} - \Gamma^\alpha{}_{\nu\mu}$$

Torsion measures the amount by which covariants derivatives do not commute when acting on a scalar

$$[\nabla_\mu, \nabla_\nu] f = -\mathcal{T}^\alpha{}_{\mu\nu} \partial_\alpha f$$

For the rest of this course will assume that torsion vanishes, unless otherwise stated, and that we have a symmetric connection,

$$\Gamma^\alpha{}_{\mu\nu} = \Gamma^\alpha{}_{\nu\mu}$$

The most important connection is the *Levi Civita connection*. This is the unique connection for which the metric is covariantly constant, in a torsion free theory,

$$\nabla_\alpha g_{\mu\nu} = 0$$

One can straightforwardly show that the Levi-Civita connection is given by

$$\Gamma^\alpha{}_{\mu\nu} = \frac{1}{2} g^{\alpha\beta} \left(g_{\beta\mu,\nu} + g_{\beta_\nu,\mu} - g_{\mu\nu,\beta} \right) \tag{5.17}$$

The Levi-Civita connection is sometimes called the metric connection, for obvious reasons.

We are now ready to return to the notion of parallel transport. A tensor is parallely transported along a curve $x^\mu = X^\mu(\lambda)$ if its absolute derivative vanishes everywhere on the curve

$$\frac{D}{D\lambda} T^{\cdots}_{\cdots} = 0$$

The absolute derivative is defined as the covariant derivative along the direction of the curve,

$$\frac{D}{D\lambda} T^{\cdots}_{\cdots} = \nabla_u T^{\cdots}_{\cdots}$$

where $u^\mu = \frac{dX^\mu}{d\lambda}$ is the tangent vector along the curve.

An affinely parametrised *geodesic* is an example of a curve for which the tangent vector is transported parallel to itself, $\nabla_u u^\mu = 0$ (see section 6.1).

Exercises

1. Suppose we change coordinates, $x^\mu \to \hat{x}^\mu(x)$. Show that if $v^\mu(x)$ transforms as a vector that

$$T^\mu_\nu(x) = \frac{\partial v^\mu(x)}{\partial x^\nu} \tag{5.18}$$

   transforms as

$$T^\mu_\nu(x) \to \hat{T}^\mu_\nu(\hat{x}) \equiv \frac{\partial \hat{v}^\mu(\hat{x})}{\partial \hat{x}^\nu} = \frac{\partial \hat{x}^\mu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial \hat{x}^\nu} T^\alpha_\beta(x) + \frac{\partial^2 \hat{x}^\mu}{\partial x^\alpha \partial x^\beta} \frac{\partial x^\beta}{\partial \hat{x}^\nu} v^\alpha(x) \tag{5.19}$$

2. Verify that $\mathcal{L}_u v^\mu$ defined by equation 5.7 transforms as a vector.
3. Derive the formula for $\mathcal{L}_u T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \ldots \nu_l}$ given by equation 5.8 from first principles, in analogy with the vector example. Verify that this Lie derivative transforms as a tensor.

4. (a) Consider the fluctuation in some vector field, $v^\mu(x) = \bar{v}^\mu(x) + \delta v^\mu(x)$ about some background value $\bar{v}^\mu(x)$. Prove that under an infinitesimal coordinate transformation, $x^\mu \to x^\mu - \xi^\mu(x)$ the vector fluctuation transforms as.

$$\delta v^\mu(x) \to \delta v^\mu(x) + \mathcal{L}_\xi \bar{v}^\mu(x) \tag{5.20}$$

Hint: let $\hat{x}^\mu = x^\mu - \xi^\mu(x)$ and use equation 4.4 to show that $v^\mu(x) \to v^\mu(\hat{x}) \equiv \hat{v}^\mu(x-\xi) = v^\mu(x) - \partial_\nu \xi^\mu v^\nu(x)$. Hence show that $\hat{v}^\mu = v^\mu + \xi^\nu \partial_\nu v^\mu - \partial_\nu \xi^\mu v^\nu$. The algebra is identical to that used in the derivation of the Lie derivative, provided we identify $\delta\lambda u^\mu$ with $-\xi^\mu$.

(b) If you are feeling confident show that equation 5.10 holds for a general tensor fluctuation.

5. Prove that the Lie derivative is Liebnizian,

$$\mathcal{L}_u(T_1 T_2) = (\mathcal{L}_u T_1)T_2 + T_1 \mathcal{L}_u T_2$$

for any tensors T_1 and T_2 (indices suppressed for brevity).

6. Show that the transformation law for the connection given by equation 5.14 is equivalent to the following transformation law,

$$\Gamma^\mu{}_{\alpha\nu}(x) \to \hat{\Gamma}^\mu{}_{\alpha\nu}(\hat{x}) = \frac{\partial \hat{x}^\mu}{\partial x^\rho}\frac{\partial x^\sigma}{\partial \hat{x}^\alpha}\frac{\partial x^\lambda}{\partial \hat{x}^\nu}\Gamma^\rho{}_{\sigma\lambda}(x) + \frac{\partial^2 x^\rho}{\partial \hat{x}^\nu \partial \hat{x}^\alpha}\frac{\partial \hat{x}^\mu}{\partial x^\rho}$$

7. Verify that the covariant derivative of a vector, defined by equation 5.13, transforms as a $(1,1)$ tensor.

8. Verify that the covariant derivative of a (k,l) tensor, defined by equation 5.16, transforms as a $(k,l+1)$ tensor.

9. By choosing a coordinate transformation $x^\mu \to \hat{x}^\mu = x^\mu + \frac{1}{2}\Gamma^\mu{}_{\alpha\beta}x^\alpha x^\beta$ show using the transformation law 5.14 that the connection vanishes to quadratic order in the coordinate.

10. Prove using the equation $\nabla_\alpha g_{\mu\nu} = 0$ that the Levi-Civita connection is given by

$$\Gamma^\alpha{}_{\mu\nu} = \frac{1}{2}g^{\alpha\beta}\left(g_{\beta\mu,\nu} + g_{\beta\nu,\mu} - g_{\mu\nu,\beta}\right) \tag{5.21}$$

6 Curvature

Gravity is geometry, and locally, the most important geometrical property is curvature. The orbits of planets around the Sun are curved precisely because spacetime is curved, because spacetime has curvature. Intuitively, what is curvature? In your mind's eye you are probably picturing the curved surface of a two dimensional sphere. You can appreciate that it is curved because you see its shape, embedded in 3 dimensional space. What you are really seeing here is the *extrinsic curvature* of the sphere. But what if you were a two dimensional ant bound to the surface of that sphere? Would you know it was curved? To do so you would need to find a way of measuring the *intrinsic* curvature of the sphere. That is a more abstract notion, and one that we will pursue in this section.

6.1 Geodesics

The simplest way to identify intrinsic curvature is to study geodesics. If they are curved, then the spacetime is curved! As you learned in your gravity course, these are the shortest paths, $x^\mu = X^\mu(\lambda)$ in a given geometry, and can be obtained by minimising the spacetime distance

$$\int \sqrt{\left|g_{\mu\nu}(X(\lambda))\frac{dX^\mu}{d\lambda}\frac{dX^\nu}{d\lambda}\right|}d\lambda \tag{6.1}$$

This yields the following differential equation

$$\frac{d^2X^\mu}{d\lambda^2} + \Gamma^\mu{}_{\alpha\beta}\frac{dX^\alpha}{d\lambda}\frac{dX^\beta}{d\lambda} = \alpha(\lambda)\frac{dX^\mu}{d\lambda} \tag{6.2}$$

where $\Gamma^\mu{}_{\alpha\beta}$ is the Levi-Civita connection and $\alpha(\lambda) = \frac{d}{d\lambda}\ln\sqrt{\left|g_{\mu\nu}(X(\lambda))\frac{dX^\mu}{d\lambda}\frac{dX^\nu}{d\lambda}\right|}$. We should also recall that in spacetime, geodesics can be categorised as timelike, null or spacelike depending on the nature of the tangent vector to the geodesic, $u^\mu = \frac{dX^\mu}{d\lambda}$.

If we choose an affine parametrisation then the RHS of the geodesic equation (6.2) vanishes and we see that it is equivalent to

$$\nabla_u u^\mu = 0 \tag{6.3}$$

This says that an affinely parametrised geodesic is a curve along which the tangent vector is transported parallel to itself. Equation 6.3, however, does not make any reference to a Levi-Civita connection and can be used to define the (affinely parametrised) geodesic even when we use any arbitrary connection. Note that torsion plays no role in geodesics since the antisymmetric part of the connection drops out of the geodesic equation.

6.1.1 Geodesic deviation

Now consider a whole family of geodesics labelled by some parameter s, and each individually parametrised by an affine parameter t. We use these parameters to make up our coordinate system in the relevant patch of the manifold, $x^\mu = (s, t, \ldots)$. The tangent vectors to the geodesic are given by $T^\mu \frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial t}$, and they satisfy the geodesic equation $\nabla_T T^\mu = 0$. The *deviation vector* represents the displacement to a neighbouring geodesic in the family, and is given by $S^\mu \frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial s}$. How does it change along an individual geodesic? This will give us information about the relative velocity and acceleration of neighbouring geodesics.

The relative velocity of a neighbouring geodesic is given by

$$v^\mu = \frac{DS^\mu}{D\lambda} = T^\nu \nabla_\nu S^\mu \tag{6.4}$$

while the relative acceleration is given by

$$a^\mu = \frac{Dv^\mu}{D\lambda} = T^\nu \nabla_\nu v^\mu \tag{6.5}$$

It is the relative acceleration that will give us more insight into curvature and the nature of the gravitational force. This stands to reason. Freely falling test particles are assumed to travel along timelike geodesics. By measuring the relative acceleration of those geodesics we are measuring the acceleration between freely falling test particles. That acceleration is due to the gravitational potential, which we are identifying with a curved geometry. An example of this would be the relative acceleration of earth and mars. Neglecting the small gravitational force of one planet on the other, their relative acceleration comes about due to the fact that they are following different accelerated trajectories in the gravitational field of the Sun.

After a small amount of algebra we find that

$$a^\mu = \nabla_T W^\mu + \nabla_W T^\mu + T^\alpha S^\beta [\nabla_\alpha, \nabla_\beta] T^\mu + \nabla_S \nabla_T T^\mu \tag{6.6}$$

where $W^\mu = T^\nu \nabla_\nu S^\mu - S^\nu \nabla_\nu T^\mu$. The last term vanishes by the geodesic equation. Also, since $T^\mu = \delta_t^\mu$ and $S^\mu = \delta_s^\mu$ are coordinate vector fields, it is clear that we have the following commutation relation,

$$T^\nu \nabla_\nu S^\mu - S^\nu \nabla_\nu T^\mu = \Gamma^\mu{}_{\alpha\nu}(S^\alpha T^\nu - T^\alpha S^\nu) = \mathcal{T}^\mu{}_{\alpha\nu} S^\alpha T^\nu$$

where $\mathcal{T}^\mu{}_{\alpha\nu}$ is the torsion. Finally, as we will see in the next section, we identify the commutator of covariant derivatives with the *Riemann tensor*, such that

$$[\nabla_\alpha, \nabla_\beta] T^\mu = -R_{\alpha\beta\nu}{}^\mu T^\nu \tag{6.7}$$

Therefore, in a torsion free theory, we are left with

$$a^\mu = -R_{\alpha\beta\nu}{}^\mu T^\nu T^\alpha S^\beta \tag{6.8}$$

This is known as the *geodesic deviation equation*. We have $a^\mu = 0$ for all families of geodesics if and only if the Riemann tensor vanishes everywhere. Otherwise some geodesics will accelerate towards or away from one another, depending on the curvature. It is worth repeating what we have already said: freely falling test particles follow geodesics, and these will accelerate towards or away from one another in the presence of curvature. This is gravity in action. To a good approximation, all planets are following different geodesics in the curved geometry sourced by the Sun, and as a result they accelerate relative to one another. This is why planets to and fro in the night sky.

The geodesic deviation equation has an analogue in Newtonian gravity, of course. A particle falling in a potential $\Phi(x)$ follows a trajectory $x^i = X^i(t)$ according to Newton's second law

$$\frac{d^2 X^i}{dt^2} = -\partial^i \Phi(X) \tag{6.9}$$

A neighbouring particle follows a trajectory $x^i = X^i(t) + S^i(t)$, where

$$\frac{d^2}{dt^2}(X^i + S^i) = -\partial^i \Phi(X + S) \tag{6.10}$$

We deduce that the relative acceleration of the two particles is given by

$$\frac{d^2S^i}{dt^2} = -(\partial^i\partial_j\Phi)S^j + \mathcal{O}(S^2) \tag{6.11}$$

This is sometimes referred to as the Newtonian deviation equation and it is the analogue of equation (6.8)

6.2 The Riemann tensor

Through geodesic deviation, we have already touched on the Riemann tensor, $R_{\mu\nu\alpha}{}^{\beta}$. To see where this arises from we first look at what happens when we commute covariant derivatives acting on 1-forms by considering the following map,

$$[\nabla_\mu, \nabla_\nu] : T_p^*(M) \to T_p^*(M) \otimes T_p^*(M) \otimes T_p^*(M) : w_\alpha \to [\nabla_\mu, \nabla_\nu]w_\alpha \tag{6.12}$$

By direct computation we see that,

$$[\nabla_\mu, \nabla_\nu]w_\alpha = R_{\mu\nu\alpha}{}^{\beta}w_\beta \tag{6.13}$$

where, in a torsion free theory,

$$R_{\mu\nu\alpha}{}^{\beta} = -2\partial_{[\mu}\Gamma^{\beta}{}_{\nu]\alpha} + 2\Gamma^{\sigma}{}_{\alpha[\mu}\Gamma^{\beta}_{\nu]\sigma} \tag{6.14}$$

Note that the Riemann tensor is given entirely in terms of the connection. We do not need to refer to a metric to define it, but we do need the metric if want to raise and lower its indices in the usual way.

The Riemann tensor shows up when we commute covariant derivatives acting on any tensor field. For a vector we have

$$[\nabla_\mu, \nabla_\nu]v^\alpha = -R_{\mu\nu\beta}{}^{\alpha}v^\beta \tag{6.15}$$

whereas for a general (k, l) tensor we have

$$[\nabla_\mu, \nabla_\nu]T^{\alpha_1\ldots\alpha_k}{}_{\beta_1\ldots\beta_l} = -\sum_{i=1}^{k} R_{\mu\nu\gamma}{}^{\alpha_i}T^{\cdots\alpha_{i-1}\gamma\alpha_{i+1}\cdots}{}_{\beta_1\ldots\beta_l} + \sum_{i=1}^{l} R_{\mu\nu\beta_i}{}^{\gamma}T^{\alpha_1\ldots\alpha_k}{}_{\ldots\beta_{i-1}\gamma\beta_{i+1}\ldots} \tag{6.16}$$

The Riemann tensor has some important properties, which we list below. For a proof of these, check out chaper 3.2 of Wald.

1. $R_{\mu\nu\alpha}{}^{\beta} = -R_{\nu\mu\alpha}{}^{\beta}$
2. $R_{[\mu\nu\alpha]}{}^{\beta} = 0$
3. $R_{\mu\nu\alpha\beta} = -R_{\mu\nu\beta\alpha}$ (only true for a Levi-Civita connection)
4. $\nabla_{[\gamma}R_{\mu\nu]\alpha}{}^{\beta} = 0$ (Bianchi identity)
5. $R_{\mu\nu\alpha\beta} = R_{\alpha\beta\mu\nu}$ (only true for a Levi-Civita connection)

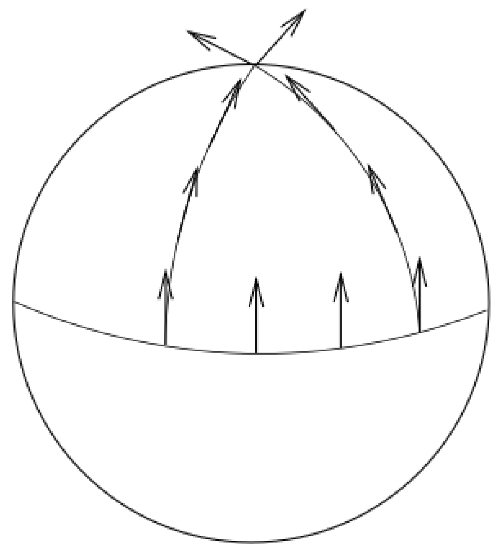


Figure 6.1: Parallel transport about a loop on a sphere.

