

# Notes on General Relativity and Cosmology

## Oxford Physics Department

Steven Balbus

Version: April 3, 2018

*These notes are intended for classroom use. There is little here that is original, but in fact, a few results are. If you would like to quote from these notes and are unable to find an original literature reference, please contact me by email first. Conversely, if you find errors in the text, I would be grateful to have these brought to my attention. I thank those of you who already have done so: C. Arrowsmith, J. Bonifacio, Y. Chen, M. Elliot, G. Johnson, A. Orr, J. Patton, W. Potter, and T. Surawy-Stepney.*

*Thanks to the reader in advance for his/her cooperation.*

Steven Balbus

# Recommended Texts

Hobson, M. P., Efstathiou, G., and Lasenby, A. N. 2006, *General Relativity: An Introduction for Physicists*, (Cambridge: Cambridge University Press) Referenced as HEL06.

A very clear, very well-blended book, admirably covering the mathematics, physics, and astrophysics of GR. Excellent presentation of black holes and gravitational radiation. The explanation of the geodesic equation and the affine connection is very clear and enlightening. Not so much on cosmology, though a nice introduction to the physics of inflation. Overall, my favourite text on this topic. (The metric has a different sign convention in HEL06 compared with Weinberg 1972 & MTW [see below], as well as these notes. Be careful.)

Weinberg, S. 1972, *Gravitation and Cosmology. Principles and Applications of the General Theory of Relativity*, (New York: John Wiley) Referenced as W72.

This has become a classic reference, but it is by now very dated. Very little material on black holes, an awkward treatment of gravitational radiation, and an unsatisfying development of the geometrical interpretation of the equations are all drawbacks. The author could not be more explicit in his aversion to anything geometrical. Gravity is a field theory with a mere geometrical “analogy,” according to Weinberg. From the introduction: “[A] student who asked why the gravitational field is represented by a metric tensor, or why freely falling particles follow geodesics, or why the field equations are generally covariant would come away with the impression that this had something to do with the fact that space-time is a Riemannian manifold.” Well, yes, actually. There is simply no way to understand the equations of gravity without immersing oneself in geometry, and this point-of-view has led to profound advances. (Please *do* come away from this course with the impression that space-time is something very like a Riemannian manifold.) The detailed sections on classical physical cosmology are this text’s main strength, and these are very fine indeed. Weinberg also has a more recent graduate text on cosmology *per se*, (*Cosmology* 2007, Oxford: Oxford University Press). This is very comprehensive, but at an advanced level and quite a difficult read.

Misner, C. W., Thorne, K. S., and Wheeler, J. A. 1973, *Gravitation*, (New York: Freeman) Reprinted 2017 by Princeton University Press. Referenced as MTW.

At 1280 pages, don’t drop this on your toe, not even the paperback version. MTW, as it is known, is often criticised for its sheer bulk and its seemingly endless meanderings. But look, I must say, in the end, there really is a lot of very good material in here, much that is difficult to find anywhere else. It is a monumental achievement. It is also the opposite of Weinberg: geometry is front and centre from start to finish, and there is lots and lots of black hole and nice gravitational radiation physics, 40+ years on more timely than ever. I heartily recommend its insightful discussion of gravitational radiation, now part of the course syllabus. There is a “Track 1” and “Track 2” for aid in navigation; Track 1 contains the essentials. (Update: A new hardbook edition has been published in October 2017 by Princeton University Press. Text is unchanged, but there is interesting new prefatory material.)

Hartle, J. B. 2003, *Gravity: An Introduction to Einstein’s General Theory of Relativity*, (San Francisco: Addison-Wesley)

This is GR Lite, at a very different level from the previous three texts. But for what it is meant to be, it succeeds very well. Coming into the subject cold, this is not a bad place to start to get the lay of the land, to understand the issues in their broadest context, and

to be treated to a very accessible presentation. General Relativity is a demanding subject. There will be times in your study of GR when it will be difficult to see the forest for the trees, when you will feel overwhelmed with the calculations, drowning in a sea of indices and Riemannian formalism. Everything will be all right: just spend some time with this text.

Ryden, Barbara 2017, *Introduction to Cosmology*, (Cambridge: Cambridge University Press)

Very recent and therefore up-to-date second edition of an award-winning text. The style is clear and lucid, the level is right, and the choice of topics is excellent. Less GR and more astrophysical in content but with a blend appropriate to the subject matter. Ryden is always very careful in her writing, making this a real pleasure to read. Warmly recommended.

A few other texts of interest:

Binney J., and Tremaine, S. 2008, *Galactic Dynamics*, (Princeton: Princeton University Press) Masterful text on galaxies with an excellent cosmology treatment in the appendix. Very readable, given the high level of mathematics.

Longair, M., 2006, *The Cosmic Century*, (Cambridge: Cambridge University Press) Excellent blend of observations, theory and history of cosmology, as part of a more general study. Good general reference for anyone interested in astrophysics.

Landau, L., and Lifschitz, E. M. 1962, *Classical Theory of Fields*, (Oxford: Pergamon) Classic advanced text; original and interesting treatment of gravitational radiation. Dedicated students only!

Peebles, P. J. E. 1993, *Principles of Physical Cosmology*, (Princeton: Princeton University Press) Authoritative and advanced treatment by the leading cosmologist of the 20th century, but in my view a difficult and sometimes frustrating read. A better reference than text.

Shapiro, S., and Teukolsky S. 1983, *Black Holes, White Dwarfs, and Neutron Stars*, (Wiley: New York) A text I grew up with. Very clear, with a nice summary of applications of GR to compact objects and good physical discussions. Level is appropriate to this course. Highly recommended.

# Notational Conventions & Miscellany

- Spacetime dimensions are labelled 0, 1, 2, 3 or (Cartesian)  $ct, x, y, z$  or (spherical)  $ct, r, \theta, \phi$ . Time is always the 0-component. Beware of extraneous factors of  $c$  in 0-index quantities, present in e.g.  $T^{00} = \rho c^2$ ,  $dx^0 = c dt$ , but absent in e.g.  $g_{00} = -1$ . (That is one reason why some like to set  $c = 1$  from the start.)
- Repeated indices are summed over, unless otherwise specified. (Einstein summation convention.)
- The Greek indices  $\kappa, \lambda, \mu, \nu$  etc. are used to represent arbitrary spacetime components in all general relativity calculations.
- The Greek indices  $\alpha, \beta$ , etc. are used to represent arbitrary spacetime components in special relativity calculations (Minkowski spacetime).
- The Roman indices  $i, j, k$  are used to represent purely spatial components in any spacetime.
- The Roman indices  $a, b, c, d$  are used to represent fiducial spacetime components for mnemonic aids, and in discussions of how to perform index-manipulations and/or permutations, where Greek indices may cause confusion.
- $*$  is used to denote a generic dummy index, always summed over with another  $*$ .
- The tensor  $\eta^{\alpha\beta}$  is numerically identical to  $\eta_{\alpha\beta}$  with  $-1, 1, 1, 1$  corresponding to the 00, 11, 22, 33 diagonal elements. Other texts may use the sign convention  $1, -1, -1, -1$ . Be careful.
- Viewed as matrices, the metric tensors  $g_{\mu\nu}$  and  $g^{\mu\nu}$  are always inverses. The respective diagonal elements of *diagonal*  $g_{\mu\nu}$  and  $g^{\mu\nu}$  metric tensors are therefore reciprocals.
- $c$  almost always denotes the speed of light. It is occasionally used as an (obvious) tensor index.  $c$  as the velocity of light is only occasionally set to unity in these notes or in the problem sets; if so it is explicitly stated. (Relativity texts often set  $c = 1$  to avoid clutter.) Newton's  $G$  is *never* unity, no matter what. And don't you even think of setting  $2\pi$  to unity.
- It is "Lorentz invariance," but "Lorenz gauge." Not a typo, actually two different blokes.

# Really Useful Numbers

$$c = 2.99792458 \times 10^8 \text{ m s}^{-1} \text{ (Exact speed of light.)}$$

$$c^2 = 8.9875517873681764 \times 10^{16} \text{ m}^2 \text{ s}^{-2} \text{ (Exact!)}$$

$$a = 7.565723 \times 10^{-16} \text{ J m}^{-3} \text{ K}^{-4} \text{ (Blackbody radiation constant.)}$$

$$G = 6.67384 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2} \text{ (Newton's } G\text{.)}$$

$$M_{\odot} = 1.98855 \times 10^{30} \text{ kg} \text{ (Mass of the Sun.)}$$

$$r_{\odot} = 6.955 \times 10^8 \text{ m} \text{ (Radius of the Sun.)}$$

$$GM_{\odot} = 1.32712440018 \times 10^{20} \text{ m}^3 \text{ s}^{-2} \text{ (Solar gravitational parameter; much more accurate than either } G \text{ or } M_{\odot} \text{ separately.)}$$

$$2GM_{\odot}/c^2 = 2.9532500765 \times 10^3 \text{ m} \text{ (Solar Schwarzschild radius.)}$$

$$GM_{\odot}/c^2 r_{\odot} = 2.1231 \times 10^{-6} \text{ (Solar relativity parameter.)}$$

$$M_{\oplus} = 5.97219 \times 10^{24} \text{ kg} \text{ (Mass of the Earth)}$$

$$r_{\oplus} = 6.371 \times 10^6 \text{ m} \text{ (Mean Earth radius.)}$$

$$GM_{\oplus} = 3.986004418 \times 10^{14} \text{ m}^3 \text{ s}^{-2} \text{ (Earth gravitational parameter.)}$$

$$2GM_{\oplus}/c^2 = 8.87005608 \times 10^{-3} \text{ m} \text{ (Earth Schwarzschild radius.)}$$

$$GM_{\oplus}/c^2 r_{\oplus} = 6.961 \times 10^{-10} \text{ (Earth relativity parameter.)}$$

$$1 \text{ AU} = 1.495978707 \times 10^{11} \text{ m} \text{ (1 Astronomical Unit by definition.)}$$

$$1 \text{ pc} = 3.085678 \times 10^{16} \text{ m} \text{ (1 parsec.)}$$

$$H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1} \text{ (Hubble constant. } h \simeq 0.7. \text{ } H_0^{-1} = 3.085678h^{-1} \times 10^{17} \text{ s} = 9.778h^{-1} \times 10^9 \text{ yr.)}$$

## For diagonal $g_{ab}$ ,

$$\Gamma_{ba}^a = \Gamma_{ab}^a = \frac{1}{2g_{aa}} \frac{\partial g_{aa}}{\partial x^b} \quad (a = b \text{ permitted, NO SUM})$$

$$\Gamma_{bb}^a = -\frac{1}{2g_{aa}} \frac{\partial g_{bb}}{\partial x^a} \quad (a \neq b, \text{ NO SUM})$$

$$\Gamma_{bc}^a = 0, \quad (a, b, c \text{ distinct})$$

## Ricci tensor:

$$R_{\mu\kappa} = \frac{1}{2} \frac{\partial^2 \ln |g|}{\partial x^{\kappa} \partial x^{\mu}} - \frac{\partial \Gamma_{\mu\kappa}^{\lambda}}{\partial x^{\lambda}} + \Gamma_{\mu\lambda}^{\eta} \Gamma_{\kappa\eta}^{\lambda} - \frac{\Gamma_{\mu\kappa}^{\eta}}{2} \frac{\partial \ln |g|}{\partial x^{\eta}} \quad (\text{FULL SUMMATION, } g = \det g_{\mu\nu})$$

# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>An overview</b>   | <b>11</b> |
| 1.1      | The legacy of Maxwell . . . . .  | 11        |
| 1.2      | The legacy of Newton . . . . .   | 12        |
| 1.3      | The need for a geometrical framework . . . . .                             | 12        |
| <b>2</b> | <b>The toolbox of geometrical theory: Minkowski spacetime</b>              | <b>15</b> |
| 2.1      | The 4-vector formalism . . . . .   | 15        |
| 2.2      | Transformation of gradients . . . . .                                      | 18        |
| 2.3      | Transformation matrix . . . . .  | 19        |
| 2.4      | Tensors . . . . .  | 21        |
| 2.4.1    | Conservation of $T^{\alpha\beta}$ . . . . .                                | 22        |
| <b>3</b> | <b>The effects of gravity</b>  | <b>24</b> |
| 3.1      | The Principle of Equivalence . . . . .                                     | 24        |
| 3.2      | The geodesic equation . . . . .  | 26        |
| 3.3      | The metric tensor . . . . .  | 28        |
| 3.4      | The relationship between the metric tensor and affine connection . . . . . | 28        |
| 3.5      | Variational calculation of the geodesic equation . . . . .                 | 29        |
| 3.6      | The Newtonian limit . . . . .  | 31        |
| <b>4</b> | <b>Tensor Analysis</b>   | <b>35</b> |
| 4.1      | Transformation laws . . . . .  | 35        |
| 4.2      | The covariant derivative . . . . .   | 37        |
| 4.3      | The affine connection and basis vectors . . . . .                          | 40        |
| 4.4      | Volume element . . . . .   | 42        |
| 4.5      | Covariant div, grad, curl, and all that . . . . .                          | 43        |
| 4.6      | Hydrostatic equilibrium . . . . .  | 44        |
| 4.7      | Covariant differentiation and parallel transport . . . . .                 | 45        |
| <b>5</b> | <b>The curvature tensor</b>  | <b>47</b> |
| 5.1      | Commutation rule for covariant derivatives . . . . .                       | 47        |
| 5.2      | Parallel transport . . . . .   | 48        |
| 5.3      | Algebraic identities of $R^{\sigma}_{\nu\lambda\rho}$ . . . . .            | 50        |
| 5.3.1    | Remembering the curvature tensor formula. . . . .                          | 50        |
| 5.3.2    | $R_{\lambda\mu\nu\kappa}$ : fully covariant form . . . . .                 | 50        |
| 5.4      | The Ricci Tensor . . . . .   | 52        |

|          |   |            |
|----------|---|------------|
| 5.5      | Curvature and Newtonian gravity . . . . .                         | 53         |
| 5.6      | The Bianchi Identities . . . . .                                  | 53         |
| <b>6</b> | <b>The Einstein Field Equations</b>                               | <b>56</b>  |
| 6.1      | Formulation . . . . .   | 56         |
| 6.2      | Coordinate ambiguities . . . . .                                  | 59         |
| 6.3      | The Schwarzschild Solution . . . . .                              | 59         |
| 6.4      | The Schwarzschild Radius . . . . .                                | 64         |
| 6.5      | Schwarzschild spacetime. . . . .                                  | 66         |
| 6.5.1    | Radial photon geodesic . . . . .                                  | 66         |
| 6.5.2    | Orbital equations . . . . .                                       | 67         |
| 6.6      | The deflection of light by an intervening body. . . . .           | 70         |
| 6.7      | The advance of the perihelion of Mercury . . . . .                | 71         |
| 6.7.1    | Newtonian orbits . . . . .  | 73         |
| 6.7.2    | The perihelion advance of Mercury . . . . .                       | 74         |
| 6.8      | Shapiro delay: the fourth protocol . . . . .                      | 75         |
| <b>7</b> | <b>Gravitational Radiation</b>                                    | <b>78</b>  |
| 7.1      | The linearised gravitational wave equation . . . . .              | 80         |
| 7.1.1    | Come to think of it... . . . .                                    | 84         |
| 7.2      | Plane waves . . . . .   | 85         |
| 7.2.1    | The transverse-traceless (TT) gauge . . . . .                     | 86         |
| 7.3      | The quadrupole formula . . . . .                                  | 88         |
| 7.4      | Radiated Energy . . . . .   | 89         |
| 7.4.1    | A useful toy problem . . . . .                                    | 89         |
| 7.4.2    | A conserved energy flux for linearised gravity . . . . .          | 90         |
| 7.5      | The energy loss formula for gravitational waves . . . . .         | 93         |
| 7.6      | Gravitational radiation from binary stars . . . . .               | 96         |
| 7.7      | Detection of gravitational radiation . . . . .                    | 99         |
| 7.7.1    | Preliminary comments . . . . .                                    | 99         |
| 7.7.2    | Indirect methods: orbital energy loss in binary pulsars . . . . . | 101        |
| 7.7.3    | Direct methods: LIGO . . . . .                                    | 104        |
| 7.7.4    | Direct methods: Pulsar timing array . . . . .                     | 108        |
| <b>8</b> | <b>Cosmology</b>  | <b>109</b> |
| 8.1      | Introduction . . . . .  | 109        |
| 8.1.1    | Newtonian cosmology . . . . .                                     | 109        |



|          |  |            |
|----------|--|------------|
| 8.1.2    | The dynamical equation of motion . . . . .                               | 111        |
| 8.1.3    | Cosmological redshift . . . . .  | 112        |
| 8.2      | Cosmology models for the impatient . . . . .                             | 113        |
| 8.2.1    | The large-scale spacetime metric . . . . .                               | 113        |
| 8.2.2    | The Einstein-de Sitter universe: a useful toy model . . . . .            | 114        |
| 8.3      | The Friedman-Robertson-Walker Metric . . . . .                           | 117        |
| 8.3.1    | Maximally symmetric 3-spaces . . . . .                                   | 118        |
| 8.4      | Large scale dynamics . . . . .   | 120        |
| 8.4.1    | The effect of a cosmological constant . . . . .                          | 121        |
| 8.4.2    | Formal analysis . . . . .  | 122        |
| 8.4.3    | The $\Omega_0$ parameter . . . . .                                       | 125        |
| 8.5      | The classic, matter-dominated universes . . . . .                        | 125        |
| 8.6      | Our Universe . . . . .   | 127        |
| 8.6.1    | Prologue . . . . .   | 127        |
| 8.6.2    | A Universe of ordinary matter and vacuum energy . . . . .                | 128        |
| 8.7      | Observational foundations of cosmology . . . . .                         | 129        |
| 8.7.1    | The first detection of cosmological redshifts . . . . .                  | 129        |
| 8.7.2    | The cosmic distance ladder . . . . .                                     | 132        |
| 8.7.3    | The parameter $q_0$ . . . . .  | 134        |
| 8.7.4    | The redshift–magnitude relation . . . . .                                | 135        |
| 8.8      | Radiation-dominated universe . . . . .                                   | 138        |
| 8.9      | The Cosmic Microwave Background Radiation (CMB) . . . . .                | 138        |
| 8.9.1    | Overview . . . . .   | 138        |
| 8.9.2    | An observable cosmic radiation background: the Gamow argument . . . . .  | 140        |
| 8.9.3    | The cosmic microwave background (CMB): subsequent developments . . . . . | 144        |
| 8.10     | Thermal history of the Universe . . . . .                                | 146        |
| 8.10.1   | Prologue . . . . .   | 146        |
| 8.10.2   | Classical cosmology: Helium nucleosynthesis . . . . .                    | 147        |
| 8.10.3   | Neutrino and photon temperatures . . . . .                               | 149        |
| 8.10.4   | Ionisation of Hydrogen . . . . .   | 151        |
| <b>9</b> | <b>The Seeds of Structure</b> . . . . .                                  | <b>153</b> |
| 9.1      | The growth of density perturbations in an expanding universe . . . . .   | 153        |
| 9.2      | Inflationary Models . . . . .  | 155        |
| 9.2.1    | “Clouds on the horizon” . . . . .  | 155        |
| 9.2.2    | The stress energy tensor of a field . . . . .                            | 156        |

|                            |     |
|----------------------------|-----|
| 9.3 A Final Word . . . . . | 165 |
|----------------------------|-----|

*Most of the fundamental ideas of science are essentially simple, and may, as a rule, be expressed in a language comprehensible to everyone.*

— Albert Einstein

# 1 An overview

## 1.1 The legacy of Maxwell

We are told by the historians that the greatest Roman generals would have their most important victories celebrated with a triumph. The streets would line with adoring crowds, cheering wildly in support of their hero as he passed by in a grand procession. But the Romans astutely realised the need for a counterpoise, so a slave would ride with the general, whispering in his ear, “All glory is fleeting.”

All glory is fleeting. And never more so than in theoretical physics. No sooner is a triumph hailed, but unforeseen puzzles emerge that couldn’t possibly have been anticipated before the breakthrough. The mid-nineteenth century reduction of all electromagnetic phenomena to four equations, the “Maxwell Equations,” is very much a case in point.

Maxwell’s equations united electricity, magnetism, and optics, showing them to be different manifestations of the same field. The theory accounted for the existence of electromagnetic waves, explained how they propagate, and that the propagation velocity is  $1/\sqrt{\epsilon_0\mu_0}$  ( $\epsilon_0$  is the permittivity, and  $\mu_0$  the permeability, of free space). This combination is numerically precisely equal to the speed of light. Light is electromagnetic radiation! The existence of electromagnetic radiation was then verified by brilliant experiments carried out by Heinrich Hertz in 1887, in which the radiation was directly generated and detected.