Since we wish to embrace each of these properties, let us assume a Levi-Civita connection from now on, unless we say otherwise.

Physically we have already seen how the Riemann tensor plays a role in geodesic deviation, giving rise to acceleration between neighbouring geodesics. It also appears when we parallel transport a vector around a closed curve, measuring the amount by which it fails to return to its initial state. Indeed, consider the parallel transport of the vector around a closed loop on the curved surface of a sphere as shown in fig 6.1. The loop starts and ends at the north pole, and in between it passes through a line of longtiude, the equator and then another line of longtiude. We clearly see how the vector has changed upon return to the north pole.

Where exactly does the Riemann tensor come in? Well, to see this consider the (infinitesimally small) loop shown in figure 6.2. We choose coordinates $x^\mu = (s, t, \dots)$ so that AB and DC are tangent to $T^\mu \frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial t}$ while BC and AD are tangent to $S^\mu \frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial s}$. AB and DC lie on $s = 0$ and $s = \delta s$ respectively, while t varies. Similarly, BC and DA lie on $t = \delta t$ and $t = 0$ respectively, while s varies. We wish to parallel transport the vector, v^μ along this loop, as we go $A \to B \to C \to D \to A$. Now,

$$\delta v^\mu_{AB} = \int_0^{\delta t} dt \frac{dv^\mu}{dt}|_{s=0} = \int_0^{\delta t} dt \left[\frac{Dv^\mu}{Dt} - \Gamma^\mu{}_{\alpha\beta} v^\alpha T^\beta \right] |_{s=0} \tag{6.17}$$

The first term vanishes by parallel transport, and we are left with

$$\delta v^\mu_{AB} = -\int_0^{\delta t} dt \Gamma^\mu{}_{\alpha\beta} v^\alpha T^\beta |_{s=0} \tag{6.18}$$

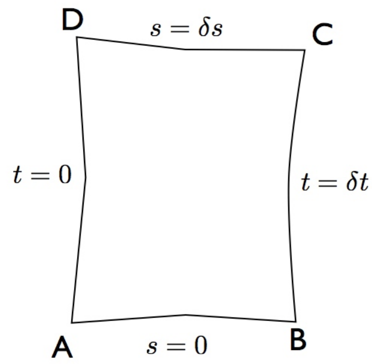


Figure 6.2: Loop around which we wish to parallely transport the vector v^μ.

Similarly we find that

$$\delta v^\mu_{BC} = -\int_0^{\delta s} ds \Gamma^\mu{}_{\alpha\beta} v^\alpha S^\beta|_{t=\delta t} \tag{6.19}$$

$$\delta v^\mu_{CD} = -\int_{\delta t}^{0} dt \Gamma^\mu{}_{\alpha\beta} v^\alpha T^\beta|_{s=\delta s} \tag{6.20}$$

$$\delta v^\mu_{DA} = -\int_{\delta s}^{0} ds \Gamma^\mu{}_{\alpha\beta} v^\alpha S^\beta|_{t=0} \tag{6.21}$$

We now see that

$$\begin{aligned}
\delta v^\mu_{AA} &= \delta v^\mu_{AB} + \delta v^\mu_{BC} + \delta v^\mu_{CD} + \delta v^\mu_{DA} \\
&= \int_0^{\delta t} dt \left[\Gamma^\mu{}_{\alpha\beta} v^\alpha T^\beta|_{s=\delta s} - \Gamma^\mu{}_{\alpha\beta} v^\alpha T^\beta|_{s=0} \right] - \int_0^{\delta s} ds \left[\Gamma^\mu{}_{\alpha\beta} v^\alpha S^\beta|_{t=\delta t} - \Gamma^\mu{}_{\alpha\beta} v^\alpha S^\beta|_{t=0} \right] \\
&\approx \delta t \delta s \left[S^\gamma \partial_\gamma (\Gamma^\mu{}_{\alpha\beta} v^\alpha) T^\beta - T^\gamma \partial_\gamma (\Gamma^\mu{}_{\alpha\beta} v^\alpha) S^\beta \right] |_{t=s=0}
\end{aligned} \tag{6.22}$$

where we have used the fact that $T^\mu = \delta^\mu_t$ and $S^\mu = \delta^\mu_s$ are coordinate vector fields.

Now,

$$S^\gamma \partial_\gamma (\Gamma^\mu{}_{\alpha\beta} v^\alpha) T^\beta|_{t=s=0} = S^\gamma \partial_\gamma (\Gamma^\mu{}_{\alpha\beta}) v^\alpha T^\beta|_{t=s=0} + S^\gamma \Gamma^\mu{}_{\alpha\beta} \partial_\gamma v^\alpha T^\beta|_{t=s=0} \tag{6.23}$$

and, because we are parallely transporting,

$$S^\gamma \partial_\gamma v^\alpha|_{t=s=0} = -\Gamma^\alpha{}_{\sigma\gamma} v^\sigma S^\gamma|_{t=s=0} \tag{6.24}$$

Thus we have that

$$S^\gamma \partial_\gamma (\Gamma^\mu{}_{\alpha\beta} v^\alpha) T^\beta|_{t=s=0} = \left[\partial_\gamma \Gamma^\mu{}_{\alpha\beta} - \Gamma^\mu{}_{\sigma\beta} \Gamma^\sigma{}_{\alpha\gamma} \right] v^\alpha S^\gamma T^\beta|_{t=s=0} \tag{6.25}$$

Similarly

$$T^\gamma \partial_\gamma (\Gamma^\mu{}_{\alpha\beta} v^\alpha) S^\beta|_{t=s=0} = [\partial_\gamma \Gamma^\mu{}_{\alpha\beta} - \Gamma^\mu{}_{\sigma\beta}\Gamma^\sigma{}_{\alpha\gamma}]\, v^\alpha T^\gamma S^\beta|_{t=s=0} \tag{6.26}$$

Given that we are in a torsion free, one can explcitly check that

$$\delta v^\mu_{AA} = \delta t \delta s R_{\beta\gamma\alpha}{}^\mu v^\alpha S^\gamma T^\beta|_{t=s=0} \tag{6.27}$$

It follows that parallel transport of a vector around a small loop will not return the vector to its initial state if the space is curved.

6.3 Other curvature tensors

The Riemann tensor may well be the mother of all curvature tensors, but what of its progeny? There are a few other curvature tensors that we ought to know about. Let's go through them, bearing in mind that we are now assuming a Levi-Civita connection.

The Ricci tensor. This is obtained by taking the appropriate trace of the Riemann tensor,

$$R_{\mu\nu} = R_{\mu\alpha\nu}{}^\alpha \tag{6.28}$$

This is a symmetric tensor, $R_{\mu\nu} = R_{\nu\mu}$, following on from property 5 of the Riemann tensor given in the previous section.

The Ricci scalar. This is obtained by taking the trace of the Ricci tensor

$$R = g^{\mu\nu} R_{\mu\nu} = R^\mu{}_\mu \tag{6.29}$$

The Einstein tensor. This is a super-important tensor, as we shall see later, and it is given by

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \tag{6.30}$$

Clearly it is a symmetric tensor. However, perhaps most importantly, it is *divergence-free*

$$\nabla^\mu G_{\mu\nu} = 0 \tag{6.31}$$

This follows from the Bianchi identity (property 4 of the Riemann tensor). Indeed, equation 6.31 is sometimes referred to as the contracted Bianchi identity for obvious reasons.

The Weyl tensor This is the traceless part of the Riemann tensor, and in D dimensions is given by

$$W_{\mu\nu\alpha\beta} = R_{\mu\nu\alpha\beta} - \frac{2}{D-2}\left(g_{\mu[\alpha}\tilde{R}_{\beta]\nu} - g_{\nu[\alpha}\tilde{R}_{\beta]\mu}\right); \qquad \tilde{R}_{\mu\nu} = R_{\mu\nu} - \frac{1}{2(D-1)} R g_{\mu\nu} \tag{6.32}$$

The Weyl tensor shares the properties 1, 2 and 3 of the Riemann tensor, and is also traceless in each of its indices. It vanishes for conformally flat metrics, $g_{\mu\nu} = e^{2\phi(x)}\eta_{\mu\nu}$

6.4 Calculating curvature

If you want to study gravitational physics then at some point you are going to have to compute a curvature tensor or two. This can be a messy business. Here we outline some standard methods for doing so.

Method 1: brute force Suppose you have a metric, $g_{\mu\nu}$. You can calculate the corresponding Levi-Civita connection,

$$\Gamma^{\mu}{}_{\alpha\beta} = \frac{1}{2} g^{\mu\nu} \left(g_{\nu\alpha,\beta} + g_{\nu\beta,\alpha} - g_{\alpha\beta,\nu} \right) \tag{6.33}$$

then plug this directly, using brute force, into our formula for the Riemann tensor,

$$R_{\mu\nu\alpha}{}^{\beta} = -2\partial_{[\mu}\Gamma^{\beta}{}_{\nu]\alpha} + 2\Gamma^{\sigma}{}_{\alpha[\mu}\Gamma^{\beta}_{\nu]\sigma} \tag{6.34}$$

You are encouraged to try this method for a conformally flat metric in Exercise 5.

Method 2: using differential geometry This is probably the simplest way to calculate the Riemann tensor, once you have mastered the art of differential geometry. You have already done a course on this, so hopefully this section won't seem like complete gobbledy gook. We start with an orthonormal basis of 1-forms, Θ^a, and calculate the curvature 2-form, $\mathcal{R}^a{}_b$ using Cartan's structural equations, with vanishing torsion,

$$d\Theta^a = -w^a{}_b \wedge \Theta^b \tag{6.35}$$

$$\mathcal{R}^a{}_b = dw^a{}_b + w^a{}_c \wedge w^c{}_b \tag{6.36}$$

Recall that the connection 1-form and the curvature 2-form have antisymmetric indices: $w_{ab} = -w_{ba}$ and $\mathcal{R}_{ab} = -\mathcal{R}_{ba}$. The curvature two form can be related to the Riemann tensor, written in an orthonormal basis via the relation,

$$\mathcal{R}^a{}_b = \frac{1}{2} R^a{}_{bcd} \Theta^c \wedge \Theta^d \tag{6.37}$$

To get the Riemann tensor in a coordinate basis we just use the vierbein

$$R^{\mu}{}_{\nu\alpha\beta} = R^a{}_{bcd} e_a{}^{\mu} e^b{}_{\nu} e^c{}_{\alpha} e^d{}_{\beta} \tag{6.38}$$

Let's do an example. Consider the cosmological metric,

$$ds^2 = -dt^2 + a(t)^2 \delta_{ij} dx^i dx^j \tag{6.39}$$

The orthonormal basis of 1-forms is given by

$$\Theta^t = dt, \qquad \Theta^i = a dx^i$$

from which we easily obtain,

$$d\Theta^t = 0, \qquad d\Theta^i = \dot{a} dt \wedge dx^i = H\Theta^t \wedge \Theta^i$$

where $H = \dot{a}/a$. It follows that the connection 1-forms are given by

$$w^t{}_i = H\Theta_i = \dot{a}dx_i, \qquad w^i{}_t = H\Theta^i = -\dot{a}dx^i, \qquad w^a{}_b = 0 \text{ otherwise}$$

It is important to ensure that these are antisymmetric, $w_{ab} = -w_{ba}$, then the solution is unique. We now compute the two form

$$\mathcal{R}^t{}_i = dw^t{}_i + w^t{}_a \wedge w^a{}_i = \ddot{a}dt \wedge dx_i = \frac{\ddot{a}}{a}\Theta^t \wedge \Theta_i \tag{6.40}$$

$$\mathcal{R}^i{}_j = dw^i{}_j + w^i{}_a \wedge w^a{}_j = H^2\Theta^i \wedge \Theta_j \tag{6.41}$$

This enables us to read off the non-trivial components of the Riemann tensor in an orthonormal basis, remembering to respect the symmetries of the Riemann tensor

$$R^t{}_{itj} = \frac{\ddot{a}}{a}\delta_{ij}, \qquad R^i{}_{jkl} = H^2(\delta^i_k\delta_{jl} - \delta^i_l\delta_{jk}) \tag{6.42}$$

All other components of the Riemann tensor are either vanishing, or follow from symmetry. Finally, we return to a coordinate basis using the vierbein to obtain,

$$R^t{}_{itj} = \frac{\ddot{a}}{a}g_{ij}, \qquad R^i{}_{jkl} = H^2(\delta^i_k g_{jl} - \delta^i_l g_{jk}) \tag{6.43}$$

Method 3: using Maple Maple contains some useful packages for working out curvature, including the "tensor" package. A word of warning though. Although the Riemann tensor is computed correctly, the Maple programmers use a different convention for the definition of the Ricci tensor. This means that their results for the Ricci tensor and Ricci scalar will differ by a sign compared with the ones used in this course, and in most of the literature you will see. Anyway, check out the following programme which computes the curvature tensors for a Schwarzschild like metric,

$$ds^2 = -V(r)dt^2 + \frac{dr^2}{V(r)} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{6.44}$$

```
> restart:

> # call the "tensor" package
> with(tensor):

> # define coordinates
> coords:=[t,  r, theta, phi]:

> # define metric components
> metric:=array(1..4, 1..4, symmetric, sparse):
> metric[1, 1]:=-V(r): metric[2, 2]:=1/V(r): metric[3, 3]:=r^2:
  metric[4, 4]:=metric[3, 3]*sin(theta)^2:

> # create the metric. Note "[-1, -1]" denotes two indices down. #
  Indices up have 1 rather than -1
> gdd:=create([-1,-1], eval(metric));
```

$$gdd := table\left(\left[index\_char = [-1, -1], compts = \begin{bmatrix} -V(r) & 0 & 0 & 0 \\ 0 & \frac{1}{V(r)} & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin(\theta)^2 \end{bmatrix}\right]\right) \quad \textbf{(1)}$$

```
> # create various objects. In the list below, we have
> # coords=coordinates
> # gdd=metric
> # guu=inverse metric
> # detg=determinant of metric
> # C1=christoffel symbol of first kind (don't worry about this)
> # C2=christoffel symbol of second kind (the Levi-Civita
  connection)
> # Rdddd=Riemann tensor, all indices down
> # Rdd=Ricci tensor, all indices down  (warning: sign convention)
> # R=Ricci scalar  (warning: sign convention)
> # Gdd= Einstein tensor, all indices down (warning: sign
  convention)
> # Cdddd=Weyl tensor
> tensorsGR( coords, gdd,guu, detg, C1, C2, Rdddd, Rdd, R, Gdd,
  Cdddd):

> #display the Ricci scalar  (warning: sign convention)
> displayGR(Ricciscalar, R);
```

The Ricci Scalar

$$R = \frac{r^2 \left(\frac{d^2}{dr^2} V(r)\right) + 4 \left(\frac{d}{dr} V(r)\right) r - 2 + 2 V(r)}{r^2} \quad \textbf{(2)}$$

```
> #display the Ricci tensor (warning: sign convention)
```

```
> displayGR(Ricci, Rdd);
```

The Ricci tensor

non-zero components :

$$R11 = -\frac{1}{2}\,\frac{V(r)\left(\left(\frac{d^2}{dr^2}V(r)\right)r+2\left(\frac{d}{dr}V(r)\right)\right)}{r}$$

$$R22 = \frac{1}{2}\,\frac{\left(\frac{d^2}{dr^2}V(r)\right)r+2\left(\frac{d}{dr}V(r)\right)}{r\,V(r)}$$

$$R33 = \left(\frac{d}{dr}V(r)\right)r-1+V(r)$$

$$R44 = \left(\frac{d}{dr}V(r)\right)r-\left(\frac{d}{dr}V(r)\right)r\cos(\theta)^2+\cos(\theta)^2-1+V(r)-V(r)\cos(\theta)^2$$

character : [-1, -1] **(3)**

```
> #display the Einstein tensor   (warning: sign convention)
> displayGR(Einstein, Gdd);
```

The Einstein Tensor

non-zero components :

$$G11 = \frac{V(r)\left(\left(\frac{d}{dr}V(r)\right)r-1+V(r)\right)}{r^2}$$

$$G22 = -\frac{\left(\frac{d}{dr}V(r)\right)r-1+V(r)}{V(r)\,r^2}$$

$$G33 = -\left(\frac{d}{dr}V(r)\right)r-\frac{1}{2}r^2\left(\frac{d^2}{dr^2}V(r)\right)$$

$$G44 = -\left(\frac{d}{dr}V(r)\right)r+\left(\frac{d}{dr}V(r)\right)r\cos(\theta)^2-\frac{1}{2}r^2\left(\frac{d^2}{dr^2}V(r)\right)+\frac{1}{2}r^2\left(\frac{d^2}{dr^2}V(r)\right)\cos(\theta)^2$$

character : [-1, -1] **(4)**

Exercises

1. Show that in a torsion free theory,
$$[\nabla_\mu, \nabla_\nu](f_1 w^{(1)}_\alpha + f_2 w^{(2)}_\alpha) = f_1[\nabla_\mu, \nabla_\nu]w^{(1)}_\alpha + f_2[\nabla_\mu, \nabla_\nu]w^{(2)}_\alpha$$
where f_1 and f_2 are scalar functions, and $w^{(1)}_\alpha$ and $w^{(2)}_\alpha$ are 1-forms.