But Maxwell’s theory, for all its success, had disquieting features when one probed. For one, there seemed to be no provision in the theory for allowing the velocity of light to change with the observer’s velocity. The speed of light is always  $1/\sqrt{\epsilon_0\mu_0}$ . A related point was that simple Galilean invariance was not obeyed, i.e. absolute velocities seemed to affect the physics, something that had not been seen before. Lorentz and Larmor in the late nineteenth century discovered that Maxwell’s equations did have a simple mathematical velocity transformation that left them invariant, but it was *not* Galilean, and most bizarrely, it involved changing the *time*. The non-Galilean character of the transformation equation relative to the “aetherial medium” hosting the waves was put down, a bit vaguely, to electromagnetic interactions between charged particles that truly changed the length of the object. In other words, the non-Galilean transformation was somehow electrodynamical in origin. As to the time change...well, one would just have to put up with it as an aetherial formality.

All was resolved in 1905 when Einstein showed how, by adopting as postulates (i) the speed of light is constant in all frames (as had already been indicated by a body of irrefutable experiments, including the famous Michelson-Morley investigation); (ii) the truly essential Galilean notion that relative uniform velocity cannot be detected by any physical experiment, that the “Lorentz transformations” (as they had become known) must follow. Moreover, electromagnetic radiation took on a reality its own. The increasingly problematic aether medium that supposedly hosted these waves could be abandoned; the waves were hosted by the vacuum itself. Finally, *all* equations of physics, not just electromagnetic phenomena, had to be invariant in form under the Lorentz transformations, even with its peculiar relative

time variable. The transformations were purely kinematic, having nothing in particular to do with electrodynamics. They were much more general. These ideas and the consequences that ensued collectively from them became known as *relativity theory*, in reference to the invariance of form with respect to relative velocities. The relativity theory stemming from Maxwell's equations is rightly regarded as one of the crown jewels of 20th century physics. In other words, a triumph.

## 1.2 The legacy of Newton

Another triumph, another problem. If indeed, all of physics had to be compatible with relativity, what of Newtonian gravity? It works incredibly well, yet it is manifestly *not* compatible with relativity, because Poisson's equation

$$\nabla^2\Phi = 4\pi G\rho \tag{1}$$

implies instantaneous transmission of changes in the gravitational field from source to potential. (Here  $\Phi$  is the Newtonian potential function,  $G$  the Newtonian gravitational constant, and  $\rho$  the mass density.) Wiggle the density locally, and throughout all of space there must instantaneously be a wiggle in  $\Phi$ , as given by equation (1).

In Maxwell's theory, the *electrostatic* potential satisfies its own Poisson equation, but the appropriate time-dependent potential obeys a wave equation:

$$\nabla^2\Phi - \frac{1}{c^2} \frac{\partial^2\Phi}{\partial t^2} = -\frac{\rho}{\epsilon_0}, \tag{2}$$

and solutions of this equation propagate signals at the speed of light  $c$ . In retrospect, this is rather simple. Mightn't it be the same for gravity?

No. The problem is that the source of the signals for the electric potential field, i.e. the charge density, behaves differently from the source for the gravity potential field, i.e. the mass density. The electrical charge of an individual bit of matter does not change when the matter is viewed in motion, but the mass does: the mass increases with velocity. This seemingly simple detail complicates everything. Moreover, in a relativistic theory, energy, like matter, is a source of a gravitational field, including the distributed energy of the gravitational field itself! A relativistic theory of gravity would have to be nonlinear. In such a time-dependent theory of gravity, it is not even clear a priori what the appropriate mathematical objects should be on either the right side or the left side of the wave equation. Come to think of it, should we be using a wave equation at all?

## 1.3 The need for a geometrical framework

In 1908, the mathematician Hermann Minkowski came along and argued that one should view the Lorentz transformations not merely as a set of rules for how coordinates (including a time coordinate) change from one constant-velocity reference frame to another, but that these coordinates should be regarded as living in their own sort of pseudo-Euclidian geometry—a *spacetime*, if you will: Minkowski spacetime.

To understand the motivation for this, start simply. We know that in ordinary Euclidian space we are free to choose any coordinates we like, and it can make no difference to the description of the space itself, for example, in measuring how far apart objects are. If  $(x, y)$

is a set of Cartesian coordinates for the plane, and  $(x', y')$  another coordinate set related to the first by a rotation, then

$$dx^2 + dy^2 = dx'^2 + dy'^2 \quad (3)$$

i.e., the distance between two closely spaced points is the same number, regardless of the coordinates used.  $dx^2 + dy^2$  is said to be an “invariant.”

Now, an abstraction. There is nothing special from a mathematical viewpoint about the use of  $dx^2 + dy^2$  as our so-called metric. Imagine a space in which the metric invariant was  $dy^2 - dx^2$ . From a purely mathematical point of view, we needn't worry about the plus/minus sign. An invariant is an invariant. However, with  $dy^2 - dx^2$  as our invariant, we are describing a Minkowski space, with  $dy = cdt$  and  $dx$  an ordinary space interval, just as before. Under Lorentz transformations,  $c^2dt^2 - dx^2$  is in fact an invariant quantity, and this is precisely what we need in order to guarantee that the speed of light is always constant—an invariant! In this case,  $c^2dt^2 - dx^2$  is always zero for light propagation along  $x$ , whatever coordinates (read “observers”) are involved, and more generally,

$$c^2dt^2 - dx^2 - dy^2 - dz^2 = 0 \quad (4)$$

will guarantee the same in any direction. We have thus taken a kinematical requirement—that the speed of light be a universal constant—and given it a geometrical interpretation in terms of an invariant quantity (a “quadratic form” as it is sometimes called) in Minkowski space. Rather, Minkowski's *spacetime*.

Pause. As the French would say, “Bof.” And so what? Call it whatever you like. Who needs obfuscating mathematical pretence? Eschew obfuscation! Lots of things add quadratically. The Lorentz transform stands on its own! I like my way! That was very much Einstein's initial take on Minkowski's pesky little meddling with his theory.

However, it is the geometrical viewpoint that is the more fundamental. In Minkowski's 1908 tour-de-force paper, we find the first mention of 4-vectors, of relativistic tensors, of the Maxwell equations in manifestly covariant form, and the realisation that the magnetic vector and electrostatic potentials combine to form a 4-vector. Gone are the comforting clocks and rods of 1905 relativity. This is more than “überflüssige Gelehrsamkeit” (superfluous erudition), Einstein's dismissive term for the whole business. In 1912, Einstein dramatically changed his opinion. His great revelation, his big idea, was that *the effect of the presence of matter (or its equivalent energy) in the universe is to distort a truly Minkowski spacetime, and this embedded distortion manifests itself as the gravitational field*. Minkowski spacetime is physical stuff. This same distortion must become, in the limit of weak gravity, familiar Newtonian theory. You thought gravity was dynamics? Nope. Gravity is a purely geometrical phenomenon.

Whoa. Now that is one big idea. It is an idea that will take the rest of this course—and beyond—to understand. How did Einstein make this leap? Why did he change his mind? Where did this notion of a gravity-geometry connection come from?

From a simple observation. In a freely falling elevator, or more safely in an aircraft executing a ballistic parabolic arch, one feels “weightless.” That is, the effect of gravity can be made to locally disappear in the appropriate reference frame—the right coordinates. This is because gravity has exactly the same effect on all types mass, regardless of composition, which is precisely what we would expect if objects were responding to background geometrical distortions instead of an applied force. In free-fall, the effective absence of gravity, we locally return to the environment of undistorted (“flat,” in mathematical parlance) Minkowski spacetime, much as a flat Euclidian tangent plane is an excellent local approximation to the surface of a curved sphere. This is why it is easy to be fooled into thinking that the earth is flat, if your view is local. “Tangent plane coordinates” on small scale road maps locally eliminate spherical geometry complications, but if we are flying to Hong Kong, the

earth's curvature is important. Einstein's notion that the effect of gravity is to cause a geometrical distortion of an otherwise flat Minkowski spacetime, and therefore that it is always possible to find coordinates in which local distortions may be eliminated *to leading order*, is the foundational insight of general relativity. It is known as the *Equivalence Principle*. We will have more to say on this topic.

Spacetime. *Spacetime*. Bringing in time, you see, is everything. Who would have thought of it in a geometrical theory? Non-Euclidean geometry, as developed by the great mathematician Bernhard Riemann, begins with just the notion we've been discussing, that any distorted space looks *locally* flat. Riemannian geometry is the natural language of gravitational theory, and Riemann himself had the notion that gravity might arise from a non-Euclidian curvature in three-dimensional space. He got nowhere, because *time was not part of his geometry*. He is thinking of space. It was the (in my view underrated) genius of Minkowski, who in showing us how to incorporate time into a purely geometrical theory, allowed Einstein to take the crucial next step, freeing himself to think of gravity in geometrical terms, without having to agonise over whether it made any sense to have time as part of a geometrical framework! In fact, the Newtonian limit is reached *not* from the leading order curvature terms in the spatial part of the geometry, but from the leading order "curvature" (if that is the word) of the *time* dimension. That is why Riemann failed.

Riemann created the mathematics of non-Euclidian geometry. Minkowski realised that the natural language of the Lorentz transformations was neither electrodynamical, nor even kinematic, it was really geometrical. But you need to include time as a component of the geometrical interpretation! Einstein took the great leap of realising that gravity arises from the distortions of Minkowski's flat spacetime created by the existence of mass/energy.

Well done. You now understand the conceptual framework of general relativity, and that is itself a giant leap. From here on, it is just a matter of the technical details. But then, you and I also can paint like Leonardo da Vinci. It is just a matter of the technical details.

*From henceforth, space by itself and time by itself, have vanished into the merest shadows, and only a blend of the two exists in its own right.*

— Hermann Minkowski

## 2 The toolbox of geometrical theory: Minkowski space-time

In what sense is general relativity “general?” In the sense that as we are dealing with a true spacetime geometry, the essential mathematical description must be the same in *any* coordinate system at all, not just among those related by constant velocity reference frame shifts, nor even just among those coordinate transformations that make tangible physical sense as belonging to some observer or another. *Any mathematically proper coordinates at all, however unusual.* Full stop.

We need coordinates for our description of the structure of spacetime, but somehow the essential physics (and other mathematical properties) must not depend on which coordinates we use, and it is no easy business to formulate a theory which satisfies this restriction. We owe a great deal to Bernhard Riemann for coming up with a complete mathematical theory for these non-Euclidian geometries. The sort of geometry in which it is always possible to find coordinates in which the space looks locally smooth is known as a *Riemannian manifold*. Mathematicians would say that an  $n$ -dimensional manifold is *homeomorphic* to  $n$ -dimensional Euclidian space. Actually, since our local invariant interval  $c^2 dt^2 - dx^2$  is not a simple sum of squares, but contains a minus sign, the manifold is said to be *pseudo-Riemannian*. Pseudo or no, the descriptive mathematical machinery is the same.

The objects that geometrical theories work with are scalars, vectors, and higher order tensors. You have certainly seen scalars and vectors before in your other physics courses, and you may have encountered tensors as well. We will need to be very careful how we define these objects, and very careful to distinguish them from objects that look like vectors and tensors (because they have many components and index labels) but actually are not.

To set the stage, we begin with the simplest geometrical objects of Minkowski spacetime that are not just simple scalars: the 4-vectors.

### 2.1 The 4-vector formalism

In their most elementary guise, the familiar Lorentz transformations from “fixed” laboratory coordinates  $(t, x, y, z)$  to moving frame coordinates  $(t', x', y', z')$  take the form

$$ct' = \gamma(ct - vx/c) = \gamma(ct - \beta x) \quad (5)$$

$$x' = \gamma(x - vt) = \gamma(x - \beta ct) \quad (6)$$

$$y' = y \quad (7)$$

$$z' = z \quad (8)$$

where  $v$  is the relative velocity (taken along the  $x$  axis),  $c$  the speed of light,  $\beta = v/c$  and

$$\gamma \equiv \frac{1}{\sqrt{1 - v^2/c^2}} \equiv \frac{1}{\sqrt{1 - \beta^2}} \quad (9)$$

is the *Lorentz factor*. The primed frame can be thought of as the frame moving with an object we are studying, that is to say the object's rest frame. To go backwards to find  $(x, t)$  as a function  $(x', t')$ , just interchange the primed and unprimed coordinates in the above equations, and then flip the sign of  $v$ . Do you understand why this works?

*Exercise.* Show that in a coordinate free representation, the Lorentz transformations are

$$ct' = \gamma(ct - \boldsymbol{\beta} \cdot \mathbf{x}) \quad (10)$$

$$\mathbf{x}' = \mathbf{x} + \frac{(\gamma - 1)}{\beta^2}(\boldsymbol{\beta} \cdot \mathbf{x})\boldsymbol{\beta} - \gamma ct\boldsymbol{\beta} \quad (11)$$

where  $c\boldsymbol{\beta} = \mathbf{v}$  is the vector velocity and boldface  $\mathbf{x}$ 's are spatial vectors. (Hint: This is not *nearly* as scary as it looks! Note that  $\boldsymbol{\beta}/\beta$  is just a unit vector in the direction of the velocity and sort out the components of the equation.)

*Exercise.* The Lorentz transformation can be made to look more rotation-like by using hyperbolic trigonometry. The idea is to place equations (5)–(8) on the same footing as the transformation of Cartesian position vector components under a simple rotation, say about the  $z$  axis:

$$x' = x \cos \theta + y \sin \theta \quad (12)$$

$$y' = -x \sin \theta + y \cos \theta \quad (13)$$

$$z' = z \quad (14)$$

Show that if we define

$$\beta \equiv \tanh \zeta, \quad (15)$$

then

$$\gamma = \cosh \zeta, \quad \gamma\beta = \sinh \zeta, \quad (16)$$

and

$$ct' = ct \cosh \zeta - x \sinh \zeta, \quad (17)$$

$$x' = -ct \sinh \zeta + x \cosh \zeta. \quad (18)$$

What happens if we apply this transformation twice, once with “angle”  $\zeta$  from  $(x, t)$  to  $(x', t')$ , then with angle  $\xi$  from  $(x', t')$  to  $(x'', t'')$ ? How is  $(x, t)$  related to  $(x'', t'')$ ?

Following on, rotations can be made to look more Lorentz-like by introducing

$$\alpha \equiv \tan \theta, \quad \Gamma \equiv \frac{1}{\sqrt{1 + \alpha^2}} \quad (19)$$

Then show that (12) and (13) become

$$x' = \Gamma(x + \alpha y) \quad (20)$$

$$y' = \Gamma(y - \alpha x) \quad (21)$$



Thus, while having a different appearance, the Lorentz and rotational transformations have mathematical structures that are similar. The Universe has both timelike and spacelike dimensional extensions; what distinguishes a timelike extension from a spacelike extension is the symmetry it exhibits. Spacelike extensions exhibit rotational (trigonometric) symmetry amongst themselves. Timelike extensions exhibit Lorentzian (hyperbolic trigonometric) symmetry with a spacelike extension, and probably nothing with another timelike extension. Maybe that is self-consistently why there can only be one timelike dimension in the Universe. The fundamental imperative for symmetry is often a powerful constraint in physics.

Of course lots of quantities besides position are vectors, and it is possible (indeed desirable) just to *define* a quantity as a vector if its individual components satisfy equations (12)–(14). Likewise, we find that many quantities in physics obey the transformation laws of equations (5–8), and it is therefore natural to give them a name and to probe their properties more deeply. We call these quantities 4-vectors. They consist of an ordinary vector  $\mathbf{V}$ , together with an extra component—a “time-like” component we will designate as  $V^0$ . (We use superscripts for a reason that will become clear later.) The “space-like” components are then  $V^1, V^2, V^3$ . The generic form for a 4-vector is written  $V^\alpha$ , with  $\alpha$  taking on the values 0 through 3. Symbolically,

$$V^\alpha = (V^0, \mathbf{V}) \quad (22)$$

We have seen that  $(ct, \mathbf{x})$  is one 4-vector. Another, you may recall, is the *4-momentum*,

$$p^\alpha = (E/c, \mathbf{p}) \quad (23)$$

where  $\mathbf{p}$  is the ordinary momentum vector and  $E$  is the total energy. Of course, we speak of relativistic momentum and energy:

$$\mathbf{p} = \gamma m \mathbf{v}, \quad E = \gamma mc^2 \quad (24)$$

where  $m$  is a particle’s rest mass. Just as

$$(ct)^2 - x^2 \quad (25)$$

is an invariant quantity under Lorentz transformations, so too is

$$E^2 - (pc)^2 = m^2 c^4 \quad (26)$$

A rather plain 4-vector is  $p^\alpha$  without the coefficient of  $m$ . This is the 4-velocity  $U^\alpha$ ,

$$U^\alpha = \gamma(c, \mathbf{v}) \quad (27)$$

Note that in the rest frame of a particle,  $U^0 = c$  (a constant) and the ordinary 3-velocity components  $\mathbf{U} = 0$ . To get to any other frame, just use (“boost with”) the Lorentz transformation along the  $\mathbf{v}$  direction. (Be careful with the sign of  $v$ ). We don’t have to worry that we boost along one axis only, whereas the velocity has three components. If you wish, just rotate the axes as you like after we’ve boosted. This sorts out all the 3-vector components and leaves the time (“0”) component untouched.

Humble in appearance, the 4-velocity is a most important 4-vector. Via the simple trick of boosting, the 4-velocity may be used as the starting point for constructing many other important physical 4-vectors. Consider, for example, a charge density  $\rho_0$  which is at rest. We may create a 4-vector which, in the rest frame, has only one component:  $\rho_0 c$  is the lonely time component and the ordinary spatial vector components are all zero. It is just like  $U^\alpha$ , only with a different normalisation constant. Now boost! The resulting 4-vector is denoted

$$J^\alpha = \gamma(c\rho_0, \mathbf{v}\rho_0) \quad (28)$$

The time component gives the charge density in any frame, and the 3- vector components are the corresponding standard current density  $\mathbf{J}$ ! This 4-current is the fundamental 4-vector of Maxwell's theory. As the source of the fields, this 4-vector source current is the basis for Maxwell's electrodynamics being a fully relativistic theory.  $J^0$  is the source of the electric field potential function  $\Phi$ , and  $\mathbf{J}$  is the source of the magnetic field vector potential  $\mathbf{A}$ . Moreover, as we will shortly see,

$$A^\alpha = (\Phi, \mathbf{A}/c) \quad (29)$$

is itself a 4-vector! From here, we can generate the electromagnetic fields themselves from the potentials by constructing a tensor...well, we are getting a bit ahead of ourselves.

## 2.2 Transformation of gradients

We have seen how the Lorentz transformation express  $x'^\alpha$  as a function of the  $x$  coordinates. It is a simple linear transformation, and the question naturally arises of how the partial derivatives,  $\partial/\partial t$ ,  $\partial/\partial x$  transform, and whether a 4-vector can be constructed from these components. This is a simple exercise. Using

$$ct = \gamma(ct' + \beta x') \quad (30)$$

$$x = \gamma(x' + \beta ct') \quad (31)$$

we find

$$\frac{\partial}{\partial t'} = \frac{\partial t}{\partial t'} \frac{\partial}{\partial t} + \frac{\partial x}{\partial t'} \frac{\partial}{\partial x} = \gamma \frac{\partial}{\partial t} + \gamma\beta c \frac{\partial}{\partial x} \quad (32)$$

$$\frac{\partial}{\partial x'} = \frac{\partial x}{\partial x'} \frac{\partial}{\partial x} + \frac{\partial t}{\partial x'} \frac{\partial}{\partial t} = \gamma \frac{\partial}{\partial x} + \gamma\beta \frac{1}{c} \frac{\partial}{\partial t} \quad (33)$$

In other words,

$$\frac{1}{c} \frac{\partial}{\partial t'} = \gamma \left( \frac{1}{c} \frac{\partial}{\partial t} + \beta \frac{\partial}{\partial x} \right) \quad (34)$$

$$\frac{\partial}{\partial x'} = \gamma \left( \frac{\partial}{\partial x} + \beta \frac{1}{c} \frac{\partial}{\partial t} \right) \quad (35)$$

and for completeness,

$$\frac{\partial}{\partial y'} = \frac{\partial}{\partial y} \quad (36)$$

$$\frac{\partial}{\partial z'} = \frac{\partial}{\partial z}. \quad (37)$$

This is *not* the Lorentz transformation (5)–(8); it differs by the sign of  $v$ . By contrast, *coordinate differentials*  $dx^\alpha$  transform, of course, just like  $x^\alpha$ :

$$cdt' = \gamma(cdt - \beta dx), \quad (38)$$

$$dx' = \gamma(dx - \beta cdt), \quad (39)$$

$$dy' = dy, \quad (40)$$

$$dz' = dz. \quad (41)$$

This has a very important consequence:

$$dt' \frac{\partial}{\partial t'} + dx' \frac{\partial}{\partial x'} = \gamma^2 \left[ (dt - \beta \frac{dx}{c}) \left( \frac{\partial}{\partial t} + \beta c \frac{\partial}{\partial x} \right) + (dx - \beta c dt) \left( \frac{\partial}{\partial x} + \beta \frac{1}{c} \frac{\partial}{\partial t} \right) \right], \quad (42)$$

or simplifying,

$$dt' \frac{\partial}{\partial t'} + dx' \frac{\partial}{\partial x'} = \gamma^2 (1 - \beta^2) \left( dt \frac{\partial}{\partial t} + dx \frac{\partial}{\partial x} \right) = dt \frac{\partial}{\partial t} + dx \frac{\partial}{\partial x} \quad (43)$$

Adding  $y$  and  $z$  into the mixture changes nothing. Thus, a scalar product exists between  $dx^\alpha$  and  $\partial/\partial x^\alpha$  that yields a *Lorentz scalar*, much as  $\mathbf{dx} \cdot \mathbf{\nabla}$ , the ordinary complete differential, is a rotational scalar. It is the fact that only certain combinations of 4-vectors and 4-gradients appear in the equations of physics that allows these equations to remain invariant in form from one reference frame to another.

It is time to approach this topic, which is the mathematical foundation on which special and general relativity is built, on a firmer and more systematic footing.

## 2.3 Transformation matrix

We begin with a simple but critical notational convention: repeated indices are summed over, unless otherwise explicitly stated. This is known as the *Einstein summation convention*, invented to avoid tedious repeated summation  $\Sigma$ 's. For example:

$$dx^\alpha \frac{\partial}{\partial x^\alpha} = dt \frac{\partial}{\partial t} + dx \frac{\partial}{\partial x} + dy \frac{\partial}{\partial y} + dz \frac{\partial}{\partial z} \quad (44)$$

I will often further shorten this to  $dx^\alpha \partial_\alpha$ . This brings us to another important notational convention. I was careful to write  $\partial_\alpha$ , not  $\partial^\alpha$ . Superscripts will be reserved for vectors, like  $dx^\alpha$  which transform like (5) through (8) from one frame to another (primed) frame moving a relative velocity  $v$  along the  $x$  axis. Subscripts will be used to indicate vectors that transform like the gradient components in equations (34)–(37). Superscript vectors like  $dx^\alpha$  are referred to as *contravariant* vectors; subscripted vectors as *covariant*. (The names will acquire significance later.) The co- contra- difference is an important distinction in general relativity, and we begin by respecting it here in special relativity.

Notice that we can write equations (38) and (39) as

$$[-cdt'] = \gamma([-cdt] + \beta dx) \quad (45)$$

$$dx' = \gamma(dx + \beta[-cdt]) \quad (46)$$

so that the 4-vector  $(-cdt, dx, dy, dz)$  is covariant, like a gradient! We therefore have

$$dx^\alpha = (cdt, dx, dy, dz) \quad (47)$$

$$dx_\alpha = (-cdt, dx, dy, dz) \quad (48)$$

It is easy to go between covariant and contravariant forms by flipping the sign of the time component. We are motivated to formalise this by introducing a matrix  $\eta_{\alpha\beta}$  defined as

$$\eta_{\alpha\beta} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (49)$$

Then  $dx_\alpha = \eta_{\alpha\beta} dx^\beta$  “lowers the index.” We will write  $\eta^{\alpha\beta}$  to raise the index, though it is a numerically identical matrix. Note that the invariant spacetime interval may be written

$$c^2 d\tau^2 \equiv c^2 dt^2 - dx^2 - dy^2 - dz^2 = -\eta_{\alpha\beta} dx^\alpha dx^\beta \quad (50)$$

The time interval  $d\tau$  is just the “proper time,” the time shown ticking on the clock in the rest frame moving with the object of interest (since in this frame all spatial differentials  $dx^i$  are zero). Though introduced as a bookkeeping device,  $\eta_{\alpha\beta}$  is an important quantity: it goes from being a constant matrix in special relativity to a function of coordinates in general relativity, mathematically embodying the departures of spacetime from simple Minkowski form when matter is present.

The standard Lorentz transformation may now be written as a matrix equation,  $dx'^\alpha = \Lambda^\alpha_\beta dx^\beta$ , where

$$\Lambda^\alpha_\beta dx^\beta = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} dx^0 \\ dx^1 \\ dx^2 \\ dx^3 \end{pmatrix} \quad (51)$$

$\Lambda^\alpha_\beta$  is symmetric in  $\alpha$  and  $\beta$ . (A possible notational ambiguity is difficult to avoid here:  $\beta$  and  $\gamma$  used as subscripts or superscripts are of course never velocity variables!) Direct matrix multiplication gives:

$$\Lambda^\alpha_\beta \Lambda^\epsilon_\gamma \eta_{\alpha\epsilon} = \eta_{\beta\gamma} \quad (52)$$

(Do it, and notice that the  $\eta$  matrix must go in the middle...why?) Then, if  $V^\alpha$  is any contravariant vector and  $W_\alpha$  any covariant vector,  $V^\alpha W_\alpha$  must be an invariant (or “scalar”) because

$$V'^\alpha W'_\alpha = V'^\alpha W'^\beta \eta_{\beta\alpha} = \Lambda^\alpha_\gamma V^\gamma \Lambda^\beta_\epsilon W^\epsilon \eta_{\beta\alpha} = V^\gamma W^\epsilon \eta_{\gamma\epsilon} = V^\gamma W_\gamma \quad (53)$$

For covariant vectors, for example  $\partial_\alpha$ , the transformation is  $\partial'_\alpha = \tilde{\Lambda}^\beta_\alpha \partial_\beta$ , where  $\tilde{\Lambda}^\beta_\alpha$  is the same as  $\Lambda^\beta_\alpha$ , but with the sign of  $\beta$  reversed:

$$\tilde{\Lambda}^\alpha_\beta = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (54)$$

Note that

$$\tilde{\Lambda}^\alpha_\beta \Lambda^\beta_\gamma = \delta^\alpha_\gamma, \quad (55)$$

where  $\delta^\alpha_\gamma$  is the Kronecker delta function. This leads immediately once again to  $V'^\alpha W'_\alpha = V^\alpha W_\alpha$ .

Notice that equation (38) says something rather interesting in terms of 4-vectors. The right side is just proportional to  $-dx^\alpha U_\alpha$ , where  $U_\alpha$  is the (covariant) 4-vector corresponding to ordinary velocity  $v$ . Consider now the case  $dt' = 0$ , a surface in  $t, x, y, z$ , spacetime corresponding to simultaneity in the frame of an observer moving at velocity  $\mathbf{v}$ . The equations of constant time in this frame are given by the requirement that  $dx^\alpha$  and  $U_\alpha$  are orthogonal.

*Exercise.* Show that the general Lorentz transformation matrix is:

$$\Lambda^\alpha_\beta = \begin{pmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & 1 + (\gamma - 1)\beta_x^2/\beta^2 & (\gamma - 1)\beta_x\beta_y/\beta^2 & (\gamma - 1)\beta_x\beta_z/\beta^2 \\ -\gamma\beta_y & (\gamma - 1)\beta_x\beta_y/\beta^2 & 1 + (\gamma - 1)\beta_y^2/\beta^2 & (\gamma - 1)\beta_y\beta_z/\beta^2 \\ -\gamma\beta_z & (\gamma - 1)\beta_x\beta_z/\beta^2 & (\gamma - 1)\beta_y\beta_z/\beta^2 & 1 + (\gamma - 1)\beta_z^2/\beta^2 \end{pmatrix} \quad (56)$$

Hint: Keep calm and use (10) and (11).

## 2.4 Tensors

There is more to relativistic life than vectors and scalars. There are objects called tensors, with more than one indexed component. But possessing indices isn't enough! All tensor components must transform in the appropriate way under a Lorentz transformation. To play off an example from Prof. S. Blundell, I could make a matrix of grocery prices with a row of dairy products ( $a_{11}$  = milk,  $a_{12}$  = butter), and a row of vegetables ( $a_{21}$  = carrots,  $a_{22}$  = spinach). If I put this collection in my shopping cart and push the cart at some velocity  $v$ , I shouldn't expect the prices to change by a Lorentz transformation!

A tensor  $T^{\alpha\beta}$  must transform according to the rule

$$T'^{\alpha\beta} = \Lambda^\alpha_\gamma \Lambda^\beta_\epsilon T^{\gamma\epsilon}, \quad (57)$$

while

$$T'_{\alpha\beta} = \tilde{\Lambda}^\gamma_\alpha \tilde{\Lambda}^\epsilon_\beta T_{\gamma\epsilon}, \quad (58)$$

and of course

$$T'^\alpha_\beta = \Lambda^\alpha_\gamma \tilde{\Lambda}^\epsilon_\beta T^\gamma_\epsilon, \quad (59)$$

You get the idea. Contravariant superscript use  $\Lambda$ , covariant subscript use  $\tilde{\Lambda}$ .

Tensors are not hard to find. Remember equation (52)? It works for  $\tilde{\Lambda}^\alpha_\beta$  as well, since it doesn't depend on the sign of  $\beta$  (or its magnitude for that matter):

$$\tilde{\Lambda}^\alpha_\beta \tilde{\Lambda}^\epsilon_\gamma \eta_{\alpha\epsilon} = \eta_{\beta\gamma} \quad (60)$$

So  $\eta_{\alpha\beta}$  is a tensor, with the same components in any frame! The same is true of  $\delta^\alpha_\beta$ , a *mixed* tensor (which is the reason for writing its indices as we have), that we must transform as follows:

$$\Lambda^\epsilon_\gamma \tilde{\Lambda}^\alpha_\beta \delta^\gamma_\alpha = \Lambda^\epsilon_\gamma \tilde{\Lambda}^\gamma_\beta = \delta^\epsilon_\beta. \quad (61)$$

Here is another tensor, slightly less trivial:

$$W^{\alpha\beta} = U^\alpha U^\beta \quad (62)$$

where the  $U$ 's are 4-velocities. This obviously transforms as tensor, since each  $U$  obeys its own vector transformation law. Consider next the tensor

$$T^{\alpha\beta} = \rho_r \langle u^\alpha u^\beta \rangle \quad (63)$$

where the  $\langle \rangle$  notation indicates an average of all the 4-velocity products  $u^\alpha u^\beta$  taken over a whole swarm of little particles, like a gas. (An average of 4-velocities is certainly itself a 4-velocity, and an average of all the little particle tensors is itself a tensor.)  $\rho_r$  is a local rest density, a scalar number. (Here,  $r$  is not an index.)

The component  $T^{00}$  is just  $\rho c^2$ , the energy density of the swarm, where  $\rho$  (without the  $r$ ) includes both a rest mass energy and a thermal contribution. (The thermal part comes from averaging the "swarming"  $\gamma$  factors in the  $u^0 = \gamma c$ .) Moreover, if, as we shall assume, the particle velocities are isotropic, then  $T^{\alpha\beta}$  vanishes if  $\alpha \neq \beta$ . Finally, when  $\alpha = \beta \neq 0$ , then  $T^{ii}$  (no sum!) is by definition the pressure  $P$  of the swarm. (Do you see how this works

out with the  $\gamma$  factors when the  $u^i$  are relativistic?) Hence, in the frame in which the *swarm* has no net translational *bulk* motion,

$$T^{\alpha\beta} = \begin{pmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix} \quad (64)$$

This is, in fact, the most general form for the so-called *energy-momentum stress tensor* for an isotropic fluid in the rest frame of the fluid.

To find  $T^{\alpha\beta}$  in any frame with 4-velocity  $U^\alpha$  we could adopt a brute force method, boost away, and apply the  $\Lambda$  matrix twice to the rest frame form, but what a waste of effort that would be! Here is a better idea. If we can find *any* true tensor that reduces to our result in the rest frame, then that tensor is the *unique* stress tensor. Proof: if a tensor is zero in any frame, then it is zero in all frames, as a trivial consequence of the transformation law. Suppose the tensor I construct, which is designed to match the correct rest frame value, may not be (you claim) correct in all frames. Hand me your tensor, the one you think is the correct choice. Now, the two tensors by definition match in the rest frame. I'll subtract one from the other to form the difference between my tensor and your tensor. The difference is also a tensor, but it vanishes in the rest frame by construction. Hence this “difference tensor” must vanish in all frames, so your tensor and mine are identical after all! Corollary: if you can prove that the two tensors are the same in any one particular frame, then they are the same in all frames. This is a very useful ploy.

The only two tensors we have at our disposal to construct  $T^{\alpha\beta}$  are  $\eta^{\alpha\beta}$  and  $U^\alpha U^\beta$ , and there is only one linear superposition that matches the rest frame value and does the trick:

$$T^{\alpha\beta} = P\eta^{\alpha\beta} + (\rho + P/c^2)U^\alpha U^\beta \quad (65)$$

This is the general form of energy-momentum stress tensor appropriate to an ideal fluid.

#### 2.4.1 Conservation of $T^{\alpha\beta}$

One of the most salient properties of  $T^{\alpha\beta}$  is that it is *conserved*, in the sense of

$$\frac{\partial T^{\alpha\beta}}{\partial x^\alpha} = 0 \quad (66)$$

Since gradients of tensors transform as tensors, this must be true in all frames. What, exactly, are we conserving?

First, the time-like 0-component of this equation is

$$\frac{\partial}{\partial t} \left[ \gamma^2 \left( \rho + \frac{Pv^2}{c^4} \right) \right] + \nabla \cdot \left[ \gamma^2 \left( \rho + \frac{P}{c^2} \right) \mathbf{v} \right] = 0 \quad (67)$$

which is the relativistic version of mass conservation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (68)$$

Elevated in special relativity, it becomes a statement of *energy conservation*. So one of the things we are conserving is energy. (And not just rest mass energy by the way, thermal energy as well!) This is good.

The spatial part of the conservation equation reads

$$\frac{\partial}{\partial t} \left[ \gamma^2 \left( \rho + \frac{P}{c^2} \right) v_i \right] + \left( \frac{\partial}{\partial x^j} \right) \left[ \gamma^2 \left( \rho + \frac{P}{c^2} \right) v_i v_j \right] + \frac{\partial P}{\partial x^i} = 0 \quad (69)$$

You may recognise this as Euler's equation of motion, a statement of momentum conservation, upgraded to special relativity. Conserving momentum is also good.

What if there are other external forces? The idea is that these are included by expressing them in terms of the divergence of their own stress tensor. Then it is the total  $T^{\alpha\beta}$  including, say, electromagnetic fields, that comes into play. What about the force of gravity? *That*, it will turn out, is on an all-together different footing.

You start now to gain a sense of the difficulty in constructing a theory of gravity compatible with relativity. The density  $\rho$  is part of the stress tensor, and it is the entire stress tensor in a relativistic theory that would have to be the source of the gravitational field, just as the entire 4-current  $J^\alpha$  is the source of electromagnetic fields. No fair just picking the component you want. Relativistic theories work with scalars, vectors and tensors to preserve their invariance properties from one frame to another. This insight is already an achievement: we can, for example, expect pressure to play a role in generating gravitational fields. Would you have guessed that? Our relativistic gravity equation maybe ought to look something like :

$$\nabla^2 G^{\mu\nu} - \frac{1}{c^2} \frac{\partial^2 G^{\mu\nu}}{\partial t^2} = T^{\mu\nu} \quad (70)$$

where  $G^{\mu\nu}$  is some sort of, I don't know, conserved tensor guy for the...spacetime geometry and stuff? In Maxwell's theory we had a 4-vector ( $A^\alpha$ ) operated on by the so-called "d'Alembertian operator"  $\nabla^2 - (1/c)^2 \partial^2 / \partial t^2$  on the left side of the equation and a source ( $J^\alpha$ ) on the right. So now we just need to find a  $G^{\mu\nu}$  tensor to go with  $T^{\mu\nu}$ . Right?

Actually, this really is a pretty good guess. It is more-or-less correct for weak fields, and most of the time gravity *is* a weak field. But...well...patience. One step at a time.

*Then there occurred to me the ‘glücklichste Gedanke meines Lebens,’ the happiest thought of my life, in the following form. The gravitational field has only a relative existence in a way similar to the electric field generated by magnetoelectric induction. Because <sup>1</sup> for an observer falling freely from the roof of a house there exists—at least in his immediate surroundings—no gravitational field.*

— Albert Einstein

### 3 The effects of gravity

The central idea of general relativity is that presence of mass (more precisely the presence of any stress-energy tensor component) causes departures from flat Minkowski spacetime to appear, and that other matter (or radiation) responds to these distortions in some way. There are then really two questions: (i) How does the affected matter/radiation move in the presence of a distorted spacetime?; and (ii) How does the stress-energy tensor distort the spacetime in the first place? The first question is purely computational, and fairly straightforward to answer. It lays the groundwork for answering the much more difficult second question, so let us begin here.

#### 3.1 The Principle of Equivalence

We have discussed the notion that by going into a frame of reference that is in free-fall, the effects of gravity disappear. In this era in which space travel is common, we are all familiar with astronauts in free-fall orbits, and the sense of weightlessness that is produced. This manifestation of the Equivalence Principle is so palpable that hearing total mishmashes like “In orbit there is no gravity” from an over-eager science correspondent is a common experience. (Our own BBC correspondent in Oxford Astrophysics, Prof. Christopher Lintott, would certainly *never* say such a thing.)

The idea behind the equivalence principle is that the  $m$  in  $F = ma$  and the  $m$  in the force of gravity  $F_g = mg$  are the same  $m$  and thus the acceleration caused by gravity,  $g$ , is invariant for any mass. We could imagine, for example, that  $F = m_I a$  and  $F_g = m_g g$ , where  $m_g$  is some kind of “massy” property that might vary from one type of body to another with the same  $m_I$ . In this case, the acceleration  $a$  is  $m_g g / m_I$ , i.e., it varies with the ratio of inertial to gravitational mass from one body to another. How well can we actually measure this ratio, or what is more to the point, how well do we know that it is truly a universal constant for all types of matter?

The answer is very, very well indeed. We don’t of course do anything as crude as directly measure the rate at which objects fall to the ground any more, à la Galileo and the tower of Pisa. As with all classic precision gravity experiments (including those of Galileo!) we

---

<sup>1</sup>With apologies to any readers who may actually have fallen off the roof of a house—safe space statement.



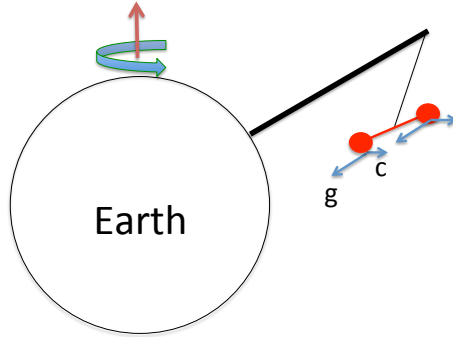


Figure 1: Schematic diagram of the Eötvös experiment. (Not to scale!) A barbell shape, the red object above, is hung from a pendulum on the surface of a rotating Earth, with two masses of two different types of material, say copper and lead. Each mass is affected by gravity ( $g$ ), pulling it to the centre of the earth with a force proportional to a *gravitational mass*  $m_g$ , and a centrifugal force ( $c$ ) proportional to the *inertial mass*  $m_I$ , due to the earth's rotation. Forces are shown as blue arrows, rotation axis as a maroon arrow. Any difference between the inertial to gravitational mass ratio (in copper and lead here) will produce an unbalanced torque about the axis of the suspending fibre of the barbell, arising from the  $g$  and  $c$  forces.

use a pendulum. The first direct measurement of the gravitational to inertial mass actually predates relativity, the so-called Eötvös experiment, after Baron Loránd Eötvös (1848-1919). (Pronounced “ötvösh” with ö as in German *schön*.)

The idea is shown in schematic form in figure [1]. Hang a pendulum from a string, but instead of hanging a big mass, hang a rod, and put two masses of two different types of material at either end. There is a force of gravity toward the center of the earth ( $g$  in the figure), and a centrifugal force ( $c$ ) due to the earth's rotation. The net force is the vector sum of these two, and if the components of the acceleration perpendicular to the string of each mass do not precisely balance, and they won't if  $m_g/m_I$  is not the same for both masses, there will be a net torque twisting the masses about the string (a quartz fibre in the actual experiment). The fact that no such twist is measured is an indication that the ratio  $m_g/m_I$  does not, in fact, vary. In practise, to achieve high accuracy, the pendulum rotates with a tightly controlled period, so that the masses would be sometimes hindered by any putative torque, sometimes pushed forward by this torque. This would imprint a frequency dependence onto the motion, and the resulting signal component at a particular frequency can be very sensitively constrained. Experiment shows that the ratio between any difference in the twisting accelerations on either mass and the average acceleration must be less than a few parts in  $10^{12}$  (Su et al. 1994, Phys Rev D, 50, 3614). With direct laser

ranging experiments to track the Moon's orbit, it is possible, in effect, to use the Moon and Earth as the masses on the pendulum as they rotate around the Sun! This gives an accuracy an order of magnitude better, a part in  $10^{13}$  (Williams et al. 2012, *Class. Quantum Grav.*, 29, 184004), an accuracy comparable to measuring the distance to the Sun to within the size of your thumbnail.

There are two senses in which the Equivalence Principle may be used, a strong sense and weak sense. The weak sense is that it is not possible to detect the effects of gravity locally in a freely falling coordinate system, that all matter behaves identically in a gravitational field independent of its composition. Experiments can test this form of the Principle directly. The strong, much more powerful sense, is that *all physical laws*, gravitational or not, behave in a freely falling coordinate system just as they do in Minkowski spacetime. In this sense, the Principle is a postulate which appears to be true.