2. Show that
$$[\nabla_\mu, \nabla_\nu]\nabla_\alpha w_\beta = R_{\mu\nu\alpha}{}^\gamma \nabla_\gamma w_\beta + R_{\mu\nu\beta}{}^\gamma \nabla_\alpha w_\gamma$$

3. Show that property 5 of the Riemann tensor given in section 6.2 implies that Ricci tensor is symmetric.

4. Using property 4 of the Riemann tensor given in section 6.2, prove the contracted Bianchi identity (6.31). *Hint: contract γ and β, then μ and α.*

5. Show that the Levi-Civita connection for a conformally flat metric, $g_{\mu\nu} = e^{2\phi(x)}\eta_{\mu\nu}$, is given by
$$\Gamma^\alpha{}_{\mu\nu} = \delta^\alpha_\mu \partial_\nu \phi + \delta^\alpha_\nu \partial_\mu \phi - \eta_{\mu\nu}\partial^\alpha \phi$$
Now use the brute force method to show that the Riemann tensor is given by
$$R_{\mu\nu\alpha}{}^\beta = 2\delta^\beta_{[\mu}\partial_{\nu]}\partial_\alpha\phi - 2\eta_{\alpha[\mu}\partial_{\nu]}\partial^\beta\phi - 2\delta^\beta_{[\mu}\partial_{\nu]}\phi\partial_\alpha\phi + 2\eta_{\alpha[\mu}\partial_{\nu]}\phi\partial^\beta\phi - 2\eta_{\alpha[\mu}\delta^\beta_{\nu]}\partial^\gamma\phi\partial_\gamma\phi$$

7 The Einstein equation

The great relativist, John Wheeler, famously said: "Matter tells space how to curve, and space tells matter how to move". As long ago as your *Gravity* course you got to grips with the second half of this statement. Matter particles follow geodesic paths on curved spacetime. Now that we have developed a better mathematical understanding of what curvature is, we are ready to try to understand the first half of Wheeler's statement. How exactly does matter tell space how to curve?

Before we dive headlong into Einstein's brilliant equation, let us pause for a moment to consider Newtonian gravity. For Newtonian gravity, Wheeler's statement can be modified slightly to read "Matter sets up the gravitational potential, and the gravitational potential tells particles how to move" . Again, the latter half of this statement is understood through Newton's second law
$$\frac{d^2x^i}{dt^2} = -\partial^i \Phi(x) \tag{7.1}$$
where the particle moves along a trajectory $x^i(t)$ under the influence of the potential $\Phi(x)$. But what about the first half of the statement? What sets up the potential? Well, we know that for a non-relativistic matter distribution of density $\rho(x)$, the gravitational potential is given by the Poisson equation
$$\partial^i \partial_i \Phi = 4\pi G_N \rho \tag{7.2}$$

where G_N is Newton's constant. The energy density is obtained from the energy δE contained in an infinitesimal spatial volume δV via the relation

$$\delta E = \rho \delta V \tag{7.3}$$

Energy density is, of course, conserved for a non-relativistic fluid $\dot{\rho} = 0$.

7.1 The stress energy tensor

Now suppose we wish to consider a relativistic matter distribution (in flat space). Things get more complicated. A relativistic fluid can exert pressure and anisotropic stress. It is reasonable to suggest that these will play a role in setting up the curvature, along with energy density. Also, from special relativity we know that the energy, δE, is not a Lorentz scalar, but the time component of a vector: the four-momentum,

$$\delta p^\mu = \begin{pmatrix} \delta E \\ \delta p^i \end{pmatrix}$$

where δp^i is the 3-momentum contained in δV. The spatial volume itself is also ill defined at this stage, because in 4 dimensional spacetime it is merely a surface and we need to specify its orientation. This requires us to introduce the spacetime normal to that surface, n_ν, and the resulting surface element $n_\nu \delta V$. What, then, is the relativistic analogue of equation 7.3? Since we have energy stored within δp^μ and spatial volume encoded by $n_\nu \delta V$, to extract the energy density we need an object with two indices, $T^{\mu\nu}$. This is the *energy-momentum* or *stress-energy* tensor, and it enables us to write

$$\delta p^\mu = T^{\mu\nu} n_\nu \delta V \tag{7.4}$$

Using an inertial frame in Minkowski space with coordinates (t, x^i), the stress-energy tensor decomposes as follows:

$$\begin{aligned} T^{tt} &= \text{energy density} & T^{tj} &= \text{energy flux} \\ T^{it} &= \text{momentum density} & T^{ij} &= \text{stress tensor} \end{aligned} \tag{7.5}$$

Energy flux and momentum density are always identical[5], $T^{ti} = T^{it}$, and because the stress tensor is symmetric $T^{ij} = T^{ji}$, we deduce that the full stress-energy tensor is symmetric, $T^{\mu\nu} = T^{\nu\mu}$. The stress tensor is hopefully familiar from classical mechanics. T^{ij} measures the ith component of the stress (force/area) exerted across a surface in the j direction. The diagonal components are the components of pressure while the off-diagonal components are usually referred to as anisotropic stress.

What is the conservation law for a relativistic fluid? Well, in flat space it is given by

$$\partial_\mu T^{\mu\nu} = 0 \tag{7.6}$$

[5] energy flux=(energy density)×(velocity)=momentum density.

The t component of this equation coincides with the conservation of mass equation, familiar from fluid mechanics,

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{\pi} = 0$$

where ρ is the energy density and $\vec{\pi}$ is the momentum density. The i component is a statement of Newton's second law,

$$\frac{\partial \pi^i}{\partial t} = -\partial_j T^{ij} = f^i$$

where f^i is the force density (see Exercise 1).

All of this can be promoted to curved space, only now the stress-energy tensor, $T^{\mu\nu}$, is a $(2,0)$ tensor in a curved spacetime. The flat space conservation law (7.6) must be replaced by its covariant counterpart,

$$\nabla_\mu T^{\mu\nu} = 0 \tag{7.7}$$

The precise form of the stress-energy tensor depends on the properties of the fluid. A particularly important example is a *perfect fluid*. This is a fluid that is completely characterised by its energy density and pressure. If the fluid has four velocity u^μ, normalised such that $g_{\mu\nu}u^\mu u^\nu = -1$, then the stress-energy tensor is given by

$$T^{\mu\nu} = \rho u^\mu u^\nu + p\gamma^{\mu\nu} \tag{7.8}$$

where $\gamma^{\mu\nu} = g^{\mu\nu} + u^\mu u^\nu$ is the projection operator on to the spatial surfaces orthogonal to u^μ. In cosmology, the stress-energy tensor takes the form of a perfect fluid, in accordance with the assumption of homogeneity and isotropy. A fluid composed of non-relativistic matter has $p = 0$, while radiation has $p = \frac{\rho}{3}$. The vacuum energy, or cosmological constant, has $p = -\rho$.

7.2 The Einstein equation

Remember, our goal was to find the covariant analogue of the Poisson equation (7.2). We have already seen how the right hand side must be promoted to the stress-energy tensor. What about the left hand side? Clearly we need a tensor with two indices, and in the spirit of Wheeler's statement, this should be related to curvature. Einstein's original idea was to use the Ricci tensor, and he proposed an equation of the form

$$R_{\mu\nu} = \kappa T_{\mu\nu} \tag{7.9}$$

However, this is clearly wrong. The right hand side is divergence free by energy conservation, $\nabla^\mu T_{\mu\nu} = 0$, but the left hand side is not, $\nabla^\mu R_{\mu\nu} \neq 0$. Einstein corrected this mistake by replacing the Ricci tensor with the Einstein tensor, which *is* divergence free. Thus we arrive at Einstein's equation in all its glory

$$G_{\mu\nu} = 8\pi G_N T_{\mu\nu} \tag{7.10}$$

According to Einstein, this is how matter tells spacetime how to curve. Note that the stress-energy tensor does not specify the geometry uniquely in four dimensions. This is easy to see. There are 10 components of the energy momentum tensor, but, taking into account its symmetries, there are 20 independent components of the Riemann tensor (see Misner, Thorne and Wheeler, p220). In this sense Einstein's theory is not completely Machian and as we shall see later, it leaves room for gravitational waves. What Einstein's equation does give us is 10 coupled non-linear partial differential equations for the 10 components of the metric. The name of the game in General Relativity is to try to solve for this metric. That is obviously very difficult in general and typically we can only make progress in the following cases

- plenty of symmetry (so the metric admits some killing vectors)
- perturbation theory (expand about a known solution, $g_{\mu\nu} = \underbrace{\bar{g}_{\mu\nu}}_{known} + \underbrace{h_{\mu\nu}}_{small}$)

In a moment, we will consider some examples with plenty of symmetry. A study of perturbation theory is postponed until section 8.

In any event, we end this section by noting that we can add a bare cosmological constant term to the left hand side of Einstein's equation without spoiling its covariance,.

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G_N T_{\mu\nu} \tag{7.11}$$

Indeed, in four dimensions, Lovelock's theorem proves that a linear combination $\alpha G_{\mu\nu} + \beta g_{\mu\nu}$ is the most general $(0, 2)$ tensor that one can construct out of the metric and its first two derivatives that is both symmetric and divergence free. Recall that the vacuum energy, ρ_{vac}, contributes a similar term to the stress-energy tensor $T^{vac}_{\mu\nu} = -\rho_{vac} g_{\mu\nu}$. Λ may therefore be absorbed into a redefinition of $\rho_{vac} \to \rho_{vac} + \frac{\Lambda}{8\pi G_N}$, or vice versa, $\Lambda \to \Lambda + 8\pi G_N \rho_{vac}$. This is important in understanding the cosmological constant problem.

7.3 Maximally symmetric solutions

Let us consider the vacuum Einstein's equations in the presence of a cosmological constant, $G_{\mu\nu} + \Lambda g_{\mu\nu} = 0$. An important class of solutions to this equation are the *maximally symmetric solutions*. These have a maximal number of Killing vectors, and a Riemann tensor of the form

$$R_{\mu\nu\alpha\beta} = \kappa(g_{\mu\alpha} g_{\nu\beta} - g_{\mu\beta} g_{\nu\alpha}) \tag{7.12}$$

One can easily show that we must have $\kappa = \frac{\Lambda}{3}$. In a static global coordinate system the metric is given by

$$ds^2 = -(1 - \kappa r^2) dt^2 + \frac{1}{1 - \kappa r^2} + r^2 (d\theta^2 + \sin^2\theta d\phi^2) \tag{7.13}$$

The solution with $\Lambda > 0$ is known as de Sitter (dS) space, while the one with $\Lambda < 0$ is known as anti-de Sitter space (AdS). Of course, for $\Lambda = 0$ we have Minkowski space.

7.4 Cosmological equations

In cosmology we have lots of symmetry: homogeneity and isotropy. For a spatially flat cosmology we have a Friedmann-Robertson-Walker metric

$$ds^2 = -dt^2 + a(t)^2 \delta_{ij} dx^i dx^j \tag{7.14}$$

We have already calculated the Riemann tensor for this metric and can use equations (6.43) to compute the Einstein tensor,

$$G_{tt} = 3H^2, \qquad G_{it} = 0, \qquad G_{ij} = -\left(H^2 + 2\frac{\ddot{a}}{a}\right) g_{ij}$$

The cosmological fluid is a perfect fluid with 4-velocity $u^\mu = (1, 0, 0, 0)$, so

$$T_{tt} = \rho, \qquad T_{it} = 0, \qquad T_{ij} = p g_{ij}$$

The tt component of the Einstein equation now gives the Friedmann Equation,

$$H^2 = \frac{8\pi G_N}{3}\rho \tag{7.15}$$

while the ij equation, combined with the Friedmann equation, gives the Raychauduri equation,

$$\frac{\ddot{a}}{a} = -\frac{4\pi G_N}{3}(\rho + 3p) \tag{7.16}$$

You will see how to solve these equations in certain cases in your cosmology course. Note further that the energy conservation equation is given by

$$\dot{\rho} + 3H(\rho + p) = 0 \tag{7.17}$$

It is straightforward to check that Raychauduri eqn (12.10) can be derived from the Friedmann equation (7.15) and the energy conservation equation (7.17).

7.5 Birkhoff's theorem

Birkhoff's theorem states that in general relativity, any spherically symmetric solution of the vacuum field equations must be static and asymptotically flat. This is an important result and implies that any spherically symmetric vacuum solution must corresponds to a section of the Schwarzschild geometry. Let's prove this.