If going into a freely falling frame eliminates gravity locally, then going from an inertial frame to an accelerating frame reverses the process and mimics the effect of gravity—again, locally. After all, if in an inertial frame

$$\frac{d^2x}{dt^2} = 0, \quad (71)$$

and we transform to the accelerating frame  $x'$  by  $x = x' + gt^2/2$ , where  $g$  is a constant, then

$$\frac{d^2x'}{dt^2} = -g, \quad (72)$$

which looks an awful lot like motion in a gravitational field.

One immediate consequence of this realisation is of profound importance: gravity affects light. In particular, if we are in an elevator of height  $h$  in a gravitational field of local strength  $g$ , *locally the physics is exactly the same as if we were accelerating upwards at  $g$* . But the effect of this on light is then easily analysed: a photon released upwards reaches a detector at height  $h$  in a time  $h/c$ , at which point the detector is moving at a velocity  $gh/c$  relative to the bottom of the elevator (at the time of release). The photon is measured to be redshifted by an amount  $gh/c^2$ , or  $\Phi/c^2$  with  $\Phi$  being the gravitational potential per unit mass at  $h$ . This is the classical gravitational redshift, the simplest nontrivial prediction of general relativity. The gravitational redshift has been measured accurately using changes in gamma ray energies (RV Pound & JL Snider 1965, *Phys. Rev.*, **140 B**, 788).

The gravitational redshift is the critical link between Newtonian theory and general relativity. It is not, after all, a distortion of space that gives rise to Newtonian gravity at the level we are familiar with, it is a distortion of the flow of *time*.

### 3.2 The geodesic equation

We denote by  $\xi^\alpha$  our freely falling inertial coordinate frame in which the effects of gravity are locally absent. In this frame, the equation of motion for a particle is

$$\frac{d^2\xi^\alpha}{d\tau^2} = 0 \quad (73)$$

with

$$c^2 d\tau^2 = -\eta_{\alpha\beta} d\xi^\alpha d\xi^\beta \quad (74)$$

being the invariant time interval. (If we are doing light, then  $d\tau = 0$ , but ultimately it doesn't really matter. Either take a limit from finite  $d\tau$ , or use any other parameter you fancy, like your wristwatch. In the end, we won't use  $\tau$  or your watch. As for  $d\xi^\alpha$ , it is just the freely-falling guy's ruler and *his* wristwatch.) Next, write this equation in any other set of coordinates you like, and call them  $x^\mu$ . Our inertial coordinates  $\xi^\alpha$  will be some function or other of the  $x^\mu$  so

$$0 = \frac{d^2\xi^\alpha}{d\tau^2} = \frac{d}{d\tau} \left( \frac{\partial\xi^\alpha}{\partial x^\mu} \frac{dx^\mu}{d\tau} \right) \quad (75)$$

where we have used the chain rule to express  $d\xi^\alpha/d\tau$  in terms of  $dx^\mu/d\tau$ . Carrying out the differentiation,

$$0 = \frac{\partial\xi^\alpha}{\partial x^\mu} \frac{d^2x^\mu}{d\tau^2} + \frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \quad (76)$$

where now the chain rule has been used on  $\partial\xi^\alpha/\partial x^\mu$ . This may not look very promising. But if we multiply this equation by  $\partial x^\lambda/\partial\xi^\alpha$ , and remember to sum over  $\alpha$  now, then the chain rule in the form

$$\frac{\partial x^\lambda}{\partial\xi^\alpha} \frac{\partial\xi^\alpha}{\partial x^\mu} = \delta_\mu^\lambda \quad (77)$$

rescues us. (We are using the chain rule repeatedly and will certainly continue to do so, again and again. Make sure you understand this, *and* that you understand what variables are being held constant when the partial derivatives are taken. Deciding what is constant is just as important as doing the differentiation!) Our equation becomes

$$\frac{d^2x^\lambda}{d\tau^2} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0, \quad (78)$$

where

$$\Gamma_{\mu\nu}^\lambda = \frac{\partial x^\lambda}{\partial\xi^\alpha} \frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu} \quad (79)$$

is known as the *affine connection*, and is a quantity of central importance in the study of Riemannian geometry and relativity theory in particular. You should be able to prove, using the chain rule of partial derivatives, an identity for the second derivatives of  $\xi^\alpha$  that we will use shortly:

$$\frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu} = \frac{\partial\xi^\alpha}{\partial x^\lambda} \Gamma_{\mu\nu}^\lambda \quad (80)$$

(How does this work out when used in equation [76]?)

No need to worry, despite the intimidating notation. (Early relativity texts liked to use **gothic font**  $\mathfrak{W}_{\mu\nu}^\lambda$  for the affine connection, which must have imbued it with a nice steampunk terror.) There is nothing especially mysterious about the affine connection. You use it all the time, probably without realising it. For example, in cylindrical  $(r, \theta)$  coordinates, when you use the combinations  $\ddot{r} - r\dot{\theta}^2$  or  $r\ddot{\theta} + 2\dot{r}\dot{\theta}$  for your radial and tangential accelerations, you are using the affine connection and the geodesic equation. In the first case,  $\Gamma_{\theta\theta}^r = -r$ ; in the second,  $\Gamma_{r\theta}^\theta = 1/r$ . (What happened to the 2?)

*Exercise.* Prove the last statements using  $\xi^x = r \cos \theta$ ,  $\xi^y = r \sin \theta$ .

*Exercise.* On the surface of a unit-radius sphere, choose any point as your North Pole, work in colatitude  $\theta$  and azimuth  $\phi$  coordinates, and show that *locally* near the North Pole  $\xi^x = \theta \cos \phi$ ,  $\xi^y = \theta \sin \phi$ . It is in this sense that the  $\xi^\alpha$  coordinates are tied to a local region of the *space* near the North Pole point. In our freely-falling coordinate system, the local coordinates are tied to a point in *spacetime*.

### 3.3 The metric tensor

In our locally inertial coordinates, the invariant spacetime interval is

$$c^2 d\tau^2 = -\eta_{\alpha\beta} d\xi^\alpha d\xi^\beta, \quad (81)$$

so that in any other coordinates,  $d\xi^\alpha = (\partial\xi^\alpha/\partial x^\mu)dx^\mu$  and

$$c^2 d\tau^2 = -\eta_{\alpha\beta} \frac{\partial\xi^\alpha}{\partial x^\mu} \frac{\partial\xi^\beta}{\partial x^\nu} dx^\mu dx^\nu \equiv -g_{\mu\nu} dx^\mu dx^\nu \quad (82)$$

where

$$g_{\mu\nu} = \eta_{\alpha\beta} \frac{\partial\xi^\alpha}{\partial x^\mu} \frac{\partial\xi^\beta}{\partial x^\nu} \quad (83)$$

is known as the *metric tensor*. The metric tensor embodies the information of how coordinate differentials combine to form the invariant interval of our spacetime, and once we know  $g_{\mu\nu}$ , we know everything, including (as we shall see) the affine connections  $\Gamma_{\mu\nu}^\lambda$ . The object of general relativity theory is to compute  $g_{\mu\nu}$  for a given distribution of mass (more precisely, a given stress energy tensor), and a key goal of this course is to find the field equations that enable us to do so.

### 3.4 The relationship between the metric tensor and affine connection

Because of the explicit reliance on the *local* freely falling inertial coordinates  $\xi^\alpha$ , the  $g_{\mu\nu}$  and  $\Gamma_{\mu\nu}^\lambda$  quantities are awkward to use in their present formulation. Fortunately, there is a direct relationship between  $\Gamma_{\mu\nu}^\lambda$  and the first derivatives of  $g_{\mu\nu}$  that will allow us to become free of local bondage, permitting us to dispense with the  $\xi^\alpha$  altogether. Though their *existence* is crucial to formulate the mathematical structure, the practical necessity of the  $\xi$ 's to carry out calculations is minimal.

Differentiate equation (83):

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} = \eta_{\alpha\beta} \frac{\partial^2 \xi^\alpha}{\partial x^\lambda \partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu} + \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial^2 \xi^\beta}{\partial x^\lambda \partial x^\nu} \quad (84)$$

Now use (80) for the second derivatives of  $\xi$ :

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} = \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\rho} \frac{\partial \xi^\beta}{\partial x^\nu} \Gamma_{\lambda\mu}^\rho + \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\rho} \Gamma_{\lambda\nu}^\rho \quad (85)$$

All remaining  $\xi$  derivatives may be absorbed as part of the metric tensor, leading to

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} = g_{\rho\nu} \Gamma_{\lambda\mu}^\rho + g_{\mu\rho} \Gamma_{\lambda\nu}^\rho \quad (86)$$

It remains only to unweave the  $\Gamma$ 's from the cloth of indices. This is done by first adding  $\partial g_{\lambda\nu}/\partial x^\mu$  to the above, then subtracting the same with indices  $\mu$  and  $\nu$  reversed.

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\lambda\mu}}{\partial x^\nu} = g_{\rho\nu} \Gamma_{\lambda\mu}^\rho + \cancel{g_{\mu\rho} \Gamma_{\lambda\nu}^\rho} + g_{\rho\nu} \Gamma_{\mu\lambda}^\rho + \cancel{g_{\rho\lambda} \Gamma_{\mu\nu}^\rho} - \cancel{g_{\rho\mu} \Gamma_{\nu\lambda}^\rho} - \cancel{g_{\rho\lambda} \Gamma_{\nu\mu}^\rho} \quad (87)$$

Remembering that  $\Gamma$  is symmetric in its bottom indices, only the  $g_{\rho\nu}$  terms survive, leaving

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\lambda\mu}}{\partial x^\nu} = 2g_{\rho\nu}\Gamma_{\mu\lambda}^\rho \quad (88)$$

Our last step is to multiply by the inverse matrix  $g^{\nu\sigma}$ , defined by

$$g^{\nu\sigma}g_{\rho\nu} = \delta_\rho^\sigma, \quad (89)$$

leaving us with the pretty result

$$\Gamma_{\mu\lambda}^\sigma = \frac{g^{\nu\sigma}}{2} \left( \frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\lambda\mu}}{\partial x^\nu} \right). \quad (90)$$

Notice that there is no mention of the  $\xi$ 's. The affine connection is completely specified by  $g^{\mu\nu}$  and the derivatives of  $g_{\mu\nu}$  in whatever coordinates you like. In practise, the inverse matrix is not difficult to find, as we will usually work with metric tensors whose off diagonal terms vanish. (Gain confidence once again by practising the geodesic equation with cylindrical coordinates  $g_{rr} = 1$ ,  $g_{\theta\theta} = r^2$  and using [90].) Note as well that with some very simple index relabeling, equation (88) leads directly to the mathematical identity

$$g_{\rho\nu}\Gamma_{\mu\lambda}^\rho \frac{dx^\mu}{d\tau} \frac{dx^\lambda}{d\tau} = \left( \frac{\partial g_{\mu\nu}}{\partial x^\lambda} - \frac{1}{2} \frac{\partial g_{\lambda\mu}}{\partial x^\nu} \right) \frac{dx^\mu}{d\tau} \frac{dx^\lambda}{d\tau}. \quad (91)$$

We'll use this in a moment.

*Exercise.* Prove that  $g^{\nu\sigma}$  is given explicitly by

$$g^{\nu\sigma} = \eta^{\alpha\beta} \frac{\partial x^\nu}{\partial \xi^\alpha} \frac{\partial x^\sigma}{\partial \xi^\beta}$$

*Exercise.* Prove the identities on page 6 of the notes for a diagonal metric  $g_{ab}$ ,

$$\Gamma_{ba}^a = \Gamma_{ab}^a = \frac{1}{2g_{aa}} \frac{\partial g_{aa}}{\partial x^b} \quad (a = b \text{ permitted, NO SUM})$$

$$\Gamma_{bb}^a = -\frac{1}{2g_{aa}} \frac{\partial g_{bb}}{\partial x^a} \quad (a \neq b, \text{ NO SUM})$$

$$\Gamma_{bc}^a = 0, \quad (a, b, c \text{ distinct})$$

### 3.5 Variational calculation of the geodesic equation

The physical significance of the relationship between the metric tensor and affine connection may be understood by a variational calculation. Of all possible paths in our spacetime from some point  $A$  to another  $B$ , which leaves the proper time an extremum (in this case, a maximum)? The motivation for this formulation is obvious: "The shortest distance between two points is a straight line," and the equations for this line-geodesic are  $d^2\xi_i/ds^2 = 0$  in Cartesian coordinates. This is an elementary property of Euclidian space. We may ask what is the shortest distance between two points in a more general curved space as well, and this question naturally lends itself to a variational approach. What is less obvious is that this mathematical machinery, which was fashioned for generalising the spacelike straight line

equation  $d^2\xi^i/ds^2 = 0$  to more general non-Euclidian geometries, also works for generalising a *dynamical* equation of the form  $d^2\xi^i/d\tau^2 = 0$ , where now we are using invariant *timelike* intervals, to geodesics embedded in distorted Minkowski geometries.

We describe our path by some external parameter  $p$ , which could be anything really, perhaps the time on your very own wristwatch in your rest frame. (I don't want to start with  $\tau$ , because  $d\tau = 0$  for light.) Then the proper time from  $A$  to  $B$  is

$$T_{AB} = \int_A^B \frac{d\tau}{dp} dp = \frac{1}{c} \int_A^B \left( -g_{\mu\nu} \frac{dx^\mu}{dp} \frac{dx^\nu}{dp} \right)^{1/2} dp \quad (92)$$

Next, vary  $x^\lambda$  to  $x^\lambda + \delta x^\lambda$  (we are regarding  $x^\lambda$  as a function of  $p$  remember), with  $\delta x^\lambda$  vanishing at the end points  $A$  and  $B$ . We find

$$\delta T_{AB} = \frac{1}{2c} \int_A^B \left( -g_{\mu\nu} \frac{dx^\mu}{dp} \frac{dx^\nu}{dp} \right)^{-1/2} \left( -\frac{\partial g_{\mu\nu}}{\partial x^\lambda} \delta x^\lambda \frac{dx^\mu}{dp} \frac{dx^\nu}{dp} - 2g_{\mu\nu} \frac{d\delta x^\mu}{dp} \frac{dx^\nu}{dp} \right) dp \quad (93)$$

(Do you understand the final term in the integral?)

Since the leading inverse square root in the integrand is just  $dp/d\tau$ ,  $\delta T_{AB}$  simplifies to

$$\delta T_{AB} = \frac{1}{2c} \int_A^B \left( -\frac{\partial g_{\mu\nu}}{\partial x^\lambda} \delta x^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} - 2g_{\mu\nu} \frac{d\delta x^\mu}{d\tau} \frac{dx^\nu}{d\tau} \right) d\tau, \quad (94)$$

and  $p$  has vanished from sight. Should we now worry about light? No, not really. We could simply set  $d\tau = K dp$ , where  $K$  is an arbitrary constant, and take the limit that  $K$  goes to zero at the end of our calculations. We will do something very much like this when we consider particle orbits around a black hole.

We next integrate the second term by parts, noting that the contribution from the end-points has been specified to vanish. Remembering that

$$\frac{dg_{\lambda\nu}}{d\tau} = \frac{dx^\sigma}{d\tau} \frac{\partial g_{\lambda\nu}}{\partial x^\sigma}, \quad (95)$$

we find

$$\delta T_{AB} = \frac{1}{c} \int_A^B \left( -\frac{1}{2} \frac{\partial g_{\mu\nu}}{\partial x^\lambda} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} + \frac{\partial g_{\lambda\nu}}{\partial x^\sigma} \frac{dx^\sigma}{d\tau} \frac{dx^\nu}{d\tau} + g_{\lambda\nu} \frac{d^2 x^\nu}{d\tau^2} \right) \delta x^\lambda d\tau \quad (96)$$

or

$$\delta T_{AB} = \frac{1}{c} \int_A^B \left[ \left( -\frac{1}{2} \frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} \right) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} + g_{\lambda\nu} \frac{d^2 x^\nu}{d\tau^2} \right] \delta x^\lambda d\tau \quad (97)$$

Finally, using equation (91), we obtain

$$\delta T_{AB} = \frac{1}{c} \int_A^B \left[ \left( \frac{dx^\mu}{d\tau} \frac{dx^\sigma}{d\tau} \Gamma_{\mu\sigma}^\nu + \frac{d^2 x^\nu}{d\tau^2} \right) g_{\lambda\nu} \right] \delta x^\lambda d\tau \quad (98)$$

Thus, if the geodesic equation (78) is satisfied,  $\delta T_{AB} = 0$  is satisfied, and the proper time is an extremum. The name “geodesic” is used in geometry to describe the path of minimum distance between two points in a manifold, and it is therefore gratifying to see that there is a correspondence between a local “straight line” with zero curvature, and the local elimination of a gravitational field with the resulting zero acceleration, along the lines of the first paragraph of this section. In the first case, the proper choice of local coordinates results in

the second derivative with respect to an invariant spatial interval vanishing; in the second case, the proper choice of coordinates means that the second derivative with respect to an invariant time interval vanishes, but *the essential mathematics is the same*.

There is often a very practical side to working with the variational method: it can be much easier to obtain the equations of motion for a given  $g_{\mu\nu}$  this way, as opposed to first calculating the  $\Gamma_{\mu\nu}^\lambda$  explicitly. The variational method quickly produces all the non-vanishing affine connection components; just read them off as the coefficients of  $(dx^\mu/d\tau)(dx^\nu/d\tau)$ . These quantities are then available for any variety of purposes.

Here is a nice trick. You should have little difficulty showing that if we apply the Euler-Lagrange variational method directly to the following functional “Lagrangian”  $\mathcal{L}$ ,

$$\mathcal{L} = g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu, \quad (99)$$

where the dot is  $d/d\tau$ , the resulting Euler-Lagrange equation

$$\frac{d}{d\tau} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}^\rho} \right) - \frac{\partial \mathcal{L}}{\partial x^\rho} = 0 \quad (100)$$

is just the standard geodesic equation of motion! This is usually the best way to proceed.

In classical mechanics, we all know that the equations of motion may be derived from a Lagrangian variational principle of least action, an integral involving the difference between kinetic and potential energies. This doesn’t seem geometrical at all. What is the connection with what we’ve just done? How do we make contact with Newtonian mechanics from the geodesic equation?

### 3.6 The Newtonian limit

We consider the case of a slowly moving mass (“slow” of course means relative to  $c$ , the speed of light) in a weak gravitational field ( $GM/rc^2 \ll 1$ ). Since  $cdt \gg |d\mathbf{x}|$ , the geodesic equation greatly simplifies:

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma_{00}^\mu \left( \frac{cdt}{d\tau} \right)^2 = 0. \quad (101)$$

Now

$$\Gamma_{00}^\mu = \frac{1}{2} g^{\mu\nu} \left( \frac{\partial g_{0\nu}}{\partial(cdt)} + \frac{\partial g_{0\nu}}{\partial(cdt)} - \frac{\partial g_{00}}{\partial x^\nu} \right) \quad (102)$$

In the Newtonian limit, the largest of the  $g$  derivatives is the spatial gradient, hence

$$\Gamma_{00}^\mu \simeq -\frac{1}{2} g^{\mu\nu} \frac{\partial g_{00}}{\partial x^\nu} \quad (103)$$

Since the gravitational field is weak,  $g_{\alpha\beta}$  differs very little from the Minkowski value:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \quad h_{\alpha\beta} \ll 1, \quad (104)$$

and the  $\mu = 0$  geodesic equation is

$$\frac{d^2 t}{d\tau^2} + \frac{1}{2} \frac{\partial h_{00}}{\partial t} \left( \frac{dt}{d\tau} \right)^2 = 0 \quad (105)$$

Clearly, the second term is zero for a static field, and will prove to be tiny when the gravitational field changes with time under nonrelativistic conditions—we are, after all, calculating the difference between proper time and observer time! Dropping this term we find that  $t$  and  $\tau$  are linearly related, so that the spatial components of the geodesic equation become

$$\frac{d^2 \mathbf{x}}{dt^2} - \frac{c^2}{2} \nabla h_{00} = 0 \quad (106)$$

Isaac Newton would say:

$$\frac{d^2 \mathbf{x}}{dt^2} + \nabla \Phi = 0, \quad (107)$$

with  $\Phi$  being the classical gravitational potential. The two views are consistent if

$$h_{00} \simeq -\frac{2\Phi}{c^2}, \quad g_{00} \simeq -\left(1 + \frac{2\Phi}{c^2}\right) \quad (108)$$

In other words, the gravitational potential force emerges as a sort of centripetal term, similar in structure to the centripetal force in the standard radial equation of motion. *This is a remarkable result.* It is by no means obvious that a purely geometrical geodesic equation can serve the role of a Newtonian gravitational potential gradient force equation, but it can. Moreover, it teaches us that the Newtonian limit of general relativity is all in the time component,  $h_{00}$ . It is now possible to measure directly the differences in the rate at which clocks run at heights separated by a few meters on the Earth's surface.

The quantity  $h_{00}$  is a dimensionless number of order  $v^2/c^2$ , where  $v$  is a velocity typical of the system, an orbital speed or just the square root of a potential. Note that  $h_{00}$  is determined by the dynamical equations only up to an additive constant. Here we have chosen the constant to make the geometry Minkowskian at large distances from any matter. At the surface of a spherical object of mass  $M$  and radius  $R$ ,

$$h_{00} \simeq 4.2 \times 10^{-6} \left(\frac{M}{M_\odot}\right) \left(\frac{R_\odot}{R}\right) \quad (109)$$

where  $M_\odot$  is the mass of the sun (about  $2 \times 10^{30}$  kg) and  $R_\odot$  is the radius of the sun (about  $7 \times 10^8$  m). The gravitational redshift fraction is half this number. As an exercise, you may wish to look up masses of planets and other types of stars and evaluate  $h_{00}$ . What is its value at the surface of a white dwarf (mass of the sun, radius of the earth)? What about a neutron star (mass of the sun, radius of Oxford)? How many decimal points are needed to see the time difference in two digital clocks at a one meter separation in height on the earth?

We are now able to relate the geodesic equation to the principle of least action in classical mechanics. In the Newtonian limit, our variational integral becomes

$$\int [c^2(1 + 2\Phi/c^2)dt^2 - d|\mathbf{x}|^2]^{1/2} \quad (110)$$

(Remember our compact notation:  $dt^2 \equiv (dt)^2$ ,  $d|\mathbf{x}|^2 = (d|\mathbf{x}|)^2$ .) Expanding the square root,

$$\int c \left(1 + \frac{\Phi}{c^2} - \frac{v^2}{2c^2} + \dots\right) dt \quad (111)$$

where  $v^2 \equiv (d|\mathbf{x}|/dt)^2$ . Thus, minimising the Lagrangian (kinetic energy minus potential energy) is the same as maximising the proper time interval! What an unexpected and beautiful connection.



What we have calculated in this section is nothing more than our old friend the gravitational redshift, with which we began our formal study of general relativity. The invariant spacetime interval  $d\tau$ , the proper time, is given by

$$c^2 d\tau^2 = -g_{\mu\nu} dx^\mu dx^\nu \quad (112)$$

For an observer at rest at location  $x$ , the time interval registered on a clock will be

$$d\tau(x) = [-g_{00}(x)]^{1/2} dt \quad (113)$$

where  $dt$  is the time interval registered at infinity, where  $-g_{00} \rightarrow 1$ . (Compare: the “proper length” on the unit sphere for an interval at constant  $\theta$  is  $\sin\theta d\phi$ , where  $d\phi$  is the length registered by an equatorial observer.) If the interval between two wave crest crossings is found to be  $d\tau(y)$  at location  $y$ , it will be  $d\tau(x)$  when the light reaches  $x$  and it will be  $dt$  at infinity. In general,

$$\frac{d\tau(y)}{d\tau(x)} = \left[ \frac{g_{00}(y)}{g_{00}(x)} \right]^{1/2}, \quad (114)$$

and in particular

$$\frac{d\tau(R)}{dt} = \frac{\nu(\infty)}{\nu} = [-g_{00}(R)]^{1/2} \quad (115)$$

where  $\nu = 1/d\tau(R)$  is, for example, an atomic transition frequency measured at rest at the surface  $R$  of a body, and  $\nu(\infty)$  the corresponding frequency measured a long distance away. Interestingly, the value of  $g_{00}$  that we have derived in the Newtonian limit is, in fact, the *exact* relativistic value of  $g_{00}$  around a point mass  $M$ ! (A black hole.) The precise redshift formula is

$$\nu_\infty = \left( 1 - \frac{2GM}{Rc^2} \right)^{1/2} \nu \quad (116)$$

The redshift as measured by wavelength becomes infinite from light emerging from radius  $R = 2GM/c^2$ , the so-called Schwarzschild radius (about 3 km for a point with the mass of the sun!).

Historically, in its infancy general relativity theory was supported empirically by the reported detection of a gravitational redshift in a spectral line observed from the surface of the white dwarf star Sirius B in 1925 by W.S. Adams. It “killed two birds with one stone,” the leading astronomer A.S. Eddington remarked. For it not only proved the existence of white dwarf stars (at the time controversial since the mechanism of pressure support was unknown), the measurement also confirmed an early and important prediction of general relativity theory: the redshift of light due to gravity.

Alas, the modern consensus is that the actual measurements were flawed! Adams knew what he was looking for and found it. We now call this “confirmation bias.” Though premature, the activity this apparently positive observation imparted to the study of white dwarfs and to relativity theory turned out to be very fruitful indeed. But we were lucky. Well-regarded but incorrect single-investigator astronomical observations have often caused much confusion and needless wrangling, as well as years of wasted effort.

The first definitive test for gravitational redshift came much later, and it was purely terrestrial: the 1959 Pound and Rebka experiment performed at Harvard University’s Jefferson Tower measured the frequency shift of a 14.4 keV gamma ray falling (if that is the word for a gamma ray) 22.6 m. Pound & Rebka were able to measure the shift in energy—just a few parts in  $10^{14}$ —by what was at the time the new and novel technique of Mössbauer spectroscopy.

*Exercise.* A novel application of the gravitational redshift is provided by Bohr's refutation of an argument put forth by Einstein purportedly showing that an experiment could in principle be designed to bypass the quantum uncertainty relation  $\Delta E \Delta t \geq h$ . The idea is to hang a box containing a photon by a spring suspended in a gravitational field  $g$ . At some precise time a shutter is opened and the photon leaves. You weigh the box before and after the photon. There is in principle no interference between the arbitrarily accurate change in box weight and the arbitrarily accurate time at which the shutter is opened. Or is there?

1.) Show that the box apparatus satisfies an equation of the form

$$M\ddot{x} = -Mg - kx$$

where  $M$  is the mass of the apparatus,  $x$  is the displacement, and  $k$  is the spring constant. Before release, the box is in equilibrium at  $x = -gM/k$ .

2.) Show that the momentum of the box after a short time interval  $\Delta t$  from when the photon escapes is

$$\delta p = -\frac{g\delta m}{\omega} \sin(\omega\Delta t) \simeq -g\delta m\Delta t$$

where  $\delta m$  is the (uncertain!) photon mass and  $\omega^2 = k/M$ . With  $\delta p \sim g\delta m\Delta t$ , the uncertainty principle then dictates an uncertain location of the box position  $\delta x$  given by  $g\delta m\delta x\Delta t \sim h$ . But this is location uncertainty, not time uncertainty.

3.) Now the gravitational redshift comes in! The flow of time depends upon location. Show that if there is an uncertainty in position  $\delta x$ , there is an uncertainty in the time of release:  $\delta t \sim (g\delta x/c^2)\Delta t$ .

4.) Finally use this in part (2) to establish  $\delta E \delta t \sim h$  with  $\delta E = \delta mc^2$ .

Why does general relativity come into *nonrelativistic* quantum mechanics in such a fundamental way? Because the gravitational redshift is relativity theory's point-of-contact with classical Newtonian mechanics, and Newtonian mechanics when blended with the uncertainty principle is the start of nonrelativistic quantum mechanics.

### *A final thought*

We Newtonian beings, with our natural mode of thinking in terms of forces and responses, would naturally say "How interesting, the force of gravity distorts the flow of time." This is the way I have been describing the gravitational redshift throughout this chapter. But Einstein has given us a more profound insight. It is not that gravity *distorts* the flow of time. An Einsteinian being, brought up from the cradle to be comfortable with a spacetime point-of-view, would, upon hearing this comment, cock their head and say: "What are you talking about? Newtonian gravity *is* the distortion of the flow of time. It is a simple geometric distortion that is brought about by the presence of matter." This is a better way to think of it. Close to the source, the effect of weak gravity is indeed a distortion in the flow of time; far from the source, the effect of weak gravity is gravitational radiation, and this, we shall see, may be thought of as a distortion of *space*.

## 4 Tensor Analysis

*Further, the dignity of the science seems to require that every possible means be explored itself for the solution of a problem so elegant and so celebrated.*

— Carl Friedrich Gauss

A mathematical equation is valid in the presence of general gravitational fields when

- i.) It is a valid equation in the absence of gravity and respects Lorentz invariance.*
- ii.) It preserves its form, not just under Lorentz transformations, but under any coordinate transformation,  $x \rightarrow x'$ .*

What does “preserves its form” mean? It means that the equation must be written in terms of quantities that transform as scalars, vectors, and higher ranked tensors under general coordinate transformations. From (ii), we see that if we can find one coordinate system in which our equation holds, it will hold in any set of coordinates. But by (i), the equation *does* hold in locally freely falling coordinates, in which the effect of gravity is locally absent. The effect of gravity is strictly embodied in the two key quantities that emerge from the calculus of coordinate transformations: the metric tensor  $g_{\mu\nu}$  and its first derivatives in  $\Gamma_{\mu\nu}^\lambda$ . This approach is known as the *Principle of General Covariance*, and it is a very powerful tool indeed.

### 4.1 Transformation laws

The simplest vector one can write down is the ordinary coordinate differential  $dx^\mu$ . If  $x'^\mu = x'^\mu(x)$ , there is no doubt how the  $dx'^\mu$  are related to the  $dx^\mu$ . It is called the chain rule, and it is by now very familiar:

$$dx'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} dx^\nu \quad (117)$$

Be careful to distinguish between the coordinates  $x^\mu$ , which can be pretty much anything, and their differentials  $dx^\mu$ , which are true vectors. Indeed, any set of quantities  $V^\mu$  that transforms in this way is known as a *contravariant vector*:

$$V'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} V^\nu \quad (118)$$

The contravariant 4-velocity, which is a 4-vector, is simply  $V^\mu = dx^\mu/d\tau$ , a generalisation of the special relativistic  $d\xi^\alpha/d\tau$ . A *covariant vector*, by contrast, transforms as

$$V'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} V_\nu \quad (119)$$

“CO LOW, PRIME BELOW.” (Sorry. Maybe you can do better.) These definitions of contravariant and covariant vectors are consistent with those we first introduced in our

discussions of the Lorentz matrices  $\Lambda_\beta^\alpha$  and  $\tilde{\Lambda}_\alpha^\beta$  in Chapter 2, but now generalised from specific linear transformations to arbitrary transformations.

The simplest covariant vector is the gradient  $\partial/\partial x^\mu$  of a scalar  $\Phi$ . Once again, the chain rule tells us how to transform from one set of coordinates to another—we've no choice:

$$\frac{\partial \Phi}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial \Phi}{\partial x^\nu} \quad (120)$$

The generalisation to tensor transformation laws is immediate. A contravariant tensor  $T^{\mu\nu}$  transforms as

$$T'^{\mu\nu} = \frac{\partial x'^\mu}{\partial x^\rho} \frac{\partial x'^\nu}{\partial x^\sigma} T^{\rho\sigma} \quad (121)$$

a covariant tensor  $T_{\mu\nu}$  as

$$T'_{\mu\nu} = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} T_{\rho\sigma} \quad (122)$$

and a mixed tensor  $T_\nu^\mu$  as

$$T_\nu'^\mu = \frac{\partial x'^\mu}{\partial x^\rho} \frac{\partial x^\sigma}{\partial x'^\nu} T_\sigma^\rho \quad (123)$$

The generalisation to mixed tensors of arbitrary rank should be self-evident.

By this definition the metric tensor  $g_{\mu\nu}$  really *is* a covariant tensor, just as its notation would lead you to believe, because

$$g'_{\mu\nu} \equiv \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x'^\mu} \frac{\partial \xi^\beta}{\partial x'^\nu} = \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\lambda} \frac{\partial \xi^\beta}{\partial x^\rho} \frac{\partial x^\lambda}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} \equiv g_{\lambda\rho} \frac{\partial x^\lambda}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} \quad (124)$$

and the same for the *contravariant*  $g^{\mu\nu}$ . (How else could it be?  $dx^\mu dx^\nu$  is clearly a contravariant tensor, and  $g_{\mu\nu} dx^\mu dx^\nu$  is by construction a scalar, which works only if  $g_{\mu\nu}$  is a covariant tensor.) However, the gradient of a vector is *not*, in general, a tensor or a vector:

$$\frac{\partial V'^\lambda}{\partial x'^\mu} = \frac{\partial}{\partial x'^\mu} \left( \frac{\partial x'^\lambda}{\partial x^\nu} V^\nu \right) = \frac{\partial x'^\lambda}{\partial x^\nu} \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial V^\nu}{\partial x^\rho} + \frac{\partial^2 x'^\lambda}{\partial x^\rho \partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} V^\nu \quad (125)$$

The first term is just what we would have wanted if we were searching for a tensor transformation law. But oh those pesky second order derivatives—the final term spoils it all. This of course vanishes when the coordinate transformation is linear (as when we found that vector derivatives are perfectly good tensors under the Lorentz transformations), but not in general. We will show in the next section that while the gradient of a vector is in general not a tensor, there is an elegant solution around this problem.

Tensors can be created and manipulated in many ways. For example, direct products of tensors are tensors:

$$W_{\rho\sigma}^{\mu\nu} = T^{\mu\nu} S_{\rho\sigma}. \quad (126)$$

A linear combination of tensors of the same rank multiplied by scalars is obviously a tensor of unchanged rank. A tensor can lower its index by multiplying by  $g_{\mu\nu}$  or raise it with  $g^{\mu\nu}$ :

$$T_\mu'^\rho \equiv g'_{\mu\nu} T'^{\nu\rho} = \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\lambda}{\partial x'^\nu} \frac{\partial x'^\nu}{\partial x^\kappa} \frac{\partial x'^\rho}{\partial x^\tau} g_{\sigma\lambda} T^{\kappa\tau} = \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x'^\rho}{\partial x^\tau} g_{\sigma\kappa} T^{\kappa\tau} = \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x'^\rho}{\partial x^\tau} T_\sigma^\tau \quad (127)$$

which indeed does transform as a tensor of mixed second rank,  $T_\mu'^\rho$ . Multiplying  $T^{\mu\nu}$  by *any* covariant tensor  $S_{\rho\mu}$  of course generates a mixed tensor  $M_\rho^\nu$ , but we adopt the very useful

convention of keeping the name  $T_\rho{}^\nu$  when multiplying by  $S_{\rho\mu} = g_{\rho\mu}$ , and thinking of the index as “being lowered.” (And of course index-raising for multiplication by  $g^{\rho\mu}$ .)

Mixed tensors can “contract” to scalars. Start with  $T^\mu{}_\nu$ . Then consider the transformation of  $T^\mu{}_\mu$ :

$$T'^\mu{}_\mu = \frac{\partial x'^\mu}{\partial x^\nu} \frac{\partial x^\rho}{\partial x'^\mu} T^\nu{}_\rho = \delta^\rho_\nu T^\nu{}_\rho = T^\nu{}_\nu \quad (128)$$

i.e.,  $T^\mu{}_\mu$  is a scalar  $T$ . Exactly the same type of calculation shows that  $T^{\mu\nu}$  is a vector  $T^\nu$ , and so on. Remember to contract “up–down:”  $T^\mu{}_\mu = T$ , *not*  $T^{\mu\mu} = T$ .

The generalisation of the familiar scalar dot product between vectors  $A^\mu$  and  $B^\mu$  is  $A^\mu B_\mu = g_{\mu\nu} A^\mu B^\nu$ . We are often interested in just the spatial part of 4-vectors, the 3-vector  $A^i$ . Then, in a non-Euclidian 3-space, the cosine of the angle  $\Delta\theta$  between two vectors may be written as the ratio

$$\cos \Delta\theta = \frac{A^i B_i}{(A^j A_j B^k B_k)^{1/2}} = \frac{g_{ij} A^i B^j}{(g_{kl} A^k A^l g_{mn} B^m B^n)^{1/2}} \quad (129)$$

the analogue of  $\mathbf{A} \cdot \mathbf{B} / (|\mathbf{A}| |\mathbf{B}|)$ . If we are given two parameterised curves, perhaps two orbits  $x^i(p)$  and  $y^i(p)$ , and wish to know the angle between them at some particular point, this angle becomes

$$\cos \Delta\theta = \frac{\dot{x}^i \dot{y}_i}{(\dot{x}^j \dot{x}_j \dot{y}^k \dot{y}_k)^{1/2}} = \frac{g_{ij} \dot{x}^i \dot{y}^j}{(g_{kl} \dot{x}^k \dot{x}^l g_{mn} \dot{y}^m \dot{y}^n)^{1/2}}$$

where the dot notation denotes  $d/dp$ . Do you see why this is so?

## 4.2 The covariant derivative

Knowing the general transformation law for a vector allows us to understand more fully the content of the geodesic equation. With  $V'^\alpha = d\xi^\alpha/d\tau$ , the geodesic equation actually tells us the proper form of the vector quantity to use for  $dV'^\alpha/d\tau$  that is *valid in any coordinate system*. To see this clearly, consider the derivative of any vector  $dA'^\alpha/d\tau$ , expressed here in local inertial coordinates  $\xi^\alpha$ , and in these coordinates this is the right form of the derivative. Then, exactly the same manipulations we used to derive the geodesic equation yield, upon going to arbitrary coordinates  $x^\mu$  and vector  $A^\mu$ :

$$\frac{dA'^\alpha}{d\tau} = \frac{d}{d\tau} \left( \frac{\partial \xi^\alpha}{\partial x^\mu} A^\mu \right) = \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{dA^\mu}{d\tau} + \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu} A^\mu \frac{dx^\nu}{d\tau}.$$

(This is of course always valid, whether  $dA'^\alpha/d\tau$  happens to vanish or not.) Multiplying by  $\partial x^\lambda / \partial \xi^\alpha$ , we find

$$\frac{\partial x^\lambda}{\partial \xi^\alpha} \frac{dA'^\alpha}{d\tau} = \frac{dA^\lambda}{d\tau} + \Gamma^\lambda_{\mu\nu} A^\mu \frac{dx^\nu}{d\tau}. \quad (130)$$

The left hand side is now the formal transformation law of  $dA'^\alpha/d\tau$  from local inertial to general coordinates. The right side tells us precisely what that expression is. Evidently, the combination on the right of this equation is a general vector quantity unto itself.

We will return to this neat result when we discuss the important concept of parallel transport, but for now, let us further note that, with  $V^\nu = dx^\nu/d\tau$ , the vectorial right side of (130) may be written as

$$V^\nu \left[ \frac{\partial A^\lambda}{\partial x^\nu} + \Gamma_{\mu\nu}^\lambda A^\mu \right] \quad (131)$$

Since  $V^\nu$  is a vector and  $A^\lambda$  is arbitrary, the expression in square brackets must be a mixed tensor of rank two: it contracts with a contravariant vector  $V^\nu$  to produce another contravariant vector. Recall that  $\Gamma_{\mu\nu}^\lambda$  vanishes in local inertial coordinates, in which we know that simple partial derivatives of vectors are valid tensors. So this prescription tells us how to upgrade the notion of a partial derivative of a vector to the status of a full tensor: to make a tensor out of a plain partial derivative of a vector, form the quantity

$$\frac{\partial A^\lambda}{\partial x^\mu} + \Gamma_{\mu\nu}^\lambda A^\nu \equiv A_{;\mu}^\lambda \quad (132)$$

the so called *covariant derivative*. Following convention, we use a semicolon to denote covariant differentiation. (Some authors use a comma for partial derivatives (e.g.  $A_{,\mu}^\nu$ ), but it is more clear to use full partial derivative notation [or  $\partial_\mu$ ], and we shall abide by this practise in these notes, if not always in lecture.) We now have a generalisation of the partial derivative to tensor form!

You know, this is really too important a result not to check in detail. We also need how to construct the covariant derivative of covariant vectors, and of more general tensors. (Talk about confusing. Notice the use of the word “covariant” twice in that last statement in two very different senses. Apologies for this awkward, but completely standard, mathematical nomenclature.) If you are already convinced that the covariant derivative really is a tensor, just skip down to right after equation (139). You won’t learn anything more than you already know in the next long paragraph, and there is a lot of calculation.