Since we are in vacuum, we have $T_{\mu\nu} = 0$ and our goal is simply to solve the vacuum Einstein equation $G_{\mu\nu} = 0$. This is true if and only if $R_{\mu\nu} = 0$. Now consider a general spherically symmetric metric in four dimensions. This has the form

$$ds^2 = -A(t,r)dt^2 + B(t,r)dtdr + C(t,r)dr^2 + D(t,r)(d\theta^2 + \sin^2\theta d\phi^2) \tag{7.18}$$

Actually, we can exploit some residual coordinate invariance to simplify the metric some more. We don't want to mess with the spherical symmetry, but we can use transformations of the form

$$t \to \alpha(t,r), \qquad r \to \beta(t,r) \tag{7.19}$$

for suitably chosen functions α and β, to reduce the metric to the following form[6]

$$ds^2 = -e^{2a(t,r)}dt^2 + e^{2b(t,r)}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{7.20}$$

The non-trivial components of the Ricci tensor are now found to be

$$R_{tt} = \ddot{b} + \dot{b}^2 - \dot{a}\dot{b} + e^{2(a-b)}\left(a'' + a'^2 - a'b' + \frac{2}{r}a'\right) \tag{7.21}$$

$$R_{rr} = -\left(a'' + a'^2 - a'b' - \frac{2}{r}b'\right) + e^{2(b-a)}\left(\ddot{b} + \dot{b}^2 - \dot{a}\dot{b}\right) \tag{7.22}$$

$$R_{tr} = \frac{2}{r}\dot{b} \tag{7.23}$$

$$R_{\theta\theta} = 1 - e^{-2b} + re^{-2b}(b' - a') \tag{7.24}$$

$$R_{\phi\phi} = R_{\theta\theta}\sin^2\theta \tag{7.25}$$

where dot denotes ∂_t and prime denotes ∂_r. Now $R_{tr} = 0$ implies that $\dot{b} = 0$. Plugging this into $R_{\theta\theta} = 0$, and differentiating with respect to time, we find that $\dot{\alpha}' = 0$. Therefore, we can write

$$b = b(r), \qquad a = f(r) + g(t) \tag{7.26}$$

Also, we now have

$$0 = e^{2(b-a)}R_{tt} + R_{rr} = \frac{2}{r}(a' + b')$$

which implies that $b = c - f$, where c is constant. We can eliminate c by redefiniing $f \to f + c$ and $g \to g - c$, so that we now have $b = -f$. Now the metric has the form

$$ds^2 = -e^{2f(r)}e^{2g(t)}dt^2 + e^{-2f(r)}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{7.27}$$

and it is clear that we can eliminate $g(t)$ by means of a coordinate transformation of the form $t \to T(t)$. Setting $g(t) = 0$ we arrive at the *static* metric

$$ds^2 = -e^{2f(r)}dt^2 + e^{-2f(r)}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{7.28}$$

It remains to show that this is equivalent to the Schwarzschild metric. From $R_{\theta\theta} = 0$ we have

$$e^{2f}(2rf' + 1) = 1 \qquad \Longrightarrow \qquad (re^{2f})' = 1 \qquad \Longrightarrow \qquad e^{2f} = 1 - \frac{\mu}{r}$$

where μ is an integration constant.

We have thus proven that the solution is (a section of) the Schwarzschild geometry.

$$ds^2 = -\left(1 - \frac{\mu}{r}\right)dt^2 + \frac{dr^2}{1 - \frac{\mu}{r}} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{7.29}$$

[6]Strictly speaking this is only true provided $\nabla_\mu D$ is spacelike, thereby enabling us to identify D with a spacelike component. One must deal with the alternative possibility separately and check that it does not yield additional solutions.

Exercises

1. Consider a volume V with surface S. Explain why the force acting on the volume due to the anisotropic stress is
$$F^j = \int_S dS n_i^{(in)} T^{ij}$$
where $n_i^{(in)}$ is the normal pointing *into* the volume. Use the divergence theorem to show that the force density, f^j defined by $F^j = \int_V dV f^j$, is given by
$$f^j = -\partial_i T^{ij}$$

2. Verify that the trace of the stress-energy tensor vanishes for radiation.

3. If the Riemann tensor is given by the maximally symmetric solution,
$$R_{\mu\nu\alpha\beta} = \kappa(g_{\mu\alpha}g_{\nu\beta} - g_{\mu\beta}g_{\nu\alpha})$$
show that the Einstein tensor is given by $G_{\mu\nu} = -3\kappa g_{\mu\nu}$.

4. Use equations (6.43) to compute the Einstein tensor for a Friedmann-Robertson-Walker metric,
$$G_{tt} = 3H^2, \qquad G_{it} = 0, \qquad G_{ij} = -\left(H^2 + 2\frac{\ddot{a}}{a}\right) g_{ij}$$

5. Verify that the energy-conservation equation $\nabla^\mu T_{\mu\nu} = 0$ gives equation (7.17) in a cosmological setting. Show further that equation that Raychauduri eqn (12.10) can be derived from the Friedmann equation (7.15) and the energy conservation equation (7.17).

6. Consider the cosmological equations sourced by vacuum energy, $\rho > 0$, with $p = -\rho$. Show that $\dot{\rho} = 0$, and explain why the maximally symmetric de Sitter solution can be written as
$$ds^2 = -dt^2 + e^{2Ht}\delta_{ij}dx^i dx^j \tag{7.30}$$
where $H^2 = \frac{8\pi G_N}{3}\rho =$ constant. This is known as a flat slicing of de Sitter space.

7. Show that a under the change of coordinates given by (7.19), the metric (7.18) transforms into
$$ds^2 = -\hat{A}(t,r)dt^2 + \hat{B}(t,r)dtdr + \hat{C}(t,r)dr^2 + \hat{D}(t,r)(d\theta^2 + \sin^2\theta d\phi^2)$$
where
$$\hat{A}(t,r) = A(\alpha,\beta)\dot{\alpha}^2 - B(\alpha,\beta)\dot{\alpha}\dot{\beta} - C(\alpha,\beta)\dot{\beta}^2 \tag{7.31}$$
$$\hat{B}(t,r) = B(\alpha,\beta)(\dot{\alpha}\beta' + \alpha'\dot{\beta}) - 2A(\alpha,\beta)\dot{\alpha}\alpha' + 2C(\alpha,\beta)\dot{\beta}\beta' \tag{7.32}$$
$$\hat{C}(t,r) = C(\alpha,\beta)\beta'^2 + B(\alpha,\beta)\alpha'\beta' - A(\alpha,\beta)\alpha'^2 \tag{7.33}$$
$$\hat{D}(t,r) = D(\alpha,\beta) \tag{7.34}$$
Convince yourself that one can choose α and β such that this metric now takes the form given by equation (7.20), provided $\nabla_\mu D$ is spacelike.

8 Weak gravity

One of the most powerful tools in GR is perturbation theory. This is when we consider small perturbations about a known background solution,

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}, \qquad |h_{\mu\nu}| \ll 1 \tag{8.1}$$

Here $\bar{g}_{\mu\nu}$ is some known metric that describes the geometry of the background spacetime. Often this is taken to be a maximally symmetric space, or a Friedmann-Robertson-Walker solution. $h_{\mu\nu}$ describes metric fluctuations on the background and is assumed to be small. We can think of it as a spin 2 particle (the graviton) propagating on the background geometry.

The first thing to check is that the inverse metric

$$g^{\mu\nu} = \bar{g}^{\mu\nu} - h^{\mu\nu} + \mathcal{O}(h^2) \tag{8.2}$$

where the indices on h are raised and lower using $\bar{g}^{\mu\nu}$ and $\bar{g}_{\mu\nu}$. We can also compute the leading order fluctuation in the connection and the Riemann tensor. For the connection, it is convenient to go to Riemann normal coordinates on the background, as we demonstrate below,

$$\begin{aligned}
\delta\Gamma^{\alpha}{}_{\mu\nu} &= \frac{1}{2}\bar{g}^{\alpha\beta}(h_{\beta\mu,\nu} + h_{\beta\nu,\mu} - \frac{1}{2}h_{\mu\nu,\beta}) - h^{\alpha\beta}(\bar{g}_{\beta\mu,\nu} + \bar{g}_{\beta\nu,\mu} - \bar{g}_{\mu\nu,\beta}) \\
&= \frac{1}{2}\eta^{\alpha\beta}(h_{\beta\mu,\nu} + h_{\beta\nu,\mu} - h_{\mu\nu,\beta}) \qquad \text{Riemann normal coordinates: } \bar{g} \to \eta \\
&= \frac{1}{2}\bar{g}^{\alpha\beta}(h_{\beta\mu;\nu} + h_{\beta\nu;\mu} - h_{\mu\nu;\beta})
\end{aligned}$$

In the last line we have promoted the right hand side to a tensor on the background, replacing $\eta \to \bar{g}$, and replacing partial derivatives with covariant derivatives. The semi-colon here denotes a covariant derivative on the background, $Q^{\cdots}_{\cdots;\mu} = \bar{\nabla}_\mu Q^{\cdots}_{\cdots}$. Since we now have a tensor equation, it holds in any coordinate system and we can write

$$\delta\Gamma^{\alpha}{}_{\mu\nu} = \frac{1}{2}\left(h^{\alpha}_{\mu;\nu} + h^{\alpha}_{\nu;\mu} - h^{;\alpha}_{\mu\nu}\right) \tag{8.3}$$

A word of warning: Riemann normal coordinates are only suitable for computing quantities containing up to one derivative. You cannot use them for second derivatives. Fortunately, the Riemann tensor only contains first derivatives of the connection, which is now known up to leading order. So, again, using Riemann normal coordinates we can show that

$$\begin{aligned}
\delta R_{\mu\nu\alpha}{}^{\beta} &= -2\partial_{[\mu}\delta\Gamma^{\beta}{}_{\nu]\alpha} + 2\bar{\Gamma}^{\sigma}{}_{\alpha[\mu}\delta\Gamma^{\beta}_{\nu]\sigma} + 2\delta\Gamma^{\sigma}{}_{\alpha[\mu}\bar{\Gamma}^{\beta}_{\nu]\sigma} \\
&= -2\partial_{[\mu}\delta\Gamma^{\beta}{}_{\nu]\alpha} \qquad \text{Riemann normal coordinates: } \bar{\Gamma} \to 0 \\
&= -2\bar{\nabla}_{[\mu}\delta\Gamma^{\beta}{}_{\nu]\alpha}
\end{aligned}$$

This is known as the Palatini identity. Again, in the last line we have promoted the right hand side to a tensor, by replacing partial derivatives with covariant derivatives on the background. Furthermore, we have a tensor equation, and plugging in equation (8.3) we find[7]

$$\delta R_{\mu\nu\alpha}{}^{\beta} = h^{\beta}_{\alpha;[\mu\nu]} + h^{\beta}_{[\mu;|\alpha|\nu]} - h_{\alpha[\mu;}{}^{\beta}{}_{\nu]} \tag{8.4}$$

The perturbed Einstein equation is, of course, given by

$$\delta G_{\mu\nu} = 8\pi G_N \delta T_{\mu\nu} \tag{8.5}$$

or equivalently

$$\delta R_{\mu\nu} = 8\pi G_N \delta \left[T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu} \right] \tag{8.6}$$

where $T = g^{\mu\nu} T_{\mu\nu}$ is the trace of the stress-energy tensor. Using the result (8.4) and commuting some covariant derivatives we can show that

$$\delta R_{\mu\nu} = -\frac{1}{2} \Delta_L h_{\mu\nu} \tag{8.7}$$

where the *Lichnerowicz operator*,

$$\Delta_L h_{\mu\nu} = \bar{\Box} h_{\mu\nu} - 2\bar{\nabla}_{(\mu} \bar{\nabla}^{\alpha} \bar{h}_{\nu)\alpha} - 2\bar{R}_{\alpha(\mu} h^{\alpha}_{\nu)} + 2\bar{R}_{\mu\alpha\nu\beta} h^{\alpha\beta} \tag{8.8}$$

and $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} h g_{\mu\nu}$ is the trace reversed metric perturbation, and $h = h^{\alpha}_{\alpha}$ is the trace. Here $\bar{R}_{\mu\nu\alpha\beta}$ is the background Riemann tensor, and $\bar{\Box} = \bar{\nabla}_{\alpha} \bar{\nabla}^{\alpha}$.

It follows that the perturbed Einstein equations are equivalent to

$$-\frac{1}{2} \Delta_L h_{\mu\nu} = 8\pi G_N \delta \left[T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu} \right] \tag{8.9}$$

8.1 Choosing a gauge

The perturbed Einstein equations are a tensor equation, which means they hold in any coordinate system. Now, under an infinitesimal change of coordinates, $x^{\mu} \to x^{\mu} - \xi^{\mu}$, we know that

$$\delta g_{\mu\nu} \to \delta g_{\mu\nu} + \mathcal{L}_{\xi} \bar{g}_{\mu\nu} \tag{8.10}$$

or equivalently

$$h_{\mu\nu} \to h_{\mu\nu} + \bar{\nabla}_{\mu} \xi_{\nu} + \bar{\nabla}_{\nu} \xi_{\mu} \tag{8.11}$$

Here we have used the fact that

$$\mathcal{L}_{\xi} \bar{g}_{\mu\nu} = \bar{\nabla}_{\mu} \xi_{\nu} + \bar{\nabla}_{\nu} \xi_{\mu}$$

We can use this freedom to change coordinates to place restrictions on $h_{\mu\nu}$ that render the equations of motion easier to solve. This is called *choosing the gauge*. For example, a popular gauge choice is called *de donder* or *harmonic* gauge. This corresponds to setting

$$\bar{\nabla}^{\mu} \bar{h}_{\mu\nu} = 0$$

[7]Note that the vertical lines around the index α in the middle term indicates that it is not included in the antisymmerization.

To see why this is possible we note that under $x^\mu \to x^\mu - \xi^\mu$, we have

$$\bar{\nabla}^\mu \bar{h}_{\mu\nu} \to \bar{\nabla}^\mu \bar{h}_{\mu\nu} + \bar{\Box}\xi_\nu - \bar{R}_{\mu\nu\alpha}{}^\mu \xi^\alpha$$

Thus we see that for a suitable choice of ξ^μ we can set everything to vanish to the right of the arrow. Of course, other gauge choices are possible. Another gauge that is commonly used on static backgrounds is *synchronous* gauge. This corresponds to

$$h_{tt} = h_{ti} = 0$$

where we have explicitly separated our background coordinates into time and space $x^\mu = (t, x^i)$.

Let us see how de donder gauge is used on maximally symmetric backgrounds. Since

$$\bar{R}_{\mu\nu\alpha\beta} = \kappa(\bar{g}_{\mu\alpha}\bar{g}_{\nu\beta} - \bar{g}_{\mu\beta}\bar{g}_{\nu\alpha}),$$

the Lichnerowicz operator is given by

$$\Delta_L h_{\mu\nu} = \bar{\Box} h_{\mu\nu} - 4\kappa\left(h_{\mu\nu} - \frac{1}{4} h g_{\mu\nu}\right)$$

Note that the contribution from $-2\bar{\nabla}_{(\mu}\bar{\nabla}^\alpha \bar{h}_{\nu)\alpha}$ vanishes on account of the gauge condition. The stress-energy tensor is given by $T_{\mu\nu} = -\frac{3\kappa}{8\pi G_N} g_{\mu\nu} + \tau_{\mu\nu}$, where we identify the fluctuation as

$$\delta T_{\mu\nu} = -\frac{3\kappa}{8\pi G_N} h_{\mu\nu} + \tau_{\mu\nu}$$

It follows from the perturbed Einstein equation (8.9) that

$$-\frac{1}{2}\bar{\Box} h_{\mu\nu} + 2\kappa\left(h_{\mu\nu} - \frac{1}{4} h g_{\mu\nu}\right) = 3\kappa h_{\mu\nu} + 8\pi G_N\left(\tau_{\mu\nu} - \frac{1}{2}\tau \bar{g}_{\mu\nu}\right) \tag{8.12}$$

where $\tau = \bar{g}^{\mu\nu}\tau_{\mu\nu}$. This equation can be split into its trace and traceless parts,

$$\begin{aligned} -\frac{1}{2}(\bar{\Box} + 2\kappa)\left(h_{\mu\nu} - \frac{1}{4} h \bar{g}_{\mu\nu}\right) &= 8\pi G_N\left(\tau_{\mu\nu} - \frac{1}{4}\tau \bar{g}_{\mu\nu}\right) \\ -\frac{1}{2}(\bar{\Box} + 6\kappa) h &= -8\pi G_N \tau \end{aligned} \tag{8.13}$$

It follows that[8]

$$h_{\mu\nu} = -\frac{16\pi G_N}{\bar{\Box} + 2\kappa}\left[\tau_{\mu\nu} - \frac{1}{2}\left(\frac{\bar{\Box} + 4\kappa}{\bar{\Box} + 6\kappa}\right)\tau \bar{g}_{\mu\nu}\right] \tag{8.14}$$

On Minkowski space, this reduces to the oft quoted result,

$$h_{\mu\nu} = -\frac{16\pi G_N}{\bar{\Box}}\left[\tau_{\mu\nu} - \frac{1}{2}\tau \eta_{\mu\nu}\right] \tag{8.15}$$

If the gravitational field is sourced by stress-energy, $\tau^s_{\mu\nu}$, then the gravitational potential energy of a probe of stress-energy, $\tau^p_{\mu\nu}$ is given by

$$V = -\frac{1}{2}\int d^3x\, h^s_{\mu\nu}(x)\tau_p^{\mu\nu}(x) = 8\pi G_N \int d^3x \left[\tau_p^{\mu\nu}(x)\frac{1}{\bar{\Box}}\tau^s_{\mu\nu}(x) - \frac{1}{2}\tau_p(x)\frac{1}{\bar{\Box}}\tau^s(x)\right] \tag{8.16}$$

[8]For an operator such as $\Box + 2\kappa$, the operator $\frac{1}{\Box + 2\kappa}$ denotes the inverse.