The first thing we need to do is to establish the transformation law for  $\Gamma_{\mu\nu}^\lambda$ . This is just repeated application of the chain rule:

$$\Gamma_{\mu\nu}^{\prime\lambda} \equiv \frac{\partial x^{\prime\lambda}}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^{\prime\mu} \partial x^{\prime\nu}} = \frac{\partial x^{\prime\lambda}}{\partial x^\rho} \frac{\partial x^\rho}{\partial \xi^\alpha} \frac{\partial}{\partial x^{\prime\mu}} \left( \frac{\partial x^\sigma}{\partial x^{\prime\nu}} \frac{\partial \xi^\alpha}{\partial x^\sigma} \right) \quad (133)$$

Carrying through the derivative,

$$\Gamma_{\mu\nu}^{\prime\lambda} = \frac{\partial x^{\prime\lambda}}{\partial x^\rho} \frac{\partial x^\rho}{\partial \xi^\alpha} \left( \frac{\partial x^\sigma}{\partial x^{\prime\nu}} \frac{\partial x^\tau}{\partial x^{\prime\mu}} \frac{\partial^2 \xi^\alpha}{\partial x^\tau \partial x^\sigma} + \frac{\partial^2 x^\sigma}{\partial x^{\prime\mu} \partial x^{\prime\nu}} \frac{\partial \xi^\alpha}{\partial x^\sigma} \right) \quad (134)$$

Cleaning up, and recognising an affine connection when we see one, helps to rid us of these meddlesome  $\xi$ ’s:

$$\Gamma_{\mu\nu}^{\prime\lambda} = \frac{\partial x^{\prime\lambda}}{\partial x^\rho} \frac{\partial x^\tau}{\partial x^{\prime\mu}} \frac{\partial x^\sigma}{\partial x^{\prime\nu}} \Gamma_{\tau\sigma}^\rho + \frac{\partial x^{\prime\lambda}}{\partial x^\rho} \frac{\partial^2 x^\rho}{\partial x^{\prime\mu} \partial x^{\prime\nu}} \quad (135)$$

This may also be written

$$\Gamma_{\mu\nu}^{\prime\lambda} = \frac{\partial x^{\prime\lambda}}{\partial x^\rho} \frac{\partial x^\tau}{\partial x^{\prime\mu}} \frac{\partial x^\sigma}{\partial x^{\prime\nu}} \Gamma_{\tau\sigma}^\rho - \frac{\partial x^\rho}{\partial x^{\prime\nu}} \frac{\partial x^\sigma}{\partial x^{\prime\mu}} \frac{\partial^2 x^{\prime\lambda}}{\partial x^\sigma \partial x^\rho} \quad (136)$$

Do you see why? (Hint: Either integrate  $\partial/\partial x^{\prime\mu}$  by parts or differentiate the identity

$$\frac{\partial x^{\prime\lambda}}{\partial x^\rho} \frac{\partial x^\rho}{\partial x^{\prime\mu}} = \delta_{\mu}^{\lambda}.)$$

Hence

$$\Gamma_{\mu\nu}^{\lambda} A^{\nu} = \left( \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma_{\tau\sigma}^{\rho} - \frac{\partial x^{\rho}}{\partial x'^{\nu}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \frac{\partial^2 x'^{\lambda}}{\partial x^{\rho} \partial x^{\sigma}} \right) \frac{\partial x'^{\nu}}{\partial x^{\eta}} A^{\eta}, \quad (137)$$

and spotting some tricky sums over  $\partial x'^{\nu}$  that turn into Kronecker delta functions,

$$\Gamma_{\mu\nu}^{\lambda} A^{\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \Gamma_{\tau\sigma}^{\rho} A^{\sigma} - \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \frac{\partial^2 x'^{\lambda}}{\partial x^{\rho} \partial x^{\sigma}} A^{\rho} \quad (138)$$

Finally, adding this to (125), the unwanted terms cancel just as they should. We thus obtain

$$\frac{\partial A'^{\lambda}}{\partial x'^{\mu}} + \Gamma_{\mu\nu}^{\lambda} A^{\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\nu}} \frac{\partial x^{\rho}}{\partial x'^{\mu}} \left( \frac{\partial A^{\nu}}{\partial x^{\rho}} + \Gamma_{\rho\sigma}^{\nu} A^{\sigma} \right), \quad (139)$$

as desired. This combination really does transform as a tensor ought to.

It is now a one-step process to deduce how covariant derivatives work for covariant vectors. With  $V^{\lambda}$  arbitrary, consider

$$V_{\lambda} A^{\lambda}_{;\mu} = V_{\lambda} \frac{\partial A^{\lambda}}{\partial x^{\mu}} + \Gamma_{\mu\nu}^{\lambda} A^{\nu} V_{\lambda} \quad (140)$$

which is a perfectly good covariant vector. Integrating by parts the first term on the right, and then switching dummy indices  $\lambda$  and  $\nu$  in the final term, this expression is identical to

$$\frac{\partial(A^{\lambda} V_{\lambda})}{\partial x^{\mu}} - A^{\lambda} \left[ \frac{\partial V_{\lambda}}{\partial x^{\mu}} - \Gamma_{\mu\lambda}^{\nu} V_{\nu} \right]. \quad (141)$$

Since the first term is the covariant gradient of a scalar, and the entire expression must be a good covariant vector, the term in square brackets must be a purely covariant tensor of rank two. We have very quickly found our generalisation for the covariant derivative of a covariant vector:

$$V_{\lambda;\mu} = \frac{\partial V_{\lambda}}{\partial x^{\mu}} - \Gamma_{\mu\lambda}^{\nu} V_{\nu} \quad (142)$$

That this really *is* a vector can also be directly verified via a calculation exactly similar to our previous one for the covariant derivative of a contravariant vector.

Covariant derivatives of *tensors* are now simple to deduce. The tensor  $T^{\lambda\kappa}$  must formally transform like a contravariant *vector* if we “freeze” one of its indices at some particular component and allow the other to take on all component values. Since the formula must be symmetric in the two indices,

$$T^{\lambda\kappa}_{;\mu} = \frac{\partial T^{\lambda\kappa}}{\partial x^{\mu}} + \Gamma_{\mu\nu}^{\lambda} T^{\nu\kappa} + \Gamma_{\mu\kappa}^{\nu} T^{\lambda\nu} \quad (143)$$

and then it should also follow

$$T_{\lambda\kappa;\mu} = \frac{\partial T_{\lambda\kappa}}{\partial x^{\mu}} - \Gamma_{\lambda\mu}^{\nu} T_{\nu\kappa} - \Gamma_{\kappa\mu}^{\nu} T_{\lambda\nu} \quad (144)$$

and of course

$$T^{\lambda}_{\kappa;\mu} = \frac{\partial T^{\lambda}_{\kappa}}{\partial x^{\mu}} + \Gamma_{\mu\nu}^{\lambda} T^{\nu}_{\kappa} - \Gamma_{\mu\kappa}^{\nu} T^{\lambda}_{\nu} \quad (145)$$

The generalisation to tensors of arbitrary rank should now be self-evident. To generate the affine connection terms, freeze all indices in your tensor, then unfreeze them one-by-one, treating each unfrozen index as either a covariant or contravariant vector, depending upon whether it is down or up. Practise this until it is second-nature.

We now can give a precise rule for how to take an equation that is valid in special relativity, and upgrade it to the general relativistic theory of gravity. *Work exclusively with 4-vectors and 4-tensors. Replace  $\eta_{\alpha\beta}$  with  $g_{\mu\nu}$ . Take standard partial derivatives and turn them into covariant derivatives. Voilà: your equation is set for the presence of gravitational fields.*

It will not have escaped your attention, I am sure, that applying (144) to  $g_{\mu\nu}$  produces

$$g_{\mu\nu;\lambda} = \frac{\partial g_{\mu\nu}}{\partial x^\lambda} - g_{\rho\nu}\Gamma_{\mu\lambda}^\rho - g_{\mu\rho}\Gamma_{\nu\lambda}^\rho = 0 \quad (146)$$

where equation (86) has been used for the last equality. *The covariant derivatives of  $g_{\mu\nu}$  vanish.* This is exactly what we would have predicted, since the partial derivatives of  $\eta_{\alpha\beta}$  vanish in special relativity, and thus the covariant derivative of  $g_{\mu\nu}$  should vanish in the presence of gravitational fields. It's just the general relativistic upgrade of  $\partial\eta_{\alpha\beta}/\partial x^\gamma = 0$ .

Here are two important technical points that are easily shown. (You should do so explicitly.)

- The covariant derivative obeys the Leibniz rule for products. For example:

$$(T^{\mu\nu}U_{\lambda\kappa})_{;\rho} = T^{\mu\nu}_{;\rho}U_{\lambda\kappa} + T^{\mu\nu}U_{\lambda\kappa;\rho},$$

$$(A^\mu V_\mu)_{;\nu} = A^\mu(V_\mu)_{;\nu} + V_\mu(A^\mu)_{;\nu} = A^\mu \frac{\partial V_\mu}{\partial x^\nu} + V_\mu \frac{\partial A^\mu}{\partial x^\nu} \quad (\Gamma\text{'s cancel!})$$

This means that you can interchange the order of lowering or raising an index and covariant differentiation; hence

- The operation of contracting two tensor indices commutes with covariant differentiation. It does not matter which you do first. Try it in the second example above by considering the covariant derivative of  $V^\mu V_\nu$  first before, then after, contracting the indices.

### 4.3 The affine connection and basis vectors

The reader may be wondering how this all relates to our notions of, say, spherical or polar geometry and their associated sets of unit vectors and coordinates. The answer is: very simply. Our discussion will be straightforward and intuitive, rather than rigorous.

A vector  $\mathbf{V}$  may be expanded in a set of basis vectors,

$$\mathbf{V} = V^a \mathbf{e}_a \quad (147)$$

where we sum over the repeated  $a$ , but  $a$  here on a **bold-faced** vector refers to a particular vector in the basis set. The  $V^a$  are the usual vector contravariant components: old friends, just numbers. Note that the sum is *not* a scalar formed from a contraction! We've used roman letters here to help avoid that pitfall.



The covariant components are associated with what mathematicians are pleased to call a *dual* basis:

$$\mathbf{V} = V_b \mathbf{e}^b \quad (148)$$

Same  $\mathbf{V}$  mind you, just different ways of representing its components. If the  $\mathbf{e}$ 's seem a little abstract, don't worry, just take them at a formal level for the moment. You've seen something very like them before in elementary treatments of vectors.

The basis and the dual basis are related by a dot product rule,

$$\mathbf{e}_a \cdot \mathbf{e}^b = \delta_a^b. \quad (149)$$

This dot product rule relates the vectors of orthonormal bases. The basis vectors transform just as good old vectors should:

$$\mathbf{e}'_a = \frac{\partial x^b}{\partial x'^a} \mathbf{e}_b, \quad \mathbf{e}'^a = \frac{\partial x'^a}{\partial x^b} \mathbf{e}^b. \quad (150)$$

Note that the dot product rule gives

$$\mathbf{V} \cdot \mathbf{V} = V^a V_b \mathbf{e}_a \cdot \mathbf{e}^b = V^a V_b \delta_a^b = V^a V_a, \quad (151)$$

as we would expect. On the other hand, expanding the differential line element  $d\mathbf{s}$ ,

$$ds^2 = \mathbf{e}_a dx^a \cdot \mathbf{e}_b dx^b = \mathbf{e}_a \cdot \mathbf{e}_b dx^a dx^b \quad (152)$$

so that we recover the metric tensor

$$g_{ab} = \mathbf{e}_a \cdot \mathbf{e}_b \quad (153)$$

Exactly the same style calculation gives

$$g^{ab} = \mathbf{e}^a \cdot \mathbf{e}^b \quad (154)$$

These last two equations tell us first that  $g_{ab}$  is the coefficient of  $\mathbf{e}^a$  in an expansion of the vector  $\mathbf{e}_b$  in the usual basis:

$$\mathbf{e}_b = g_{ab} \mathbf{e}^a, \quad (155)$$

and tell us second that  $g^{ab}$  is the coefficient of  $\mathbf{e}_a$  in an expansion of the vector  $\mathbf{e}^b$  in the dual basis:

$$\mathbf{e}^b = g^{ab} \mathbf{e}_a \quad (156)$$

We've recovered the rules for raising and lowering indices, in this case for the entire basis vector.

Basis vectors change with coordinate position, as pretty much all vectors do in general. We *define* an thrice-indexed object  $\Gamma_{ac}^b$  by

$$\frac{\partial \mathbf{e}_a}{\partial x^c} = \Gamma_{ac}^b \mathbf{e}_b \quad (157)$$

so that

$$\Gamma_{ac}^b = \mathbf{e}^b \cdot \partial_c \mathbf{e}_a \equiv \partial_c (\mathbf{e}_a \cdot \mathbf{e}^b) - \mathbf{e}_a \cdot \partial_c \mathbf{e}^b = -\mathbf{e}_a \cdot \partial_c \mathbf{e}^b. \quad (158)$$

(Remember the shorthand notation  $\partial/\partial x^c = \partial_c$ .) The last equality implies the expansion

$$\frac{\partial \mathbf{e}^b}{\partial x^c} = -\Gamma_{ac}^b \mathbf{e}^a \quad (159)$$

Consider  $\partial_c g_{ab} = \partial_c (\mathbf{e}_a \cdot \mathbf{e}_b)$ . Using (157),

$$\partial_c g_{ab} = (\partial_c \mathbf{e}_a) \cdot \mathbf{e}_b + \mathbf{e}_a \cdot (\partial_c \mathbf{e}_b) = \Gamma_{ac}^d \mathbf{e}_d \cdot \mathbf{e}_b + \mathbf{e}_a \cdot \Gamma_{bc}^d \mathbf{e}_d, \quad (160)$$

or finally

$$\partial_c g_{ab} = \Gamma_{ac}^d g_{db} + \Gamma_{bc}^d g_{ad}, \quad (161)$$

exactly what we found in (86)! This leads, in turn, precisely to (90), the equation for the affine connection in terms of the  $g$  partial derivatives. We now have a more intuitive understanding of what the  $\Gamma$ 's really represent: they are expansion coefficients for the derivatives of basis vectors, which is how we are used to thinking of the extra acceleration terms in non Cartesian coordinates when we first encounter them in our first mechanics courses. In Cartesian coordinates, the  $\Gamma_{ac}^b$  just go away.

Finally, consider

$$\partial_a (V^b \mathbf{e}_b) = (\partial_a V^b) \mathbf{e}_b + V^b \partial_a \mathbf{e}_b = (\partial_a V^b) \mathbf{e}_b + V^b \Gamma_{ab}^c \mathbf{e}_c \quad (162)$$

Taking the dot product with  $\mathbf{e}^d$ :

$$\mathbf{e}^d \cdot \partial_a (V^b \mathbf{e}_b) = \partial_a V^d + V^b \Gamma_{ab}^d \equiv V^d_{;a}, \quad (163)$$

just the familiar covariant derivative of a contravariant vector. This one you should be able to do yourself:

$$\mathbf{e}_d \cdot \partial_a (V_b \mathbf{e}^b) = \partial_a V_d - V_b \Gamma_{ad}^b \equiv V_{d;a}, \quad (164)$$

the covariant derivative of a covariant vector. This gives us some understanding as to why the true tensors formed from the partial derivatives of a vector  $\mathbf{V}$  are not simply  $\partial_a V^d$  and  $\partial_a V_d$ , but rather  $\mathbf{e}^d \cdot \partial_a (V^b \mathbf{e}_b)$  and  $\mathbf{e}_d \cdot \partial_a (V_b \mathbf{e}^b)$  respectively. The  $\Gamma$ -terms then emerge as derivatives of the  $\mathbf{e}$  basis vectors. Our terse, purely coordinate notation avoids the use of the  $\mathbf{e}$  bases, but at a cost of missing a deeper and ultimately simplifying mathematical structure. We can see an old maxim of mathematicians in action: good mathematics starts with good definitions.

## 4.4 Volume element

The transformation of the metric tensor  $g_{\mu\nu}$  may be thought of as a matrix equation:

$$g'_{\mu\nu} = \frac{\partial x^\kappa}{\partial x'^\mu} g_{\kappa\lambda} \frac{\partial x^\lambda}{\partial x'^\nu} \quad (165)$$

Remembering that the determinant of the product of matrices is the product of the determinants, we find

$$g' = \left| \frac{\partial x}{\partial x'} \right|^2 g \quad (166)$$

where  $g$  is the determinant of  $g_{\mu\nu}$  (just the product of the diagonal terms for the diagonal metrics we will be using), and the notation  $|\partial x'/\partial x|$  indicates the Jacobian of the transformation  $x \rightarrow x'$ . The significance of this result is that there is another quantity that also transforms with a Jacobian factor: the volume element  $d^4x$ .

$$d^4x' = \left| \frac{\partial x'}{\partial x} \right| d^4x. \quad (167)$$

This means

$$\sqrt{-g'} d^4 x' = \sqrt{-g} \left| \frac{\partial x}{\partial x'} \right| \left| \frac{\partial x'}{\partial x} \right| d^4 x = \sqrt{-g} d^4 x. \quad (168)$$

In other words,  $\sqrt{-g} d^4 x$  is the invariant volume element of curved spacetime. The minus sign is used merely as an absolute value to keep the quantities positive. In flat Minkowski space time,  $d^4 x$  is invariant by itself.

Euclidian example: in going from Cartesian ( $g = 1$ ) to cylindrical polar ( $g = R^2$ ) to spherical coordinates ( $g = r^4 \sin^2 \theta$ ), we have  $dx dy dz = R dR dz d\phi = r^2 \sin \theta dr d\theta d\phi$ . You knew that. For a diagonal  $g_{\mu\nu}$ , our formula gives a volume element of

$$\sqrt{|g_{11}g_{22}g_{33}g_{00}|} dx^1 dx^2 dx^3 dx^0,$$

just the product of the *proper* differential intervals. That also makes sense.

## 4.5 Covariant div, grad, curl, and all that

The ordinary partial derivative of a scalar transforms generally as covariant vector, so in this case there is no distinction between a covariant and standard partial derivative. Another easy result is

$$V_{\mu;\nu} - V_{\nu;\mu} = \frac{\partial V_\mu}{\partial x^\nu} - \frac{\partial V_\nu}{\partial x^\mu}. \quad (169)$$

(The affine connection terms are symmetric in the two lower indices, so they cancel.) More interesting is

$$V_{;\mu}^\mu = \frac{\partial V^\mu}{\partial x^\mu} + \Gamma_{\mu\lambda}^\mu V^\lambda \quad (170)$$

where by definition

$$\Gamma_{\mu\lambda}^\mu = \frac{g^{\mu\rho}}{2} \left( \frac{\partial g_{\rho\mu}}{\partial x^\lambda} + \frac{\partial g_{\rho\lambda}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\rho} \right) \quad (171)$$

Now,  $g^{\mu\rho}$  is symmetric in its indices, whereas the last two  $g$  derivatives combined are anti-symmetric in the same indices, so that combination disappears entirely. We are left with

$$\Gamma_{\mu\lambda}^\mu = \frac{g^{\mu\rho}}{2} \frac{\partial g_{\rho\mu}}{\partial x^\lambda} \quad (172)$$

In this course, we will be dealing entirely with diagonal metric tensors, in which  $\mu = \rho$  for nonvanishing entries, and  $g^{\mu\rho}$  is the reciprocal of  $g_{\mu\rho}$ . In this simple case,

$$\Gamma_{\mu\lambda}^\mu = \frac{1}{2} \frac{\partial \ln |g|}{\partial x^\lambda} \quad (173)$$

where  $g$  is as usual the determinant of  $g_{\mu\nu}$ , here just the product of the diagonal elements. Though our result seems specific to diagonal  $g_{\mu\nu}$ , W72 pp. 106-7, shows that this result is true for *any*  $g_{\mu\nu}$ .<sup>2</sup>

---

<sup>2</sup>Sketchy proof for the mathematically inclined: For matrix  $M$ , trace  $\text{Tr}$ , differential  $\delta$ , to first order in  $\delta$  we have  $\delta \ln \det M = \ln \det(M + \delta M) - \ln \det M = \ln \det M^{-1}(M + \delta M) = \ln \det(1 + M^{-1}\delta M) = \ln(1 + \text{Tr } M^{-1}\delta M) = \text{Tr } M^{-1}\delta M$ . Can you supply the missing details?

The covariant divergence (170) becomes

$$V_{;\mu}^{\mu} = \frac{1}{\sqrt{|g|}} \frac{\partial(\sqrt{|g|}V^{\mu})}{\partial x^{\mu}} \quad (174)$$

a neat and tidy result. Note that

$$\int \sqrt{|g|} d^4x V_{;\mu}^{\mu} = 0 \quad (175)$$

if  $V^{\mu}$  vanishes sufficiently rapidly at infinity. (Why?)

We cannot leave the covariant derivative without discussing  $T_{;\mu}^{\mu\nu}$ , the covariant divergence of  $T^{\mu\nu}$ . (And similarly for the divergence of  $T_{\nu}^{\mu}$ .) Conserved stress tensors are, after all, general relativity's "coin of the realm." We have:

$$T_{;\mu}^{\mu\nu} = \frac{\partial T^{\mu\nu}}{\partial x^{\mu}} + \Gamma_{\mu\lambda}^{\mu} T^{\lambda\nu} + \Gamma_{\mu\lambda}^{\nu} T^{\mu\lambda}, \text{ or } T_{\nu;\mu}^{\mu} = \frac{\partial T_{\nu}^{\mu}}{\partial x^{\mu}} + \Gamma_{\mu\lambda}^{\mu} T_{\nu}^{\lambda} - \Gamma_{\mu\nu}^{\lambda} T_{\lambda}^{\mu} \quad (176)$$

and using (173), we may condense this to

$$T_{;\mu}^{\mu\nu} = \frac{1}{\sqrt{|g|}} \frac{\partial(\sqrt{|g|}T^{\mu\nu})}{\partial x^{\mu}} + \Gamma_{\mu\lambda}^{\nu} T^{\mu\lambda}, \text{ or } T_{\nu;\mu}^{\mu} = \frac{1}{\sqrt{|g|}} \frac{\partial(\sqrt{|g|}T_{\nu}^{\mu})}{\partial x^{\mu}} - \Gamma_{\mu\nu}^{\lambda} T_{\lambda}^{\mu}. \quad (177)$$

For an antisymmetric *contravariant* tensor, call it  $A^{\mu\nu}$ , the last term of the first equality drops out because  $\Gamma$  is symmetric in its lower indices:

$$A_{;\mu}^{\mu\nu} = \frac{1}{\sqrt{|g|}} \frac{\partial(\sqrt{|g|}A^{\mu\nu})}{\partial x^{\mu}} \text{ if } A^{\mu\nu} \text{ antisymmetric.} \quad (178)$$

## 4.6 Hydrostatic equilibrium

You have been patient and waded through a sea of indices, and it is time to be rewarded. We will do our first real physics problem in general relativity: hydrostatic equilibrium.