8.2 Decomposing into tensors, vectors and scalars

In your *Modern Cosmology* course you will study perturbations upon cosmological backgrounds

$$\bar{g}_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a(t)^2\delta_{ij}dx^i dx^j$$

It is convenient to classify these perturbations with respect to the diffeomorphism group on the spatial slices (in this case, coordinate transformations on 3 dimensional Euclidean space). We therefore write the metric perturbations as

$$\begin{aligned} h_{tt} &= -2\phi \\ h_{ti} &= a(\beta_i + \partial_i\beta), \\ h_{ij} &= a^2\left[\tilde{h}_{ij} + 2\partial_{(i}\nu_{j)} + 2\left(\partial_i\partial_j - \frac{\delta_{ij}}{3}\Delta\right)\alpha + 2\psi\delta_{ij}\right] \end{aligned} \tag{8.17}$$

where $\Delta = \partial^k\partial_k$. Here we have four scalars:

$$\phi(t,x), \qquad \beta(t,x), \qquad \alpha(t,x), \qquad \psi(t,x)$$

two vectors:

$$\beta_i(t,x), \qquad \nu_i(t,x)$$

and one tensor:

$$\tilde{h}_{ij}(t,x)$$

The vectors are divergence free (ie $\partial^i\beta_i = \partial^i\nu_i = 0$) and as a result each encode two degrees of freedom. The tensor is transverse-tracefree (ie $\partial^j\tilde{h}_{ij} = 0 = \delta^{ij}\tilde{h}_{ij}$) and as a result also encodes two degrees of freedom. In total, then we have 4×scalar+4×vector+2×tensor=10 degrees of freedom, which is exactly what we expect.

Having done this, it is natural to decompose the infinitesimal coordinate transformations in the same way. We therefore write these as

$$t \to t - \epsilon, \qquad x^i \to x^i - \xi^i - \partial^i\xi \tag{8.18}$$

where $\epsilon(t,x)$ and $\xi(t,x)$ correspond to the scalar transformations, and $\xi^i(t,x)$ corresponds to the vector transformation (there is no room for a tensor transformation). One can also decompose the components of the stress-energy tensor in this way.

The reason all of this is a good thing to do is that to linear order the different classes do not mix. In other words, scalars only couple to scalars, vectors only couple to vectors, and tensors only couple to tensors. To convince yourself that this must be true try to imagine how, say, a vector could source a scalar. The only operators we have available are derivative operators with respect to time and space. To get a scalar from the vector you would need to take the divergence, but by construction the vectors are divergence free, so this is futile.

8.2.1 Linearised gravity around a non-relativistic source

Let us proceed with the following example. Consider perturbations on a Minkowski background ($a(t) \equiv 1$) in the presence of a non-relativistic source

$$\delta T_{tt} = \rho(x), \qquad \delta T_{ti} = 0, \qquad \delta T_{ij} = 0$$

The stress-energy tensor contains a single scalar component. The perturbed Einstein equations can be decomposed into the scalar, vector and tensor equations. The scalar equations of motion are given by

$$\begin{aligned} \Delta\Phi - 3\ddot{\Psi} &= 4\pi G_N \rho && (8.19) \\ -2\partial_i \dot{\Psi} &= 0 && (8.20) \\ -(-\partial_t^2 + \Delta)\Psi\delta_{ij} - \partial_i\partial_j(\Phi + \Psi) &= 4\pi G_N \rho \delta_{ij} && (8.21) \end{aligned}$$

where we have introduced

$$\Phi(t, x) = \phi + \dot{\beta} - \ddot{\alpha}, \qquad \Psi(t, x) = \psi - \frac{\Delta}{3}\alpha \tag{8.22}$$

The vector equations of motion, meanwhile, are given by

$$\begin{aligned} -\frac{1}{2}\Delta B_i &= 0 && (8.23) \\ -\partial_{(i}\dot{B}_{j)} &= 0 && (8.24) \end{aligned}$$

where we have introduced

$$B_i(t, x) = \beta_i - \dot{\nu}_i \tag{8.25}$$

Finally the tensor equations of motion are given by

$$-\frac{1}{2}(-\partial_t^2 + \Delta)\tilde{h}_{ij} = 0 \tag{8.26}$$

These equations of motion have been written explicitly in terms of *gauge invariants*, Φ, Ψ, B_i and $\tilde{h}_{ij}$. These are quantities that are invariant under the transformations (8.18). Because the Einstein equation is a covariant equation, we can always write the perturbative equations in terms of gauge invariants. To see how the gauge invariants come about consider the action of (8.18) on each mode,

$$\begin{aligned} &\phi \to \phi + \dot{\epsilon}, \qquad \beta \to \beta + \dot{\xi} - \epsilon, \qquad \alpha \to \alpha + \xi, \qquad \psi \to \psi + \tfrac{\Delta}{3}\xi && (8.27) \\ &\beta_i \to \beta_i + \dot{\xi}_i, \qquad \nu_i \to \nu_i + \xi_i && (8.28) \\ &\tilde{h}_{ij} \to \tilde{h}_{ij} && (8.29) \end{aligned}$$

We immediately see that the tensor mode is gauge invariant (as it always is!), and it is straightforward to check that the combinations given by Φ, Ψ and B_i are also gauge invariant,

$$\Phi \to \Phi, \qquad \Psi \to \Psi, \qquad B_i \to B_i$$

To solve the system, it suffices to solve for the gauge invariants. The tensor equation permits wave solutions for the tensor mode, $\tilde{h}_{ij}$. These are gravitational waves. We are currently on the lookout for evidence of these in Nature, through gravitational wave experiments such as LIGO.

Now consider the vectors. Assuming asymptotically vanishing boundary conditions, equations (8.23) and (8.24) suggest that there can be no vector modes, $B_i = 0$.

Finally, we consider the scalars. The traceless part of (8.21) implies that

$$\Phi + \Psi = 0$$

where we have once again assumed asymptotically vanishing boundary conditions. This result is typical of GR. It is violated in some modified theories of gravity and is often used as a means to look for experimental evidence of modified gravity. The remaining scalar equations combine to give, $\Phi = \Phi(x)$ (ie independent of time), and

$$\Delta\Phi = 4\pi G_N \rho$$

This is just Poisson's equation (7.2). How remarkable! Linearised gravity on Minkowski space with a non-relativistic source reproduces the fundamental equation of Newtonian gravity from first principles. You've got to admit, Einstein really was a genius.

Finally, we note that the scalar mode is not freely propagating (there is no wave component). In GR there are only two freely propagating degrees of freedom in four dimensions, and these are contained within the gravitational waves.

Exercises

1. If the metric is given by $g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}$, where h is small, prove that the inverse metric is given by $g^{\mu\nu} = \bar{g}^{\mu\nu} - h^{\mu\nu}$ to leading order.
2. Given the form of the perturbed Riemann tensor (8.4), show that the perturbed Ricci tensor is given by
$$\delta R_{\mu\nu} = -\frac{1}{2}\Delta_L h_{\mu\nu}$$
where the Lichnerowicz operator,
$$\Delta_L h_{\mu\nu} = \bar{\Box} h_{\mu\nu} - 2\bar{\nabla}_{(\mu}\bar{\nabla}^{\alpha}\bar{h}_{\nu)\alpha} - 2\bar{R}_{\alpha(\mu}h^{\alpha}_{\nu)} + 2\bar{R}_{\mu\alpha\nu\beta}h^{\alpha\beta}$$
Hint: you will need to commute some covariant derivatives.
3. Prove that the perturbed Ricci scalar is given by
$$\delta R = \left(\bar{\nabla}^{\mu}\bar{\nabla}^{\nu} - \bar{g}^{\mu\nu}\bar{\Box} - \bar{R}^{\mu\nu}\right) h_{\mu\nu}$$
Hence show that if $h_{\mu\nu} \to h_{\mu\nu} + \bar{\nabla}_{\mu}\xi_{\nu} + \bar{\nabla}_{\nu}\xi_{\mu}$ that $\delta R \to \delta R + \mathcal{L}_{\xi}\bar{R}$.
4. Prove that
$$\mathcal{L}_{\xi}\bar{g}_{\mu\nu} = \bar{\nabla}_{\mu}\xi_{\nu} + \bar{\nabla}_{\nu}\xi_{\mu}$$
Hint: use Riemann normal coordinates

5. Prove that on a maximally symmetric space with cosmological constant Λ, the Lichnerowicz operator is given by

$$\Delta_L h_{\mu\nu} = \bar{\Box} h_{\mu\nu} - 2\bar{\nabla}_{(\mu}\bar{\nabla}^{\alpha}\bar{h}_{\nu)\alpha} - \frac{4\Lambda}{3}\left(h_{\mu\nu} - \frac{1}{4}h g_{\mu\nu}\right)$$

6. Derive the transformation laws (8.27) to (8.29) and hence show that the following quantities are gauge invariant,

$$\Phi = \phi + \dot{\beta} - \ddot{\alpha}, \qquad \Psi = \psi - \frac{\Delta}{3}\alpha, \qquad B_i = \beta_i - \dot{\nu}_i$$

9 The gravitational action

As with any theory we would like to be able to derive the Einstein's equations from an action principle. Treating the metric as our fundamental field, we therefore wish to derive an action, $S = S[g_{\mu\nu}, \ldots]$ for which

$$\frac{\delta S}{\delta g^{\mu\nu}} = 0 \implies G_{\mu\nu} = 8\pi G_N T_{\mu\nu}$$

9.1 The measure

Now the first thing to realise is that d^4x is not a suitable measure. This is because it is not covariant, transforming under a general coordinate transformation $x^\mu \to \hat{x}^\mu(x)$ as

$$d^4x \to d^4\hat{x} = |\det J^\mu_\nu| d^4x \tag{9.1}$$

where the Jacobian matrix

$$J^\mu_\nu = \frac{\partial \hat{x}^\mu}{\partial x^\nu}$$

However, we also know that the metric transforms such that

$$g_{\mu\nu}(\hat{x}) = \frac{\partial \hat{x}^\alpha}{\partial x^\mu}\frac{\partial \hat{x}^\beta}{\partial x^\nu}\hat{g}_{\alpha\beta}(\hat{x}) = J^\alpha_\mu J^\beta_\nu \hat{g}_{\alpha\beta}(\hat{x})$$

and therefore

$$\det g_{\mu\nu} = |\det J^\mu_\nu|^2 \det \hat{g}_{\mu\nu}$$

Combining this with (9.1) we immediately see that

$$d^4x\sqrt{-\det g_{\mu\nu}} = d^4\hat{x}\sqrt{-\det \hat{g}_{\mu\nu}}$$

In other words, the quantity $d^4x\sqrt{-\det g_{\mu\nu}}$ transforms as a scalar under a general coordinate transformation, and is therefore a suitable measure. We often use the following shorthand for the determinant of the metric $g = \det g_{\mu\nu}$, and the covariant measure is written as

$$d^4x\sqrt{-g} \tag{9.2}$$

This generalises to any dimension and any signature as $d^n x\sqrt{|g|}$. Furthermore, the "volume" of a manifold M is now given by $V = \int_M d^n x\sqrt{|g|}$. Let us consider an illustrative example. Take three dimensional Euclidean space written in a spherically symmetric coordinate system

$$ds^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)$$

The measure now takes the form

$$d^4x\sqrt{|g|} = drd\theta d\phi r^2 \sin\theta$$

which we recognise as the appropriate measure used for integration with spherical polars.

9.2 The Einstein-Hilbert action

Now that we have a covariant measure, we seek a scalar Lagrangian density. The simplest scalar we can think of that is built out of the curvature is the Ricci scalar. We therefore propose the following action

$$S = S_{EH} + S_m \tag{9.3}$$

where

$$S_{EH}[g_{\mu\nu}] = \frac{1}{16\pi G_N}\int d^4x\sqrt{-g}R \tag{9.4}$$

is known as the *Einstein-Hilbert action*, and $S_m = S_m[g_{\mu\nu};\Psi_n]$ is the action for the matter fields, Ψ_n, coupled to gravity.

Let us consider the variation of the Einstein-Hilbert action. Well,

$$\delta S_{EH} = \frac{1}{16\pi G_N}\int d^4x\delta\sqrt{-g}R + \sqrt{-g}\delta R \tag{9.5}$$

The variation of the metric determinant is easily obtained using the general relation, $\det X = \exp(tr\log X)$. It follows that

$$\delta\sqrt{-g} = \frac{1}{2}\sqrt{-g}g^{\mu\nu}\delta g_{\mu\nu}$$

Furthermore, since $\delta g_{\mu\nu} = -g_{\mu\alpha}g_{\mu\beta}\delta g^{\alpha\beta}$, we also have that

$$\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}g^{\mu\nu}\delta g^{\mu\nu}$$

The variation of the Ricci scalar was derived in Exercise 3 of the previous chapter. We can use that result to infer the following

$$\delta R = \left(\nabla^\mu\nabla^\nu - g^{\mu\nu}\Box - R^{\mu\nu}\right)\delta g_{\mu\nu} = \left(-\nabla_\mu\nabla_\nu + g_{\mu\nu}\Box + R_{\mu\nu}\right)\delta g^{\mu\nu}$$

Plugging all of this in we find that

$$\begin{aligned}\delta S_{EH} &= \frac{1}{16\pi G_N}\int d^4x\sqrt{-g}\left(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right)\delta g^{\mu\nu} - \sqrt{-g}(\nabla_\mu\nabla_\nu - g_{\mu\nu}\Box)\delta g^{\mu\nu} \\ &= \frac{1}{16\pi G_N}\int d^3x\sqrt{-g}G_{\mu\nu}\delta g^{\mu\nu} - \sqrt{-g}\nabla_\alpha\left[(\delta^\alpha_{(\mu}\nabla_{\nu)} - g_{\mu\nu}\nabla^\alpha)\delta g^{\mu\nu}\right] \end{aligned} \tag{9.6}$$

What do we do with the last term? Fortunately there exists the following useful relation,

$$\nabla_\mu F^\mu = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}F^\mu) \tag{9.7}$$

and we recognise the last term in (9.6) as a total derivative that may be discarded. Returning to the full action, we now have that

$$\frac{\delta S}{\delta g^{\mu\nu}} = \frac{1}{16\pi G_N}\sqrt{-g}\left(G_{\mu\nu} + \frac{16\pi G_N}{\sqrt{-g}}\frac{\delta S_m}{\delta g^{\mu\nu}}\right) \tag{9.8}$$

Therefore, if we identify the stress-energy tensor with,

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_m}{\delta g^{\mu\nu}} \tag{9.9}$$

then we do indeed have that

$$\frac{\delta S}{\delta g^{\mu\nu}} = 0 \implies G_{\mu\nu} = 8\pi G_N T_{\mu\nu}$$

The Einstein-Hilbert action has the unusual feature that one can treat the metric and the connection as independent fields and still obtain the same dynamics. Variation with respect to the metric yields the Einstein equations while variation with respect to the connection forces the Levi-Civita connection dynamically. This procedure is known as the Palatini variation. For further details see Wald p454. A word of warning though. The Einstein-Hilbert action is rather special in having this property. In a generic alternative gravity theory, changing the variational principle will change the dynamics.

9.3 The matter action

We now turn our attention to the matter action. The simplest source one can imagine is the vacuum energy, or cosmological constant. This makes a constant contribution to the matter action, of the form

$$S_m^{\text{vac}} = -\sigma\int d^4x\sqrt{-g} \implies T_{\mu\nu}^{\text{vac}} = -\sigma g_{\mu\nu} \tag{9.10}$$

Now consider a scalar field. In flat space, a canonical scalar field with a potential $V(\phi)$ has an action

$$S = \int d^4x - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi)$$

There are many ways in which we might promote this to curved space. The minimal approach involves introducing the covariant measure and promoting partial derivatives into covariant derivatives

$$S_m^{\text{scalar}} = \int d^4x\sqrt{-g}\left[-\frac{1}{2}\nabla_\mu\phi\nabla^\mu\phi - V(\phi)\right]$$
$$\implies T_{\mu\nu}^{\text{scalar}} = \nabla_\mu\phi\nabla_\nu\phi - g_{\mu\nu}\left[\frac{1}{2}(\nabla_\alpha\phi)^2 + V(\phi)\right] \tag{9.11}$$

We say that the scalar field is *minimally coupled* to gravity. One can consider more exotic scenarios in which the scalar couples to the curvature at the level of the action, eg terms like $\sqrt{-g}G_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi$ or $\sqrt{-g}R\phi^2$. This is called non-minimal coupling.

What about Maxwell fields? Again, in flat space, these are described by the following action,

$$S = -\frac{1}{4}\int d^4x F_{\mu\nu}F^{\mu\nu}$$

where the field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. To promote this to curved space we once again take the minimal approach, introducing the covariant measure and promoting partial derivatives into covariant derivatives

$$S_m^{\text{Maxwell}} = -\frac{1}{4}\int d^4x\sqrt{-g}F_{\mu\nu}F^{\mu\nu}, \qquad F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$$

$$\implies T_{\mu\nu}^{\text{Maxwell}} = F_{\mu\alpha}F_\nu{}^\alpha - \frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} \quad (9.12)$$

Actually, the connection drops out of the defintiion of the field strength due to the antisymmetrization. Also. note that the trace of the energy-momentum tensor is zero. This is an important property of the Maxwell field in four dimensions, and is related to the fact that it is a conformally invariant theory. One can also consider non-minimally coupled Maxwell theory.

Fermions are a little trickier. In flat space a fermion of mass, m, is described by an action,

$$S = -\int d^4x\bar{\psi}(\gamma^a\partial_a + m)\psi$$

where $\bar{\psi} = i\psi^\dagger\gamma^0$ and γ^a are the Gamma matrices. To promote this to curved space we need to define the covariant derivative acting on fermions

$$\nabla_\mu\psi = \partial_\mu\psi - \Gamma_\mu\psi$$

where $\Gamma_\mu = \frac{1}{4}\gamma_a\gamma_b g^{\alpha\beta}e^a{}_\alpha\nabla_\mu e^b{}_\beta$, and $\nabla_\mu e^b{}_\nu = \partial_\mu e^b{}_\nu - \Gamma^\alpha{}_{\mu\nu}e^b{}_\alpha$. The fermion action becomes

$$S_m^{\text{fermion}} = -\int d^4x\sqrt{-g}\bar{\psi}(\gamma^a e_a{}^\mu\nabla_\mu + m)\psi \quad (9.13)$$

For further details see Parker & Toms

9.4 The Gibbons-Hawking term

We now return to the Einstein-Hilbert action. Recall that in varying this action, we threw away a total derivative term of the form,

$$-\frac{1}{16\pi G_N}\int d^4x\sqrt{-g}\nabla_\alpha(\left[\delta^\alpha_{(\mu}\nabla_{\nu)} - g_{\mu\nu}\nabla^\alpha)\delta g^{\mu\nu}\right] = -\frac{1}{16\pi G_N}\int d^4x\partial_\alpha\left[\sqrt{-g}(\delta^\alpha_{(\mu}\nabla_{\nu)} - g_{\mu\nu}\nabla^\alpha)\delta g^{\mu\nu}\right]$$

What happens if the spacetime manifold, M has a non-trivial boundary ∂M? Then, taking n_μ to be the (spacelike) outward pointing unit normal to the boundary, we pick up a boundary term,

$$\begin{aligned}
\delta S_{EH}^{boundary} &= -\frac{1}{16\pi G_N}\int_M d^4x\partial_\alpha\left[\sqrt{-g}(\delta^\alpha_{(\mu}\nabla_{\nu)} - g_{\mu\nu}\nabla^\alpha)\delta g^{\mu\nu}\right] \\
&= -\frac{1}{16\pi G_N}\int_{\partial M} d^3x\sqrt{-\gamma}n_\alpha(\delta^\alpha_{(\mu}\nabla_{\nu)} - g_{\mu\nu}\nabla^\alpha)\delta g^{\mu\nu} \\
&= -\frac{1}{16\pi G_N}\int_{\partial M} d^3x\sqrt{-\gamma}(n_{(\mu}\nabla_{\nu)} - g_{\mu\nu}n^\alpha\nabla_\alpha)\delta g^{\mu\nu} \\
&= \frac{1}{16\pi G_N}\int_{\partial M} d^3x\sqrt{-\gamma}(n_{(\mu}\gamma^\alpha_{\nu)}\nabla_\alpha - \gamma_{\mu\nu}n^\alpha\nabla_\alpha)\delta g^{\mu\nu} \\
&= \frac{1}{16\pi G_N}\int_{\partial M} d^3x\sqrt{-\gamma}(n^{(\mu}\gamma^{\nu)\alpha} - \gamma^{\mu\nu}n^\alpha)\nabla_\alpha\delta g_{\mu\nu}
\end{aligned}$$

Here $\gamma_{\mu\nu} = g_{\mu\nu} - n_\mu n_\nu$ is the *induced metric* on the boundary (this just means the metric for the boundary manifold). It corresponds to a projection operator that projects out the components normal to the boundary. We have also used the fact that

$$\nabla_\nu = \gamma^\alpha_\nu\nabla_\alpha + n_\nu n^\alpha\nabla_\alpha$$

as well as the relation

$$\sqrt{-g}|_{\partial M} = \sqrt{-\gamma} \tag{9.14}$$

This is easily proven using *Gaussian-Normal coordinates* adapted to the boundary. Gaussian-Normal coordinates are defined as $x^\mu = (z, x^i)$, with the boundary at $z =$ constant, and

$$g_{zz} = 1, \qquad g_{zi} = 0$$

Then the unit normal to the boundary is given by $n_z = 1,\ n_i = 0$, and we have

$$\gamma_{zz} = 1, \qquad \gamma_{zi} = 0, \qquad \gamma_{ij} = g_{ij}$$

The result (9.14) now follows trivially.

The boundary term, $\delta S_{EH}^{boundary}$, does not vanish with respect to Dirichilet boundary conditions in which the metric is taken to be fixed on the boundary, owing to the presence of the normal derivative. This fact has been noticed by Gary Gibbons and Stephen Hawking, and they proposed adding the following boundary term to the Einstein-Hilbert action[9],

$$S_{GH} = \frac{1}{8\pi G_N}\int_{\partial M} d^3x\sqrt{-\gamma}K \tag{9.15}$$

This is known as the *Gibbons-Hawking* term. $K = \gamma^{\mu\nu}K_{\mu\nu}$ is the trace of the extrinsic curvature of the boundary. The extrinsic curvature is defined as the Lie derivative of the induced metric with respect to the outward pointing normal,

$$K_{\mu\nu} = \frac{1}{2}\mathcal{L}_n\gamma_{\mu\nu} \tag{9.16}$$

[9] For spacelike boundaries (initial and final surfaces) this term was proposed a little earlier, by York

In Gauss-Normal coordinates it is easy to convince yourself that

$$K_{zz} = 0, \qquad K_{zi} = 0, \qquad K_{ij} = \frac{1}{2}\partial_z g_{ij}$$

To see why the Gibbons-Hawking term helps, we first note that

$$\delta S_{GH} = \frac{1}{8\pi G_N}\int_{\partial M} d^3x\sqrt{-\gamma}\left(K_{\mu\nu} - K\gamma_{\mu\nu}\right)\delta\gamma^{\mu\nu} + \sqrt{-\gamma}\gamma^{\mu\nu}\left(\delta K_{\mu\nu} - \frac{1}{2}K\delta\gamma_{\mu\nu}\right) \tag{9.17}$$

To proceed, we work exclusively in Gaussian-Normal coordinates, noting that

$$\begin{aligned}
&\gamma^{\mu\nu}\left(\delta K_{\mu\nu} - \frac{1}{2}K\delta\gamma_{\mu\nu}\right) + \frac{1}{2}(n^{(\mu}\gamma^{\nu)\alpha} - \gamma^{\mu\nu}n^\alpha)\nabla_\alpha\delta g_{\mu\nu} \\
&\quad = \frac{1}{2}g^{ij}\left(\partial_z\delta g_{ij} - \frac{1}{2}\delta g_{ij}g^{kl}\partial_z g_{kl}\right) - \frac{1}{2}g^{ij}\partial_z\delta g_{ij} + \frac{1}{2}\delta g_{ij}g^{jk}\Gamma^i{}_{jz} \\
&\quad = -\frac{1}{4}\delta g_{ij}g^{ij}g^{kl}\partial_z g_{kl} + \frac{1}{4}\delta g_{ij}g^{jk}g^{il}\partial_z g_{lj} \\
&\quad = \frac{1}{2}\delta g_{ij}\left(K^{ij} - Kg^{ij}\right) \\
&\quad = \frac{1}{2}\delta\gamma_{\mu\nu}\left(K^{\mu\nu} - K\gamma^{\mu\nu}\right) \\
&\quad = -\frac{1}{2}\left(K_{\mu\nu} - K\gamma_{\mu\nu}\right)\delta\gamma^{\mu\nu}
\end{aligned}$$

Plugging all this together, we see that,

$$\delta S_{EH}^{boundary} + \delta S_{GH} = \frac{1}{4\pi G_N}\int_{\partial M} d^3x\sqrt{-\gamma}\left(K_{\mu\nu} - K\gamma_{\mu\nu}\right)\delta\gamma^{\mu\nu} \tag{9.18}$$

The normal derivatives have dropped out completely. The full gravitational action

$$S_{grav} = \frac{1}{16\pi G_N}\int_M d^4x\sqrt{-g}R + \frac{1}{8\pi G_N}\int_{\partial M} d^3x\sqrt{-\gamma}K \tag{9.19}$$

is well defined for Dirichilet boundary conditions, such that

$$\delta S_{grav} = \frac{1}{16\pi G_N}\int_M d^4x\sqrt{-g}G_{\mu\nu}\delta g^{\mu\nu} + \frac{1}{4\pi G_N}\int_{\partial M} d^3x\sqrt{-\gamma}\left(K_{\mu\nu} - K\gamma_{\mu\nu}\right)\delta\gamma^{\mu\nu} \tag{9.20}$$

Exercises

1. Prove the following relation

$$\nabla_\mu F^\mu = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}F^\mu)$$

Hint: you should make use of equation (8.3) and the fact that $\delta\sqrt{-g} = \frac{1}{2}\sqrt{-g}g^{\mu\nu}\delta g_{\mu\nu}$.

2. Prove each of the following results:
   (a) $S_m^{\text{scalar}} = \int d^4x\sqrt{-g}\left[-\frac{1}{2}\nabla_\mu\phi\nabla^\mu\phi - V(\phi)\right] \implies T_{\mu\nu}^{\text{scalar}} = \nabla_\mu\phi\nabla_\nu\phi - g_{\mu\nu}\left[\frac{1}{2}(\nabla_\alpha\phi)^2 + V(\phi)\right]$
   (b) $S_m^{\text{Maxwell}} = -\frac{1}{4}\int d^4x\sqrt{-g}F_{\mu\nu}F^{\mu\nu} \implies T_{\mu\nu}^{\text{Maxwell}} = F_{\mu\alpha}F_\nu{}^\alpha - \frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}$
3. Prove that in Gauss-Normal coordinates (z, x^i) adapted to a surface with $z =$ constant, we have
$$K_{zz} = 0, \qquad K_{zi} = 0, \qquad K_{ij} = \frac{1}{2}\partial_z g_{ij}$$

10 Brans-Dicke gravity

Whilst General Relativity is a remarkable theory, it does have its limitations as we saw in chapter 3. Furthermore a theory is only a great theory *in comparision* to other theories. With all this is mind it is well worth considering alternatives to GR.

Perhaps the most well known alternative to GR is a theory proposed by Brans and Dicke. Motivated by Dirac's large number hypothesis[10], they promoted Newton's constant to a field and proposed a theory described by the following action,

$$S_{BD} = \frac{1}{16\pi}\int d^4x\sqrt{-g}\left[\phi R - \frac{w}{\phi}(\nabla\phi)^2\right] + S_m[g_{\mu\nu};\Psi_n] \tag{10.1}$$

The field equations for this theory are given by

$$\begin{aligned} \phi G_{\mu\nu} - \nabla_\mu\nabla_\nu\phi - \frac{w}{\phi}\nabla_\mu\phi\nabla_\nu\phi + g_{\mu\nu}\left(\Box\phi + \frac{w}{2\phi}(\nabla\phi)^2\right) &= 8\pi T_{\mu\nu} & (10.2)\\ (2w+3)\Box\phi &= 8\pi T & (10.3)\end{aligned}$$

One of the features of Brans-Dicke theory is that we can perform a conformal transformation of the metric so that its looks more like the Einstein-Hilbert action. To this end we make the following field redefinitions

$$g_{\mu\nu} = e^{2\alpha\phi}\tilde{g}_{\mu\nu}, \qquad \phi = e^{-2\alpha\psi}, \qquad \alpha = \sqrt{\frac{4\pi}{2w+3}} \tag{10.4}$$

The action then takes on the more familiar form of GR + a minimally coupled scalar,

$$S_{BD} = \frac{1}{16\pi}\int d^4x\sqrt{-\tilde{g}}\left[\tilde{R} - \frac{1}{2}(\tilde{\nabla}\psi)^2\right] + S_m[e^{2\alpha\psi}\tilde{g}_{\mu\nu};\Psi_n] \tag{10.5}$$

Because of this, the metric $\tilde{g}_{\mu\nu}$ is referred to as the *Einstein frame* metric, while its cousin, $g_{\mu\nu}$, is referred to as the *Jordan frame* metric. Classically, all physical observables are completely independent of the choice of frame, as of course they must be.

In the Einstein frame, we explicitly see that matter now couples to the scalar, ϕ as well as the new metric $\tilde{g}_{\mu\nu}$. This means we have a fifth force mediated by ϕ acting on standard model fields, and results in the violation of the strong equivalence principle. The strength of the force is controlled by α. Fifth force tests constrain α to be very small, or equivalently, w to be very large. The current limit has $w \gtrsim 40000$.

[10] Dirac had noticed that $G_N \propto 1/t_U$ and $M_U \propto t_U^2$ where t_U is the age of the Universe and M_U is its mass.