In Newtonian mechanics, you will recall that hydrostatic equilibrium represents a balance between a pressure gradient and the force of gravity. In general relativity this is completely encapsulated in the condition

$$T_{;\mu}^{\mu\nu} = 0$$

applied to the energy-momentum stress tensor (65), upgraded to covariant status:

$$T^{\mu\nu} = P g^{\mu\nu} + (\rho + P/c^2) U^{\mu} U^{\nu} \quad (179)$$

Our conservation equation is

$$0 = T_{;\mu}^{\mu\nu} = g^{\mu\nu} \frac{\partial P}{\partial x^{\mu}} + [(\rho + P/c^2) U^{\mu} U^{\nu}]_{;\mu} \quad (180)$$

where we have made use of the Leibniz rule for the covariant derivative of a product, and the fact that the  $g_{\mu\nu}$  covariant derivative vanishes. Using (177):

$$0 = g^{\mu\nu} \frac{\partial P}{\partial x^{\mu}} + \frac{1}{|g|^{1/2}} \frac{\partial}{\partial x^{\mu}} [|g|^{1/2} (\rho + P/c^2) U^{\mu} U^{\nu}] + \Gamma_{\mu\lambda}^{\nu} (\rho + P/c^2) U^{\mu} U^{\lambda} \quad (181)$$

In static equilibrium, all the  $U$  components vanish except  $U^0$ . To determine this, we use

$$g_{\mu\nu}U^\mu U^\nu = -c^2 \quad (182)$$

the upgraded version of special relativity's  $\eta_{\alpha\beta}U^\alpha U^\beta = -c^2$ . Thus,

$$(U^0)^2 = -\frac{c^2}{g_{00}}, \quad (183)$$

and with

$$\Gamma_{00}^\nu = -\frac{g^{\mu\nu}}{2} \frac{\partial g_{00}}{\partial x^\mu}, \quad (184)$$

our equation reduces to

$$0 = g^{\mu\nu} \left[ \frac{\partial P}{\partial x^\mu} + (\rho c^2 + P) \frac{\partial \ln |g_{00}|^{1/2}}{\partial x^\mu} \right] \quad (185)$$

Since  $g_{\mu\nu}$  has a perfectly good inverse, the term in square brackets must be zero:

$$\frac{\partial P}{\partial x^\mu} + (\rho c^2 + P) \frac{\partial \ln |g_{00}|^{1/2}}{\partial x^\mu} = 0 \quad (186)$$

This is the general relativistic equation of hydrostatic equilibrium. Compare this with the Newtonian counterpart:

$$\nabla P + \rho \nabla \Phi = 0 \quad (187)$$

The difference for a static problem is the replacement of  $\rho$  by  $\rho + P/c^2$  for the inertial mass density, and the use of  $\ln |g_{00}|^{1/2}$  for the potential (to which it reduces in the Newtonian limit).

If  $P = P(\rho)$ ,  $P' \equiv dP/d\rho$ , equation (186) may be formally integrated:

$$\int \frac{P'(\rho) d\rho}{P(\rho) + \rho c^2} + \ln |g_{00}|^{1/2} = \text{constant}. \quad (188)$$

*Exercise.* Solve the GR equation of hydrostatic equilibrium exactly for the case  $|g_{00}| = (1 - 2GM/rc^2)^{1/2}$  (e.g., near the surface of a neutron star) and  $P = K\rho^\gamma$  for  $\gamma \geq 1$ .

## 4.7 Covariant differentiation and parallel transport

Recall the geodesic equation,

$$\frac{d^2 x^\lambda}{d\tau^2} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0. \quad (189)$$

Writing the vector  $dx^\lambda/d\tau$  as  $V^\lambda$  this becomes

$$\frac{dV^\lambda}{d\tau} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} V^\nu = 0, \quad (190)$$

a covariant formulation of the statement that the vector  $V^\lambda$  is conserved along a geodesic path. But the covariance property of this statement has nothing to do with the specific

identity of  $V^\lambda$  with  $dx^\lambda/d\tau$ . We have seen that the left-side of this equation is a genuine vector for *any*  $V^\lambda$  as long as  $V^\lambda$  itself is a bona fide contravariant vector. The right side simply tells us that the fully covariant left side expression happens to vanish. Therefore, just as we “upgrade” from special to general relativity the partial derivative,

$$\frac{\partial V^\alpha}{\partial x^\beta} \rightarrow \frac{\partial V^\lambda}{\partial x^\mu} + \Gamma_{\mu\nu}^\lambda V^\nu \equiv V^\lambda_{;\mu} \quad (191)$$

we upgrade the derivative along a path  $x(\tau)$  in the same way by multiplying by  $dx^\mu/d\tau$  and summing over the index  $\mu$ :

$$\frac{dV^\alpha}{d\tau} \rightarrow \frac{dV^\lambda}{d\tau} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} V^\nu \equiv \frac{DV^\lambda}{D\tau} \quad (192)$$

$DV^\lambda/D\tau$  is a true vector; the transformation

$$\frac{DV'^\lambda}{D\tau} = \frac{\partial x'^\lambda}{\partial x^\mu} \frac{DV^\mu}{D\tau} \quad (193)$$

may be verified directly. (The inhomogeneous contributions from the  $\Gamma$  transformation and the derivatives of the derivatives of the coordinate transformation coefficients cancel in a manner exactly analogous to our original covariant partial derivative calculation.)

Exactly the same reasoning is used to define the covariant derivative for a covariant vector, which follows from (142):

$$\frac{dV_\lambda}{d\tau} - \Gamma_{\mu\lambda}^\nu \frac{dx^\mu}{d\tau} V_\nu \equiv \frac{DV_\lambda}{D\tau}. \quad (194)$$

The same is true for tensors, e.g.:

$$\frac{dT_\lambda^\sigma}{d\tau} + \Gamma_{\mu\nu}^\sigma \frac{dx^\nu}{d\tau} T_\lambda^\mu - \Gamma_{\lambda\nu}^\mu \frac{dx^\nu}{d\tau} T_\mu^\sigma \equiv \frac{DT_\lambda^\sigma}{D\tau}. \quad (195)$$

As you might expect, the Leibniz rule holds for the covariant derivative along a path, e.g.

$$\frac{D(V^\mu V_\mu)}{D\tau} = V^\mu \frac{DV_\mu}{D\tau} + V_\mu \frac{DV^\mu}{D\tau},$$

with affine connection terms always cancelling when they need to.

When a vector or tensor quantity is carried along a path, and the object does not change in a locally inertially reference frame ( $d/d\tau = 0$ ), this statement becomes, in arbitrary coordinates,  $D/D\tau = 0$ . This is the same physical result expressed in a covariant language. (Once again this works because the statements agree in the inertial coordinates, and then their zero difference is zero in any coordinate frame.) The condition  $D/D\tau = 0$  is known as *parallel transport*. A steady vector, for example, may always point along the  $y$  axis as we move it around in the  $xy$  plane, but its  $r$  and  $\theta$  components *will* have to change in order to keep this true! The prescription for how those components change is the content of the parallel transport equation.

Suppose that we do a round trip following the rule of parallel transport, and we come back to our exact starting point. Does a vector/tensor have to have the same value it began with? You might think that the answer must be yes, but it turns out to be more complicated than that. Indeed, it is a most interesting question...

The stage is now set to introduce the key tensor embodying the gravitational distortion of spacetime.

## 5 The curvature tensor

*The properties which distinguish space from other conceivable triply-extended magnitudes are only to be deduced from experience...At every point the three-directional measure of curvature can have an arbitrary value if only the effective curvature of every measurable region of space does not differ noticeably from zero.*

— G. F. B. Riemann

### 5.1 Commutation rule for covariant derivatives

The covariant derivative shares many properties with the ordinary partial derivative: it is a linear operator, it obeys the Leibniz rule, and it allows true tensor status to be bestowed upon partial derivatives (or derivatives with respect to a scalar parameter) in any coordinates. A natural question arises. Ordinarily, partial derivatives commute: the order in which they are taken does not matter, provided that appropriate smoothness conditions are present. Does the same commutation work for covariant derivatives? Does  $V^\mu_{;\sigma;\tau}$  equal  $V^\mu_{;\tau;\sigma}$ ?

Just do it.

$$V^\mu_{;\sigma} = \frac{\partial V^\mu}{\partial x^\sigma} + \Gamma^\mu_{\nu\sigma} V^\nu \equiv T^\mu_\sigma \quad (196)$$

Then

$$T^\mu_{\sigma;\tau} = \frac{\partial T^\mu_\sigma}{\partial x^\tau} + \Gamma^\mu_{\nu\tau} T^\nu_\sigma - \Gamma^\nu_{\sigma\tau} T^\mu_\nu, \quad (197)$$

or

$$T^\mu_{\sigma;\tau} = \frac{\partial^2 V^\mu}{\partial x^\tau \partial x^\sigma} + \frac{\partial}{\partial x^\tau} (\Gamma^\mu_{\lambda\sigma} V^\lambda) + \Gamma^\mu_{\nu\tau} \left( \frac{\partial V^\nu}{\partial x^\sigma} + \Gamma^\nu_{\lambda\sigma} V^\lambda \right) - \Gamma^\nu_{\sigma\tau} \left( \frac{\partial V^\mu}{\partial x^\nu} + \Gamma^\mu_{\lambda\nu} V^\lambda \right) \quad (198)$$

The first term and the last group (proportional to  $\Gamma^\nu_{\sigma\tau}$ ) are manifestly symmetric in  $\sigma$  and  $\tau$ , and so will vanish when the same calculation is done with the indices reversed and then subtracted off. A bit of inspection shows that the same is true for all the remaining terms proportional to the partial derivatives of  $V^\mu$ . The residual terms from taking the covariant derivative commutator are then

$$T^\mu_{\sigma;\tau} - T^\mu_{\tau;\sigma} = \left[ \frac{\partial \Gamma^\mu_{\lambda\sigma}}{\partial x^\tau} - \frac{\partial \Gamma^\mu_{\lambda\tau}}{\partial x^\sigma} + \Gamma^\mu_{\nu\tau} \Gamma^\nu_{\lambda\sigma} - \Gamma^\mu_{\nu\sigma} \Gamma^\nu_{\lambda\tau} \right] V^\lambda, \quad (199)$$

which we may write as

$$T^\mu_{\sigma;\tau} - T^\mu_{\tau;\sigma} = R^\mu_{\lambda\sigma\tau} V^\lambda \quad (200)$$

Now, both sides of this equation must be tensors, and  $V^\lambda$  is itself some arbitrary vector. (Good thing those partial derivatives of  $V^\lambda$  cancelled out.) This means that the quantity

$R^\mu_{\lambda\sigma\tau}$  must transform its coordinate indices as a tensor. That it actually does so may also be verified explicitly in a nasty calculation (if you really want to see it spelt out in detail, see W72 pp.132-3). We conclude that

$$R^\mu_{\lambda\sigma\tau} = \frac{\partial \Gamma^\mu_{\lambda\sigma}}{\partial x^\tau} - \frac{\partial \Gamma^\mu_{\lambda\tau}}{\partial x^\sigma} + \Gamma^\mu_{\nu\tau} \Gamma^\nu_{\lambda\sigma} - \Gamma^\mu_{\nu\sigma} \Gamma^\nu_{\lambda\tau} \quad (201)$$

is a true tensor. It is known as the *Riemann curvature tensor*. In fact, it may be shown (W72 p. 134) that this is the only tensor that is *linear* in the second derivatives of  $g_{\mu\nu}$ , and contains only its first and second derivatives.

Why do we refer to this peculiar mixed tensor as the “curvature tensor?” We may begin to answer this by noting that it vanishes in ordinary flat Minkowski spacetime—simply choose Cartesian coordinates to do the calculation. Then, because  $R^\mu_{\lambda\sigma\tau}$  is a tensor, if it is zero in one set of coordinates, it is zero in all. Commuting covariant derivatives makes sense in this case, since they amount to ordinary derivatives. Evidently, distortions from Minkowski space are essential for a nonvanishing curvature tensor, an intuition we will strengthen in the next section.

*Exercise.* What is the (much simpler) form of  $R^\mu_{\lambda\sigma\tau}$  in local inertial coordinates? It is often convenient to work in such coordinates to prove a result, and then generalise it to arbitrary coordinates using the fact that  $R^\mu_{\lambda\sigma\tau}$  is a tensor.

## 5.2 Parallel transport

We move on to the yet more striking example of parallel transport. Consider a vector  $V_\lambda$  whose covariant derivative along a curve  $x(\tau)$  vanishes. Then,

$$\frac{dV_\lambda}{d\tau} = \Gamma^\mu_{\lambda\nu} \frac{dx^\nu}{d\tau} V_\mu \quad (202)$$

Consider next a tiny round trip journey over a closed path on which  $V_\lambda$  is changing by the above prescription. If we remain in the neighbourhood of some point  $X^\rho$ , with  $x^\rho$  passing through  $X^\rho$  at some instant  $\tau_0$ ,  $x^\rho(\tau_0) = X^\rho$ , we Taylor expand as follows:

$$\Gamma^\mu_{\lambda\nu}(x) = \Gamma^\mu_{\lambda\nu}(X) + (x^\rho - X^\rho) \frac{\partial \Gamma^\mu_{\lambda\nu}}{\partial X^\rho} + \dots \quad (203)$$

$$V_\mu[x(\tau)] = V_\mu(X) + dV_\mu + \dots = V_\mu(X) + (x^\rho - X^\rho) \Gamma^\sigma_{\mu\rho}(X) V_\sigma(X) + \dots \quad (204)$$

(where  $x^\rho - X^\rho$  is  $dx^\rho$  from the right side of the parallel transport equation), whence

$$\Gamma^\mu_{\lambda\nu}(x) V_\mu(x) = \Gamma^\mu_{\lambda\nu} V_\mu + (x^\rho - X^\rho) V_\sigma \left( \frac{\partial \Gamma^\sigma_{\lambda\nu}}{\partial X^\rho} + \Gamma^\sigma_{\mu\rho} \Gamma^\mu_{\lambda\nu} \right) + \dots \quad (205)$$

where all quantities on the right (except  $x$ !) are evaluated at  $X$ . Integrating

$$dV_\lambda = \Gamma^\mu_{\lambda\nu}(x) V_\mu(x) dx^\nu \quad (206)$$

around a tiny closed path  $\oint$ , and using (206) and (205), we find that there is a change in the starting value  $\Delta V_\lambda$  arising from the term linear in  $x^\rho$  given by

$$\Delta V_\lambda = \left( \frac{\partial \Gamma^\sigma_{\lambda\nu}}{\partial X^\rho} + \Gamma^\sigma_{\mu\rho} \Gamma^\mu_{\lambda\nu} \right) V_\sigma \oint x^\rho dx^\nu \quad (207)$$



The integral  $\oint x^\rho dx^\nu$  certainly doesn't vanish. (Try integrating it around a unit square in the  $xy$  plane.) But it *is* antisymmetric in  $\rho$  and  $\nu$ . (Integrate by parts and note that the integrated term vanishes, being an exact differential.) That means the only part of the  $\Gamma\Gamma$  and  $\partial\Gamma/\partial X$  terms that survives the  $\rho\nu$  summation is the part that is antisymmetric in  $(\rho, \nu)$ . Since any object depending on two indices, say  $A(\rho, \nu)$ , can be written as a symmetric part plus an antisymmetric part,

$$\frac{1}{2}[A(\rho, \nu) + A(\nu, \rho)] + \frac{1}{2}[A(\rho, \nu) - A(\nu, \rho)],$$

we find

$$\Delta V_\lambda = \frac{1}{2} R^\sigma_{\lambda\nu\rho} V_\sigma \oint x^\rho dx^\nu \quad (208)$$

where

$$R^\sigma_{\lambda\nu\rho} = \left( \frac{\partial \Gamma^\sigma_{\lambda\nu}}{\partial X^\rho} - \frac{\partial \Gamma^\sigma_{\lambda\rho}}{\partial X^\nu} + \Gamma^\sigma_{\mu\rho} \Gamma^\mu_{\lambda\nu} - \Gamma^\sigma_{\mu\nu} \Gamma^\mu_{\lambda\rho} \right) \quad (209)$$

is precisely the curvature tensor. Parallel transport of a vector around a closed curve does not change the vector, unless the enclosed area has a nonvanishing curvature tensor. In fact, “the enclosed area” can be given a more intuitive interpretation if we think of integrating around a very tiny square in the  $\rho\nu$  plane. Then the closed loop integral is just the directed area  $dx^\rho dx^\nu$ :

$$\Delta V_\lambda = \frac{1}{2} R^\sigma_{\lambda\nu\rho} V_\sigma dx^\rho dx^\nu. \quad (210)$$

The conversion of a tiny closed loop integral to an enclosed surface area element reminds us of Stokes theorem, and it will not be surprising to see that there is an analogy here to the identity “divergence of curl equals zero”. We will make good use of this shortly.

**Exercise. A laboratory demonstration.** Take a pencil and move it round the surface of a flat desktop without rotating the pencil. Do the same on a cylinder, something like a mailing tube. (Avoid the endcap!) Moving the pencil around a closed path, *always parallel to itself*, will in either case not change its orientation. Now do the same on the surface of a spherical globe. Take a small pencil, pointed poleward, and move it from the equator along the  $0^\circ$  meridian through Greenwich till you hit the north pole. Now, once again parallel to itself, move the pencil down the  $90^\circ\text{E}$  meridian till you come to the equator. Finally, once again parallel to itself, slide the pencil along the equator to return to the starting point at the prime meridian.

Has the pencil orientation changed from its initial one? Explain.

Curvature<sup>3</sup>, or more precisely the departure of spacetime from Minkowski structure, reveals itself through the existence of the curvature tensor  $R^\sigma_{\lambda\nu\rho}$ . If spacetime is Minkowski-flat, every component of the curvature tensor vanishes. An important consequence is that parallel transport around a closed loop can result in a vector or tensor not returning to its original value, if the closed loop encompasses matter (or its energy equivalent). An experiment was proposed in the 1960's to measure the precession of a gyroscope orbiting the earth due to the effects of the spacetime curvature tensor. This eventually evolved into a satellite known as Gravity Probe B, a \$750,000,000 mission, launched in 2004. Alas, it was plagued by technical problems for many years, and its results were controversial because of unexpectedly high noise levels (due to solar activity). A final publication of science results in

---

<sup>3</sup>“Curvature” is one of these somewhat misleading mathematical labels that has stuck, like “imaginary” numbers. The name implies an external dimension into which the space is curved or embedded, an unnecessary complication. The space is simply *distorted*.

2011 claims to have verified the predictions of general relativity to high accuracy, including an even smaller effect known as “frame dragging” from the earth’s rotation, but my sense is that there is lingering uneasiness in the physics community regarding the handling of the noise. Do an internet search on Gravity Probe B and judge for yourself!

When GPB was first proposed in the early 1960’s, tests of general relativity were very few and far between. Any potentially observable result was novel and worth exploring. Since that time, experimental GR has evolved *tremendously*. We now live in a world of gravitational lensing, exquisitely sensitive Shapiro time delays, and stunning confirmations of gravitational radiation, first via the binary pulsar system PSR1913+16, and more recently the direct signal detection of a number of sources via advanced LIGO. All of these will be discussed in later chapters. At this point it borders on ludicrous to entertain serious doubt that the crudest leading order general relativity parallel transport prediction is correct. (In fact, it looks like we have seen this effect directly in close binary pulsar systems.) Elaborately engineered artificial gyroscopes, precessing by teeny-tiny amounts in earth orbit, don’t seem very exciting any more to 21st century physicists.