10.1 Weak Brans-Dicke gravity

Let us now consider linearised theory in Brans-Dicke gravity. We will work in the Jordan frame and begin by rewriting the field equations as

$$R_{\mu\nu} - \frac{\nabla_\mu \nabla_\nu \phi}{\phi} = \frac{8\pi}{\phi}\left[T_{\mu\nu} - \left(\frac{w+1}{2w+3}\right) T g_{\mu\nu}\right] + \frac{w}{\phi^2}\nabla_\mu\phi\nabla_\nu\phi \tag{10.6}$$

$$(2w+3)\Box\phi = 8\pi T \tag{10.7}$$

We now perform linearised perturbation theory about a Minkowski vacuum with constant $\phi = \phi_0$. In other words we take,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad \phi = \phi_0 + \delta\phi, \qquad |h_{\mu\nu}|, |\delta\phi|, |T_{\mu\nu}| \ll 1$$

This gives

$$\delta R_{\mu\nu} - \frac{\partial_\mu \partial_\nu \delta\phi}{\phi_0} = \frac{8\pi}{\phi_0}\left[T_{\mu\nu} - \left(\frac{w+1}{2w+3}\right) T \eta_{\mu\nu}\right] \tag{10.8}$$

By choosing a gauge $\partial^a \bar{h}_{\alpha\nu} = \frac{\partial_\nu \delta\phi}{\phi_0}$, we arrive at the following result,

$$h_{\mu\nu} = -\frac{16\pi}{\phi_0 \bar{\Box}}\left[T_{\mu\nu} - \left(\frac{w+1}{2w+3}\right) T \eta_{\mu\nu}\right] \tag{10.9}$$

This should be compared with the equivalent result in General Relativity given by equation (8.15). Furthermore, if the gravitational field is sourced by stress-energy, $\tau^s_{\mu\nu}$, then the gravitational potential energy of a probe of stress-energy, $\tau^p_{\mu\nu}$ is given by

$$V = -\frac{1}{2}\int d^3x h^s_{\mu\nu}(x)\tau_p^{\mu\nu}(x) = \frac{8\pi}{\phi_0}\int d^3x \left[\tau_p^{\mu\nu}(x)\frac{1}{\bar{\Box}}\tau^s_{\mu\nu}(x) - \left(\frac{w+1}{2w+3}\right)\tau_p(x)\frac{1}{\bar{\Box}}\tau^s(x)\right] \tag{10.10}$$

Again, this should be compared with corresponding result in GR (8.16). In fact, let us be a little more precise. We know that GR gives a very accurate prediction to planetary orbits and for light bending in the solar system. The Sun can be treated as a non-relativistic source with $\tau^s_{\mu\nu} = \rho_s \delta^t_\mu \delta^t_\nu$. For light bending, the probe is a photon whose stress-energy tensor is traceless $\tau_p = 0$. The GR result gives

$$V_{GR}^{\text{photon}} = 8\pi G_N \int d^3x \tau_p^{\mu\nu}(x)\frac{1}{\bar{\Box}}\tau^s_{\mu\nu}(x) \tag{10.11}$$

while the BD result gives

$$V_{GR}^{\text{photon}} = \frac{8\pi}{\phi_0}\int d^3x \tau_p^{\mu\nu}(x)\frac{1}{\bar{\Box}}\tau^s_{\mu\nu}(x) \tag{10.12}$$

We can match the two by simply taking $\phi_0 = 1/G_N$.

Now consider the planets. These are also non-relativistic and have $\tau^p_{\mu\nu} = \rho_p \delta^t_\mu \delta^t_\nu$. In GR we have

$$V_{GR}^{\text{planet}} = -8\pi G_N \int d^3x \rho_p(x)\frac{1}{\bar{\Box}}\rho_s(x) \tag{10.13}$$

whereas in BD gravity we have

$$V_{BD}^{\rm planet} = -\frac{16\pi}{\phi_0}\left(\frac{w+2}{2w+3}\right)\int d^3x \rho_p(x)\frac{1}{\square}\rho_s(x) \tag{10.14}$$

As we have already set $\phi_0 = 1/G_N$ in order to be compatible with light bending, we note that

$$\frac{V_{BD}^{\rm planet} - V_{GR}^{\rm planet}}{V_{GR}^{\rm planet}} = \frac{1}{2w+3} \tag{10.15}$$

So unless w is very large we see that the theory deviates considerably from GR. Of course, we could have chosen ϕ_0 differently so that the planetary results matched, but then we would have run into problems with light bending. The deviation from GR is greatest as $w \to -3/2$. This is the strong coupling limit of Brans-Dicke gravity when matter couples to the scalar with infinite strength.

10.2 Chameleons

We have seen that fifth force tests constrain the Brans Dicke scalar to be very weakly coupled to matter. However, this is not the only way to screen a fifth force. We can tolerate a much stronger matter coupling provided we introduce a mass for the scalar. If the mass exceeds meV then the range of the fifth force is pushed down below $\frac{\hbar c}{\rm meV} \sim 0.2$ mm, making it inaccessible to table top tests of gravity.

Quintessence scenarios correspond to the BD theories in the presence of a quintessence potential, $V(\psi)$

$$S_\psi = \int d^4x\sqrt{-\tilde{g}}\left[\frac{M_{pl}^2}{2}\tilde{R} - \frac{1}{2}(\tilde{\nabla}\psi)^2 - V(\psi)\right] + S_m[e^{2\alpha\psi/M_{pl}}\tilde{g}_{\mu\nu};\Psi_n] \tag{10.16}$$

where we have explicitly included some factors of M_{pl} to make contact with the literature. Note that matter is minimally coupled to the Jordan frame metric $g_{\mu\nu} = e^{2\alpha\psi/M_{pl}}\tilde{g}_{\mu\nu}$.

In order for the quintessence field to mimic dark energy the potential must be sufficiently flat, and in particular, the mass $m^2(\psi) = V''(\psi) \lesssim H_0^2$ where $H_0 \sim 10^{-33}$ eV is the current Hubble scale. Such a light mass would allow a long range force to be mediate and one is tempted to conclude that the coupling to matter should be extremely weak in order pass fifth force constraints. However, this statement is too quick. One should not neglect the role played by the environment. Environmental effects alter the form of the *effective* potential and it is possible for the field to pick up a large mass in a dense region of matter.

To see, we note that the equations of motion are given by

$$M_{pl}^2\tilde{G}_{\mu\nu} = \partial_\mu\psi\partial_\nu\psi - \left[\frac{1}{2}(\partial\psi)^2 + V\right]g_{\mu\nu} + \tilde{T}_{\mu\nu}, \tag{10.17}$$

$$\nabla^2\psi = V'(\psi) - \frac{\alpha}{M_{pl}}\tilde{T}^\mu_\mu, \tag{10.18}$$

where $\tilde{T}_{\mu\nu} = \frac{2}{\sqrt{-\tilde{g}}} \frac{\partial S_m}{\partial \tilde{g}^{\mu\nu}}$ is the energy momentum tensor in the Einstein frame. This is related to the energy momentum tensor in the Jordan frame, $T_{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\partial S_m}{\partial g^{\mu\nu}}$, as follows

$$\tilde{T}_{\mu\nu} = e^{2\alpha\psi/M_{pl}} T_{\mu\nu}. \tag{10.19}$$

For our purposes, it is sufficient to approximate the geometry in the Einstein frame as Minkowski space, $\tilde{g}_{\mu\nu} \approx \eta_{\mu\nu}$. We will also assume a non-relativistic matter distribution, so that $g^{\mu\nu} T_{\mu\nu} \approx -\rho$. It is convenient to use the energy density $\tilde{\rho} = \rho e^{3\alpha\psi/M_{pl}}$, which is conserved in the Einstein frame[11] , so that equation (10.18) now gives

$$\nabla^2 \psi = V'(\psi) + \frac{\alpha}{M_{pl}} \tilde{\rho} e^{\alpha\psi/M_{pl}}. \tag{10.20}$$

Now let us define the effective potential

$$V_{eff}(\psi) = V(\psi) + \tilde{\rho} e^{\alpha\psi/M_{pl}}, \tag{10.21}$$

so that the scalar equation of motion (10.20) now reads $\nabla^2\psi = V'_{eff}(\psi)$. In a stationary, homogeneous distribution of matter, the field ψ will sit at the minimum of the *effective* potential, ψ_*, where $V'_{eff}(\psi_*) = 0$. In general, the field now picks up an *effective* mass, given by $m^2_{eff} = V''_{eff}(\psi_*)$. However, the depth and position of this minumum changes over with the sourounding density of matter, hence providing of a mechanism to have light fields on cosmological scales but heavy near the Earth. In the space of solutions, there are configurations of the field which interpolate between the cosmological minimum and the Earth's one. We call these *chameleons* because they change their character depending on their environment.

10.3 $f(R)$ gravity

Another popular alternative to GR is $f(R)$ gravity, describe by the following action

$$S = \int d^4x \sqrt{-g} f(R) + S_m[g_{\mu\nu}; \Psi_n] \tag{10.22}$$

where f is some function. This is entirely equivalent to a BD theory with $w = 0$ and a potential. To show this we simply rewrite the action using a Lagrange multiplier, ϕ and an auxiliary field λ,

$$S = \int d^4x \sqrt{-g} \left[f(\lambda) + \phi(R - \lambda) \right] + S_m[g_{\mu\nu}; \Psi_n] \tag{10.23}$$

The λ equation of motion gives, $\phi = f'(\lambda)$. Plugging this back in we get

$$S = \int d^4x \sqrt{-g} \left[\phi R - V(\phi) \right] + S_m[g_{\mu\nu}; \Psi_n] \tag{10.24}$$

where

$$V(\phi) = \phi\lambda(\phi) - f(\lambda(\phi))$$

and $\lambda(\phi)$ is given by inverting the relation

$$\phi = f'(\lambda)$$

[11] To see this we note that $\rho \times$ spatial volume = constant, and that spatial volume scales like $(e^{\alpha\psi/M_{pl}})^3$.

11 Kaluza Klein gravity

Kaluza-Klein (KK) theory grew out of an attempt to unify gravity and electrodynamics. The basic idea was to consider General Relativity on a $4+1$ dimensional manifold where one of the spatial dimensions was taken to be small and compact. One can perform a harmonic expansion of all fields along the extra dimension, and compute an effective $3+1$ dimensional theory by integrating out the heavy modes. This idea has been embraced by string theorists who compactify 10 dimensional string theories and 11 dimensional supergravity/Mtheory on compact manifolds of 6 or 7 dimensions respectively, often "switching on fluxes" and "wrapping branes" on the compact space. Each different compactification gives a different effective 4-dimensional theory, so much so that we now talk about an entire landscape of effective theories.

To understand the generic features of KK compactifications, it is sufficient to describe the dimensional reduction of General Relativity on a circle, S^1. We first define General Relativity in $D = d+1$ dimensions, via the generalised Einstein-Hilbert action

$$S[\gamma] = \frac{1}{16\pi G_D} \int d^D X \sqrt{-\gamma} \mathcal{R} \tag{11.1}$$

where G_D is Newton's constant in D dimensions, γ_{AB} is the D dimensional metric with corresponding Ricci tensor, $\mathcal{R}_{AB}$ and Ricci scalar, $\mathcal{R} = \gamma^{AB}\mathcal{R}_{AB}$. Note that we are neglecting the matter Lagrangian for brevity. We are assuming that one of the spatial dimensions is compactified on a circle of radius $L/2\pi$. To this end we can define coordinates $X^A = (x^\mu, z)$, where the coordinate z lies along the compact direction, such that $0 \leq z < L$

We can expand the metric as a Fourier series of the form

$$\gamma_{AB}(x, z) = \sum_n \gamma_{AB}^{(n)}(x) e^{inz/L}. \tag{11.2}$$

We find that we get an infinite number of fields in d dimensions. Modes with $n \neq 0$ correspond to massive fields with mass $|n|/L$, whereas the zero mode corresponds to a massless field. As we take L to be smaller and smaller we see that the mass of the first massive field becomes very large. This means that if we compactify on a small enough circle we can truncate to massless modes in the 4-dimensional theory. Massive modes will only get excited by scattering processes whose energy lies at or above the compactification scale $1/L$. This also applies to matter fields arising in particle physics. Indeed, particle physics imposes by far the strongest constraints on the size of the extra dimension. Standard Model processes have been well tested with great precision down to distances of the order, $(\text{TeV})^{-1}$ with no evidence of extra dimensions emerging. Assuming that the extra dimensions are universal, that is the Standard Model fields can extend all the way into them, we infer that $L \lesssim 10^{-19}$ m. The natural scale of the compact dimensions is usually taken be Planckian, $L \sim l_{pl}$.

Let us now focus on the zero modes, $\gamma_{AB}^{(0)}(x)$. We could define $\gamma_{\mu\nu}^{(0)}$, $\gamma_{\mu z}^{(0)}$ and $\gamma_{zz}^{(0)}$ to be the d-dimensional fields $g_{\mu\nu}(x)$, $A_\mu(x)$ and $\phi(x)$. In effective field theory language, these will correspond to the metric, gauge field, and dilaton respectively. In order that our results are more transparent we will actually define the components of the metric in the following way:

$$\gamma^{(0)}_{\mu\nu} = e^{2\alpha\phi} g_{\mu\nu} + e^{2\beta\phi} A_\mu A_\nu, \qquad \gamma^{(0)}_{\mu z} = e^{2\beta\phi} A_\mu, \qquad \gamma^{(0)}_{zz} = e^{2\beta\phi}. \tag{11.3}$$

where $\alpha = 1/\sqrt{2(d-1)(d-2)}$ and $\beta = -(d-2)\alpha$. Since we have truncated to the massless fields, we can integrate out the z part of the action (11.2). We find that the d-dimensional effective action is given by

$$S_{\text{eff}}[g, A, \phi] = \frac{L}{16\pi G_D} \int d^d x \sqrt{-g} \left(R - \frac{1}{2}(\nabla\phi)^2 - \frac{1}{4} e^{-2(d-1)\alpha\phi} F^2 \right) \tag{11.4}$$

where $F^2 = F_{\mu\nu}F^{\mu\nu}$ and $F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$ is the electromagnetic field strength. The curvature associated with the d dimensional metric, $g_{\mu\nu}$, is described by the Ricci tensor, $R_{\mu\nu}$ and Ricci scalar, $R = g^{\mu\nu}R_{\mu\nu}$. What we now have is an Einstein-Maxwell-Dilaton system in d dimensions. Of course, Kaluza and Klein were particularly interested in the case of $d = 4$. They were frustrated by the presence of the dilaton, ϕ, in the resulting 4-dimensional effective theory. The point is that one cannot simply set the dilaton to zero and retain a non-trivial Maxwell field, since this would be in conflict with the field equations arising from (11.2),

$$G_{\mu\nu} = \frac{1}{2}\left[\nabla_\mu\phi\nabla_\nu\phi - \frac{1}{2}(\nabla\phi)^2 g_{\mu\nu} + e^{-2(d-1)\alpha\phi}\left(F_{\mu\alpha}F_\nu{}^\alpha - \frac{1}{4}F^2 g_{\mu\nu}\right)\right] \tag{11.5}$$

$$\nabla^\mu \left(e^{-2(d-1)\alpha\phi} F_{\mu\nu}\right) = 0 \tag{11.6}$$

$$\Box\phi = -\frac{1}{2}(d-1)\alpha e^{-2(d-1)\alpha\phi} F^2 \tag{11.7}$$

where $G_{\mu\nu}$ is the Einstein tensor in d dimensions. Making use of the appropriate jargon, we say that switching off the dilaton does not represent a *consistent truncation* of the higher dimensional theory. We should also note that the *physical* size of the compact dimension is not necessarily given by L but by $Le^{\beta\phi(x)}$. If L is to represent an accurate measure of the compactification scale, we are therefore implicitly assuming that ϕ is stabilised close to zero. For this to happen we need to generate a potential for ϕ that admits a stable solution– this is known as the problem of moduli stabilization.

12 Braneworld gravity

The braneworld paradigm represents a radical alternative to the standard Kaluza-Klein scenario discussed in the previous section. In the KK scenario, the extra dimensions must be small and compact, the size of the internal space constrained by collider experiments to be below the inverse TeV scale. In the braneworld scenario the extra dimensions can be much larger, perhaps even infinite in extent! This is made possible by relaxing the assumption of universal extra dimensions.

In the braneworld picture the Standard Model fields are *not* universal, rather they are confined to lie on a $3+1$ dimensional hypersurface, known as the *brane*, embedded in some higher dimensional spacetime, known as the *bulk*. Tests of Standard Model processes can only constrain how far the brane may extend into the bulk, or in other words, the brane thickness. They do not constrain the size of the bulk itself. These can only come from gravitational experiments, since gravity is the only force that can be mediated through the bulk spacetime. As is well known, on small scales gravity is much weaker than the other three fundamental forces. This makes it difficult to test at short distances. In fact, the gravitational interaction has only been probed down to ~ 0.1 mm, with torsion-balance tests of the inverse square law. It is too simplistic to suggest that this translates into an upper bound on the radius of the bulk. Gravity is intimately related to geometry, and one can even warp the bulk geometry such that an infinitely large extra dimension is still allowed by experiment.

Let us give an overview of the governiing equations in a generic five dimensional braneworld model (there are many!) Imagine we have a five dimensional bulk space split into a series of domains separated from one another by 3-branes[12]. The 3-branes may be thought of as the boundaries to the various domains. It is convenient to label each brane, Σ_i, by an index i, so that the action is given by

$$S = \int_{bulk} d^5x\sqrt{-\gamma}\left(\frac{M_5^3}{2}\mathcal{R} + \mathcal{L}_{bulk}\right) + \sum_{branes}\int_{brane} d^4x\sqrt{-g}\left[-\Delta\left(M_5^3 K\right) + \mathcal{L}_{brane}\right] \quad (12.1)$$

where γ_{ab} is the bulk metric with corresponding Ricci scalar, $\mathcal{R}$, M_5 is the bulk Planck scale, and $\mathcal{L}_{bulk}$ is the Lagrangian density describing the bulk field content. In principle both M_5 and $\mathcal{L}_{bulk}$ can vary from domain to domain. For each brane, $g_{\mu\nu}$ is the induced metric and $\mathcal{L}_{brane}$ is the Lagrangian density describing the field content on that particular brane. $K = g^{\mu\nu}K_{\mu\nu}$ is the trace of extrinsic curvature, $K_{\mu\nu}$. This should be evaluated on either side of the brane as it can differ from side to side. Labelling the two sides of a given brane using L and R, we define $K_{\mu\nu}|_{L,R} = \frac{1}{2}\mathcal{L}_{n|_{L,R}}g_{\mu\nu}$, ie the Lie derivative of the induced metric, with respect to the unit normal $n^a|_{L,R}$. The unit normal on both sides points from L to R. Note that what appears in the action is the jump[13]

$$\Delta\left(M_5^3 K\right) = M_5^3 K|_R - M_5^3 K|_L$$

This corresponds to the Gibbons-Hawking boundary term for the bulk domains on each side of the brane.

Now there are two (completely equivalent) ways to treat the brane contributions at the level of the field equations. One approach is to treat them as delta-function sources in the Einstein equations. However, our preferred approach is to explicitly separate the field equations in the bulk from the boundary conditions at the brane. Then the bulk equations of motion are given by the bulk Einstein equations

$$\mathcal{G}_{ab} = \mathcal{R}_{ab} - \frac{1}{2}\mathcal{R}\gamma_{ab} = \frac{1}{M_5^3}T^{bulk}_{ab} \quad (12.2)$$

[12] A p-brane is a timelike hypersurface of $p+1$ dimensions.

[13] We define the jump of any quantity Q across a brane as $\Delta Q = Q|_R - Q|_L$.

where $T^{bulk}_{ab} = -\frac{2}{\sqrt{-\gamma}}\frac{\delta}{\delta\gamma^{ab}}\int_{bulk} d^5x\sqrt{-\gamma}\mathcal{L}_{bulk}$ is the bulk energy momentum tensor. The boundary conditions at Σ_i are given by the Israel junction conditions

$$\Delta\left[M_5^2(K_{\mu\nu} - Kg_{\mu\nu})\right] = -T^{brane}_{\mu\nu} \tag{12.3}$$

where $T^{(i)}_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta}{\delta g^{\mu\nu}}\int_{brane} d^4x\sqrt{-g}\mathcal{L}_{brane}$ is the brane energy momentum tensor. The Israel equations can be obtained from variation of the action with respect to the induced metric on the brane.

12.1 Randall-Sundrum gravity

Randall-Sundrum gravity was originally proposed to explain the hierarchy between the Planck scale and the electoweak scale, but it soon developed into an alternative to compactification. In this model, we have up to two 3-branes separated by a section of five dimensional anti-de Sitter space. The branes are located at the boundaries $z = 0,\ z_c$, and we impose $\mathbb{Z}_2$ symmetry across each brane. In the first Randall-Sundrum model, dubbed RS1, z_c is finite, whereas in the second model, RS2, we take $z_c \to \infty$, so we effectively have a single brane.

Neglecting Gibbons-Hawking boundary terms for brevity, the action describing the Randall-Sundrum model is given by

$$\begin{aligned} S = \frac{M_5^3}{2}\int d^4x \int_{-z_c}^{z_c} dz\sqrt{-\gamma}\,(\mathcal{R} - 2\Lambda) \\ -\sigma_+\int_{z=0} d^4x\sqrt{-g^{(+)}} - \sigma_-\int_{z=z_c} d^4x\sqrt{-g^{(-)}}\ . \end{aligned} \tag{12.4}$$

where γ_{ab} is the bulk metric and $g^{(+)}_{\mu\nu},\ g^{(-)}_{\mu\nu}$ are the metrics on the branes at $z = 0,\ z_c$ respectively. M_5 is the five dimensional Planck scale and is related to the five dimensional Newton's constant via the standard relation $G_5 = 1/8\pi M_5^3$. We also include a *negative* bulk cosmological constant, $\Lambda = -6k^2$. If we fine-tune the brane tensions against Λ, such that

$$\sigma_+ = -\sigma_- = 6M_5^3 k = \frac{3k}{4\pi G_5}$$

we admit a background solution in which the branes exhibit four-dimensional Poincaré invariance

$$ds^2 = e^{-2k|z|}\eta_{\mu\nu}dx^\mu dx^\nu + dz^2 \quad \text{for } -z_c \le z \le z_c. \tag{12.5}$$

The $\mathbb{Z}_2$ symmetry about $z = 0$ is explicit whereas the other boundary condition should be understood. The metric (12.5) contains an exponential warp factor which is seen graphically in figure 12.1. In between branes, we recognise this as a section of anti-de Sitter space, written in Poincaré coordinates. Notice the peak in the warp factor at the positive tension brane and the trough at the negative tension brane.

One can consider weak gravity around this solution. The analysis is rather complicated as it corresponds to a coupled brane-bulk system governed by equations (12.2) and (12.18). We don't have time for that here, so let us simply quote some results.

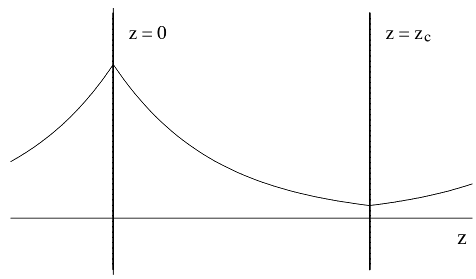


Figure 12.1: The behaviour of the warp factor in the Randall-Sundrum model. The positive tension brane is at $z = 0$, while the negative tension brane is at $z = z_c$.

By integrating out the $4D$ zero mode we are able to derive the $4D$ effective Planck scale on a given brane

$$M_\pm^2 = \pm \frac{M_5^3}{k}\left(1 - e^{\mp 2kz_c}\right) \tag{12.6}$$

where $\pm$ labels the sign of the corresponding brane tension. In terms of the effective Newton's constants, we have

$$G_\pm = G_5 k \left(\frac{\pm 1}{1 - e^{\mp 2kz_c}}\right) \tag{12.7}$$

The low energy $4D$ effective theory on either brane is not GR but Brans-Dicke gravity. The extra scalar comes from fluctuations in the brane separation and is sometimes referred to as the *radion* . The value of the Brans-Dicke parameter depends on the brane, and is given by,

$$w^{(\pm)} = \frac{3}{2}\left(e^{\pm 2kz_c} - 1\right) \tag{12.8}$$

We know that observations require this parameter to be large $w > 40000$. This is problematic for RS1 gravity (two branes), so we must generate a mass for the radion to suppress its fluctuations. The Goldberger-Wise mechanism does exactly that and the distance between the branes is stabilised.

In contrast, GR can be recovered on the one remaining brane (with positive tension) in the RS2 model. This is easily seen from the fact $w_{BD}^{(+)}$ can be made arbitrarily large with increasing brane separation.

12.2 Brane cosmology

There are two obvious reasons why cosmology offers an interesting arena in which to develop the braneworld paradigm. The first is that cosmological branes possess a high degree of symmetry, and this makes it possible to solve the field equations. The second is that cosmological physics can be tested by a number of observations, ranging from supernova data to the abundance of light elements. In this section we will study the cosmology of co-dimension one branes, focusing on the RS2 scenario with a single $\mathbb{Z}_2$ symmetric brane.

Braneworld cosmology can be studied using two different formalisms: The *brane based* formalism, and the *bulk based* formalism. These two approaches are completely equivalent and yield a background cosmology governed by the following Friedmann equations

$$\begin{aligned} H^2 + \frac{\kappa}{a^2} &= \frac{\Lambda_4}{3} + \frac{8\pi G_4}{3}\rho\left(1+\frac{\rho}{2\sigma}\right) + \frac{\mu}{a^4}, && (12.9) \\ \dot{H} - \frac{\kappa}{a^2} &= -4\pi G_4(\rho+P)\left(1+\frac{\rho}{\sigma}\right) - \frac{2\mu}{a^4}, && (12.10) \end{aligned}$$

where $H = \dot{a}/a$ is the Hubble parameter along the brane, $a(t)$ is the scale factor, and $\kappa = 0, \pm 1$ describes the spatial curvature. The brane is sourced by a tension, σ, and a cosmological fluid with energy density, $\rho(t)$, and pressure, $P(t)$. The parameters Λ_4 and G_4 denote the effective cosmological constant and Newton's constant on the brane, respectively. As in the standard scenario, the Raychaudhuri Equation (12.10) follows from the Friedmann Equation (12.9) and energy conservation,

$$\dot{\rho} + 3H(\rho + P) = 0. \tag{12.11}$$

We will review the derivation of this cosmology using the *bulk based* formalism. For the moment, however, let us comment on a few of its important features. From Equations (12.9) and (12.10) we see that the corrections to the standard cosmology manifest themselves in a term $\propto \rho^2$, and a *dark radiation* term $\frac{\mu}{a^4} = \frac{8\pi G_4}{3}\rho_{\text{weyl}}$. The latter corresponds to a non-local "Weyl" contribution and can only be fixed by specifying the bulk geometry. In the holographic description of RS2, the ρ^2 corrections contribute to the conformal anomaly, while the dark radiation is identified with thermal excitations of the CFT.

We will now develop the bulk based formalism for cosmological brane, assuming $\mathbb{Z}_2$ symmetry. brevity. The bulk based formalism requires us to solve for the bulk geometry. Since we are interested in cosmological branes (with constant curvature Euclidean 3-spaces), we study the Einstein equations, $\mathcal{R}_{ab} - \frac{1}{2}\mathcal{R}\gamma_{ab} = \frac{6}{l^2}\gamma_{ab}$, with the following metric ansatz

$$ds^2 = \gamma_{ab}dx^a dx^b = e^{2\nu}A^{-2/3}(-dt^2 + dz^2) + A^{2/3}q_{ij}dx^i dx^j, \tag{12.12}$$

where A and ν are undetermined functions of t and z, and as before $q_{ij}(x)$ is the metric of a hyper-surface of constant curvature, $\kappa = 0, \pm 1$. Now, in an extremely elegant calculation, Bowcock *et al.* were able to prove a generalised form of Birkhoff's theorem, showing that the bulk geometry is necessarily given by

$$ds^2 = -V(r)d\tau^2 + \frac{dr^2}{V(r)} + r^2 q_{ij}dx^i dx^j, \tag{12.13}$$

where

$$V(r) = \frac{r^2}{l^2} + \kappa - \frac{\mu}{r^2}. \tag{12.14}$$

For $\mu > 0$, the metric in Eq. (12.13) takes the form of a (topological) Schwarzschild black hole in anti-de Sitter space. Here we have written the solution in an explicitly time independent coordinate system, meaning that we can no longer say that we have a static brane sitting at a fixed coordinate position. On the contrary, we now have a dynamic brane, whose trajectory in the these coordinates is more complicated.

To construct the brane solution, we treat it as an embedding,

$$\tau = \tau(t), \qquad r = a(t), \tag{12.15}$$

of the bulk geometry given in Eq. (12.13). The induced metric on the brane is then

$$ds^2 = g_{\mu\nu} d\xi^\mu d\xi^\nu = \left(-V(a)\dot{\tau}^2 + \frac{\dot{a}^2}{V(a)} \right) dt^2 + a^2(t) q_{ij} dx^i dx^j \tag{12.16}$$

where over-dots denote $\partial/\partial t$. We are free to choose t to correspond to the proper time with respect to an observer comoving with the brane. This imposes the condition

$$-V(a)\dot{\tau}^2 + \frac{\dot{a}^2}{V(a)} = -1, \tag{12.17}$$

ensuring that the brane takes the standard FLRW form.. The function $a(t)$ is then immediately identified with the scale factor along the brane.

The boundary condition at the brane are given by the so-called Israel junction conditions

$$K_{\mu\nu} - K\gamma_{\mu\nu} = 3\sigma_* g_{\mu\nu} - 4\pi G_5 \mathcal{T}_{\mu\nu}, \tag{12.18}$$

with $\sigma_* = \frac{4\pi G_5 \sigma}{3}$, where σ is the brane tension and $\mathcal{T}_{\mu\nu}$ is the energy-momentum tensor of additional matter excitations. We must compute the extrinsic curvature, $K_{\mu\nu} = \frac{1}{2}\mathcal{L}_n g_{\mu\nu}$, defined as the Lie derivative of the normal pointing *into* the bulk. Assuming we cut away the AdS boundary and retain the region $r < a(t)$, we find that the inward pointing unit normal is given by

$$n_a = (-\dot{a}, \dot{\tau}, 0, 0, 0), \tag{12.19}$$

where we are free to specify that $\dot{\tau} > 0$. The components of extrinsic curvature are then given by

$$K^i_j = \frac{V\dot{\tau}}{a}\delta^i_j, \qquad K^t_t = -\left(\frac{\ddot{a} + V'/2}{V\dot{\tau}} \right). \tag{12.20}$$

In the presence of a cosmological fluid $\mathcal{T}^\mu_\nu = diag(-\rho, P, P, P)$, the junction conditions in Eq. (12.18) yield the following:

$$\frac{V\dot{\tau}}{a} = \sigma_*(1 + \frac{\rho}{\sigma}), \tag{12.21}$$

$$\frac{\ddot{a} + \frac{1}{2}V'}{V\dot{\tau}} = \sigma_* \left[1 - 2\frac{\rho}{\sigma} - 3\frac{P}{\sigma} \right]. \tag{12.22}$$

Making use of Equation (12.17) we then arrive at the modified Friedmann Equations (12.9) and (12.10).