## 5.3 Algebraic identities of $R^\sigma_{\nu\lambda\rho}$

### 5.3.1 Remembering the curvature tensor formula.

It is helpful to have a mnemonic for generating the curvature tensor. The hard part is keeping track of the indices. Remember that the tensor itself is just a sum of derivatives of  $\Gamma$ , and quadratic products of  $\Gamma$ . That part is easy to remember, since the curvature tensor has “dimensions” of  $1/x^2$ , where  $x$  represents a coordinate. To remember the coordinate juggling of  $R^a_{bcd}$  start with:

$$\frac{\partial \Gamma^a_{bc}}{\partial x^d} + \Gamma^*_{bc} \Gamma^a_{d*}$$

where the first  $abcd$  ordering is simple to remember since it follows the same placement in  $R^a_{bcd}$ , and  $*$  is a dummy variable. For the second  $\Gamma\Gamma$  term, remember to just write out the lower  $bcd$  indices straight across, making the last unfilled space a dummy index  $*$ . The counterpart dummy index that is summed over must then be the upper slot on the other  $\Gamma$ , since there is no self-contracted  $\Gamma$  in the full curvature tensor. There is then only one place left for upper  $a$ . To finish off, just subtract the same thing with  $c$  and  $d$  reversed. Think of it as swapping your CD’s. We arrive at:

$$R^a_{bcd} = \frac{\partial \Gamma^a_{bc}}{\partial x^d} - \frac{\partial \Gamma^a_{bd}}{\partial x^c} + \Gamma^*_{bc} \Gamma^a_{d*} - \Gamma^*_{bd} \Gamma^a_{c*} \quad (211)$$

### 5.3.2 $R_{\lambda\mu\nu\kappa}$ : fully covariant form

The fully covariant form of the stress tensor can be written so that it involves only second-order derivatives of  $g_{\mu\nu}$  and products of  $\Gamma$ s, with no  $\Gamma$  partial derivatives. The second-order  $g$ -derivatives, which are linear terms, will be our point of contact with Newtonian theory from the full field equations. But hang on, we have a bit of heavy weather ahead.

We define

$$R_{\lambda\mu\nu\kappa} = g_{\lambda\sigma} R^\sigma_{\mu\nu\kappa} \quad (212)$$

or

$$R_{\lambda\mu\nu\kappa} = g_{\lambda\sigma} \left[ \frac{\partial \Gamma^\sigma_{\mu\nu}}{\partial x^\kappa} - \frac{\partial \Gamma^\sigma_{\mu\kappa}}{\partial x^\nu} + \Gamma^\eta_{\mu\nu} \Gamma^\sigma_{\kappa\eta} - \Gamma^\eta_{\mu\kappa} \Gamma^\sigma_{\nu\eta} \right] \quad (213)$$

Remembering the definition of the affine connection (90), the right side of (213) is

$$\begin{aligned} & \frac{g_{\lambda\sigma}}{2} \frac{\partial}{\partial x^\kappa} \left[ g^{\sigma\rho} \left( \frac{\partial g_{\rho\mu}}{\partial x^\nu} + \frac{\partial g_{\rho\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\rho} \right) \right] - \frac{g_{\lambda\sigma}}{2} \frac{\partial}{\partial x^\nu} \left[ g^{\sigma\rho} \left( \frac{\partial g_{\rho\mu}}{\partial x^\kappa} + \frac{\partial g_{\rho\kappa}}{\partial x^\mu} - \frac{\partial g_{\mu\kappa}}{\partial x^\rho} \right) \right] \\ & + g_{\lambda\sigma} (\Gamma_{\mu\nu}^\eta \Gamma_{\kappa\eta}^\sigma - \Gamma_{\mu\kappa}^\eta \Gamma_{\nu\eta}^\sigma) \end{aligned} \quad (214)$$

The  $x^\kappa$  and  $x^\nu$  partial derivatives will operate on the  $g^{\sigma\rho}$  term and the  $g$ -derivative terms. Let us begin with the second group, the  $\partial g/\partial x$  derivatives, as it is simpler. With  $g_{\lambda\sigma} g^{\sigma\rho} = \delta_\lambda^\rho$ , the terms that are linear in the second order  $g$ -derivatives are

$$\frac{1}{2} \left( \frac{\partial^2 g_{\lambda\nu}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\kappa \partial x^\lambda} - \frac{\partial^2 g_{\lambda\kappa}}{\partial x^\nu \partial x^\mu} + \frac{\partial^2 g_{\mu\kappa}}{\partial x^\nu \partial x^\lambda} \right) \quad (215)$$

If you can sense the beginnings of the classical wave equation lurking in these linear second order derivatives, which are the leading terms when  $g_{\mu\nu}$  departs only a little from  $\eta_{\mu\nu}$ , then you are very much on the right track.

We are not done of course. We have the terms proportional to the  $x^\kappa$  and  $x^\nu$  derivatives of  $g^{\sigma\rho}$ , which certainly do not vanish in general. But the covariant derivative of the metric tensor  $g_{\lambda\sigma}$  *does* vanish, so invoke this sleight-of-hand integration by parts:

$$g_{\lambda\sigma} \frac{\partial g^{\sigma\rho}}{\partial x^\kappa} = -g^{\sigma\rho} \frac{\partial g_{\lambda\sigma}}{\partial x^\kappa} = -g^{\sigma\rho} (\Gamma_{\kappa\lambda}^\eta g_{\eta\sigma} + \Gamma_{\kappa\sigma}^\eta g_{\eta\lambda}) \quad (216)$$

where in the final equality, equation (144) has been used. By bringing  $g^{\sigma\rho}$  out from the partial derivative, it recombines with the first order  $g$ -derivatives to form affine connections once again. All the remaining terms of  $R_{\lambda\mu\nu\kappa}$  from (214) are now of the form  $g\Gamma\Gamma$ :

$$- (\Gamma_{\kappa\lambda}^\eta g_{\eta\sigma} + \cancel{\Gamma_{\kappa\sigma}^\eta g_{\eta\lambda}}) \Gamma_{\mu\nu}^\sigma + (\Gamma_{\nu\lambda}^\eta g_{\eta\sigma} + \cancel{\Gamma_{\nu\sigma}^\eta g_{\eta\lambda}}) \Gamma_{\mu\kappa}^\sigma + g_{\lambda\sigma} (\cancel{\Gamma_{\mu\nu}^\eta \Gamma_{\kappa\eta}^\sigma} - \cancel{\Gamma_{\mu\kappa}^\eta \Gamma_{\nu\eta}^\sigma}), \quad (217)$$

It is not obvious at first, but with a little colour coding and index agility to help, you should be able to see four of these six  $g\Gamma\Gamma$  terms cancel out—the second group with the fifth, the fourth group with the sixth—leaving only the first and third terms:

$$g_{\eta\sigma} (\Gamma_{\nu\lambda}^\eta \Gamma_{\mu\kappa}^\sigma - \Gamma_{\kappa\lambda}^\eta \Gamma_{\mu\nu}^\sigma) \quad (218)$$

Adding together the terms in (215) and (218), we arrive at

$$R_{\lambda\mu\nu\kappa} = \frac{1}{2} \left( \frac{\partial^2 g_{\lambda\nu}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\kappa \partial x^\lambda} - \frac{\partial^2 g_{\lambda\kappa}}{\partial x^\nu \partial x^\mu} + \frac{\partial^2 g_{\mu\kappa}}{\partial x^\nu \partial x^\lambda} \right) + g_{\eta\sigma} (\Gamma_{\nu\lambda}^\eta \Gamma_{\mu\kappa}^\sigma - \Gamma_{\kappa\lambda}^\eta \Gamma_{\mu\nu}^\sigma) \quad (219)$$

**Exercise.** What is  $R_{\lambda\mu\nu\kappa}$  in local inertial coordinates?

Note the following important symmetry properties for the indices of  $R_{\lambda\mu\nu\kappa}$ . Because each of these identities may be expressed as a vanishing tensor equation (left side minus right side equals zero), they may be established generally by choosing any particular coordinate frame we like. We choose a simple locally flat frame in which the  $\Gamma$  vanish. These results may then be verified just from the terms linear in the  $g$  derivatives in (219):

$$R_{\lambda\mu\nu\kappa} = R_{\nu\kappa\lambda\mu} \quad (\text{symmetry}) \quad (220)$$

$$R_{\lambda\mu\nu\kappa} = -R_{\mu\lambda\nu\kappa} = -R_{\lambda\mu\kappa\nu} = R_{\mu\lambda\kappa\nu} \quad (\text{antisymmetry}) \quad (221)$$

$$R_{\lambda\mu\nu\kappa} + R_{\lambda\kappa\mu\nu} + R_{\lambda\nu\kappa\mu} = 0 \quad (\text{cyclic}) \quad (222)$$

## 5.4 The Ricci Tensor

The Ricci tensor is the curvature tensor contracted on its (raised) first and third indices,  $R^a_{\text{ }bad}$ . In terms of the covariant curvature tensor:

$$R_{\mu\kappa} = g^{\lambda\nu} R_{\lambda\mu\nu\kappa} = g^{\lambda\nu} R_{\nu\kappa\lambda\mu} \text{ (by symmetry)} = g^{\nu\lambda} R_{\nu\kappa\lambda\mu} = R_{\kappa\mu} \quad (223)$$

so that the Ricci tensor is symmetric.

The Ricci tensor is an extremely important tensor in general relativity. Indeed, we shall very soon see that  $R_{\mu\nu} = 0$  is Einstein's Laplace equation. There is enough information here to calculate the deflection of light by a gravitating body or the advance of a planet's orbital perihelion! What is tricky is to guess the general relativistic version of the Poisson equation, and no, it is *not*  $R_{\mu\nu}$  proportional to the stress energy tensor  $T_{\mu\nu}$ . (It wouldn't be very tricky then, would it?) Notice that while  $R^{\lambda}_{\mu\nu\kappa} = 0$  implies that the Ricci tensor vanishes, the converse does not follow:  $R_{\mu\nu} = 0$  does not necessarily mean that the full curvature tensor (covariant or otherwise) vanishes.

**Exercise. Fun with the Ricci tensor.** Prove first that

$$R_{\mu\kappa} = \frac{\partial \Gamma^{\lambda}_{\mu\lambda}}{\partial x^{\kappa}} - \frac{\partial \Gamma^{\lambda}_{\mu\kappa}}{\partial x^{\lambda}} + \Gamma^{\eta}_{\mu\lambda} \Gamma^{\lambda}_{\kappa\eta} - \Gamma^{\eta}_{\mu\kappa} \Gamma^{\lambda}_{\lambda\eta}.$$

The expression given with the “Really Useful Numbers” on page 6 of these notes has a somewhat different form to the above. Can you show that the two expressions are actually one and the same?

Next show that

$$R_{\mu\kappa} = -g^{\lambda\nu} R_{\mu\lambda\nu\kappa} = -g^{\lambda\nu} R_{\lambda\mu\kappa\nu} = g^{\lambda\nu} R_{\mu\lambda\kappa\nu},$$

and that  $g^{\lambda\mu} R_{\lambda\mu\nu\kappa} = g^{\nu\kappa} R_{\lambda\mu\nu\kappa} = 0$ . Why does this mean that  $R_{\mu\kappa}$  is the only second rank covariant tensor that can be formed from contracting  $R_{\lambda\mu\nu\kappa}$ ?

We are not quite through contracting. We may form the curvature scalar

$$R \equiv R^{\mu}_{\mu}, \quad (224)$$

another very important quantity in general relativity.

**Exercise. The curvature scalar is unique.** Prove that

$$R = g^{\nu\lambda} g^{\mu\kappa} R_{\lambda\mu\nu\kappa} = -g^{\nu\lambda} g^{\mu\kappa} R_{\mu\lambda\nu\kappa}$$

and that

$$g^{\lambda\mu} g^{\nu\kappa} R_{\lambda\mu\nu\kappa} = 0.$$

Justify the title of this exercise.

The Ricci tensor can be a real nuisance to evaluate explicitly in terms of the partial derivatives of  $g_{\mu\nu}$ , with sums within sums to be carried out. We are, alas, often faced with this task. Fortunately, for the case of a diagonal metric tensor, a more explicit formula is available, as the next exercise shows. It is still messy!

**Exercise. The Ricci tensor for a diagonal  $g_{\mu\nu}$ . Off syllabus, only for the zealous!** Using the equations on page 29 for the affine connection  $\Gamma^a_{bc}$  for an assumed diagonal metric, derive the following two forms for the Ricci tensor. First, for components of the form  $R_{aa}$  (no sum on repeated roman indices throughout this exercise), show that

$$R_{aa} = \frac{1}{2} \partial_a \partial_a \ln g_{|a} - \frac{1}{4} (\partial_a \ln g_{aa}) \partial_a \ln g_{|a} + \frac{1}{2} \sum_{\lambda \neq a} \left[ \partial_{\lambda} \left( \frac{\partial_{\lambda} g_{aa}}{g_{\lambda\lambda}} \right) - \frac{(\partial_{\lambda} g_{aa})^2}{g_{aa} g_{\lambda\lambda}} + \frac{(\partial_a \ln g_{\lambda\lambda})^2}{2} + \frac{(\partial_{\lambda} g_{aa}) \partial_{\lambda} \ln g}{2 g_{\lambda\lambda}} \right].$$

The notation  $g_{|a}$  denotes the absolute value of the determinant of  $g_{\mu\nu}$  *without* the factor  $g_{aa}$ . For components of the form  $R_{ab}$ ,  $a \neq b$ , show that

$$R_{ab} = \frac{1}{2} \partial_a \partial_b \ln g_{|ab} - \frac{1}{4} [(\partial_a \ln g_{|a}) \partial_b \ln g_{aa} + (\partial_b \ln g_{|b}) \partial_a \ln g_{bb} - 2(\partial_b \ln g_{aa}) \partial_a \ln g_{bb} - \sum_{\lambda \neq a,b} (\partial_a \ln g_{\lambda\lambda}) \partial_b \ln g_{\lambda\lambda}].$$

The notation  $g_{|ab}$  denotes the absolute value of the determinant of  $g_{\mu\nu}$  without the factors  $g_{aa}$  and  $g_{bb}$ . These expressions look complicated, but in practice are not difficult to use, as several of the terms will often vanish identically.

(Note: I am unaware of these two formulae appearing elsewhere in the literature, but I suspect they are to be found somewhere. Student feedback is welcome.)

## 5.5 Curvature and Newtonian gravity

That curvature should be at the heart of our understanding of gravity is perhaps not too surprising if we think carefully about the Poisson equation:

$$\nabla^2 \Phi = 4\pi G \rho \quad (225)$$

where  $\Phi$  is the usual gravitational potential, and  $\rho$  the mass density. If you have an old-fashioned enough book on fluid dynamics available (I like *An Introduction to Fluid Dynamics* by G. Batchelor) you will come across a formula that relates the  $\nabla^2$  operator directly to the curvature of an ordinary 2D surface:

$$\frac{\partial^2 \zeta}{\partial x^2} + \frac{\partial^2 \zeta}{\partial y^2} = \frac{1}{R_1} + \frac{1}{R_2} \quad (226)$$

where  $\zeta$  is the (assumed small) displacement of the surface from planar, and the  $R$ 's are the two “principal radii of curvature.” (Pick a point and think of fitting as smoothly as possible, into the deformed surface, arcs of two circles at right angles to one another. The  $R$ 's are the radii of these circles.) So the Laplacian operator is on its own quite literally a measure of spatial curvature. Moreover, if we have a surface membrane with a pressure discontinuity  $\Delta P(x, y)$  across it, then the displacement satisfies the dynamical force balance (Young-Laplace equation):

$$\gamma \left( \frac{\partial^2 \zeta}{\partial x^2} + \frac{\partial^2 \zeta}{\partial y^2} \right) = \Delta P, \quad (227)$$

where  $\gamma$  is the assumed constant surface tension of the membrane (dimensions of energy per area). This is identical to the 2D Poisson equation. So, at least in a two-dimensional world, we can think of the Newtonian potential  $\Phi$  as the displacement of a surface membrane and  $\rho$  as the applied pressure difference. That makes  $1/(4\pi G)$ , in this analogy, the surface tension!

The origins of Riemannian Geometry began with Gauss trying to describe the mathematics of two-dimensional surfaces with arbitrary curvature. Riemann then worked out the notion of curvature in its full generality. It is of some interest therefore, to see geometrical curvature already built into 2D Newtonian gravity at a fundamental level.

## 5.6 The Bianchi Identities

The fully covariant curvature tensor obeys a very important set of differential identities, analogous to  $\mathbf{div}(\mathbf{curl})=0$ . These are the Bianchi identities. We shall prove the Bianchi

identities in our favourite freely-falling inertial coordinates with  $\Gamma = 0$ , and since we will be showing that a tensor is zero in these coordinates, it must be zero in all coordinate systems. In  $\Gamma = 0$  coordinates,

$$R_{\lambda\mu\nu\kappa;\eta} = \frac{1}{2} \frac{\partial}{\partial x^\eta} \left( \frac{\partial^2 g_{\lambda\nu}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\kappa \partial x^\lambda} - \frac{\partial^2 g_{\lambda\kappa}}{\partial x^\mu \partial x^\nu} + \frac{\partial^2 g_{\mu\kappa}}{\partial x^\nu \partial x^\lambda} \right) \quad (228)$$

The Bianchi identities follow from cycling  $\kappa$  goes to  $\nu$ ,  $\nu$  goes to  $\eta$ ,  $\eta$  goes to  $\kappa$ . Leave  $\lambda$  and  $\mu$  alone. Repeat. Add the original  $R_{\lambda\mu\nu\kappa;\eta}$  and the two cycled expressions together. You will find that this gives

$$R_{\lambda\mu\nu\kappa;\eta} + R_{\lambda\mu\eta\nu;\kappa} + R_{\lambda\mu\kappa\eta;\nu} = 0 \quad (229)$$

An easy way to check the bookkeeping on this is just to focus upon the  $g$ 's: once you've picked a particular value of  $\partial^2 g_{ab}$  in the numerator, the other  $\partial x^c$  indices downstairs are unambiguous, since as coordinate derivatives their order is immaterial. The first term in (229) is then as shown:  $(g_{\lambda\nu}, -g_{\mu\nu}, -g_{\lambda\kappa}, g_{\mu\kappa})$ . Cycle to get the second group for the second Bianchi term,  $(g_{\lambda\eta}, -g_{\mu\eta}, -g_{\lambda\nu}, g_{\mu\nu})$ . The final term then is  $(g_{\lambda\kappa}, -g_{\mu\kappa}, -g_{\lambda\eta}, g_{\mu\eta})$ . Look: every  $g_{ab}$  has its opposite when you add these all up, so the sum is clearly zero.

We would like to get equation (229) into the form of a *single* vanishing covariant tensor divergence, for reasons that will soon become very clear. Toward this goal, contract  $\lambda$  with  $\nu$ , remembering the symmetries in (221). (E.g., in the second term on the left side of [229], swap  $\nu$  and  $\eta$  before contracting, changing the sign.) We find

$$R_{\mu\kappa;\eta} - R_{\mu\eta;\kappa} + R^\nu_{\mu\kappa\eta;\nu} = 0. \quad (230)$$

Next, contract  $\mu$  with  $\kappa$ :

$$R_{;\eta} - R^\mu_{\eta;\mu} - R^\nu_{\eta;\nu} = 0 \quad (231)$$

(Did you understand the manipulations to get that final term on the left? First set things up with:

$$R^\nu_{\mu\kappa\eta;\nu} = g^{\nu\sigma} R_{\sigma\mu\kappa\eta;\nu} = -g^{\nu\sigma} R_{\mu\sigma\kappa\eta;\nu}$$

Now it is easy to raise  $\mu$  and contract with  $\kappa$ :

$$-g^{\nu\sigma} R^\mu_{\sigma\mu\eta;\nu} = -g^{\nu\sigma} R_{\sigma\eta;\nu} = -R^\nu_{\eta;\nu})$$

Cleaning things up, our contracted identity (231) becomes:

$$(\delta^\mu_\eta R - 2R^\mu_\eta)_{;\mu} = 0. \quad (232)$$

Raising  $\eta$  (we are allowed, of course, to bring  $g^{\nu\eta}$  *inside* the covariant derivative to do this—why?), and dividing by  $-2$  puts this identity into its classic “zero-divergence” form:

$$\left( R^{\mu\nu} - g^{\mu\nu} \frac{R}{2} \right)_{;\mu} = 0 \quad (233)$$

The generic tensor combination  $A^{\mu\nu} - g^{\mu\nu} A/2$  will appear repeatedly in our study of gravitational radiation.

Einstein did not know equation (233) when he was struggling mightily with his theory. But to be fair, neither did most mathematicians! The identities were actually first discovered by the German mathematician A. Voss in 1880, then independently in 1889 by Ricci. These results were then quickly forgotten, even, it seems, by Ricci himself. Bianchi then rediscovered them on his own in 1902, but they were still not widely known in the mathematics

community in 1915. This was a pity, because the Bianchi identities streamline the derivation of the field equations. Indeed, they are referred to as the “royal road to the Gravitational Field Equations ” by Einstein’s biographer A. Pais. It seems to have been the great mathematician H. Weyl who in 1917 first recognised the importance of the Bianchi identities for relativity, yet the particular derivation we have followed was not formulated until 1922, by Harward. A messy legacy.

The reason for the identities’ importance is precisely analogous to Maxwell’s penetrating understanding of the mathematical restrictions that the curl operator imposes on the type of field it generates: why does the displacement current need to be added to the equation  $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$ ? Because taking the divergence of this equation in this form gives zero identically on the left—the divergence of the curl is zero—so the right hand source term must also have a vanishing divergence. In other words, it must become a statement of some sort of *physical* conservation law. Maxwell needed and invoked a physical “displacement current,”  $(1/c^2)\partial \mathbf{E}/\partial t$ , adding it to the right side of the equation, because  $\nabla \cdot \mathbf{J}$  is manifestly *not* always zero. The ensuing physical conservation law corresponded to the conservation of electric charge, now built directly into the fundamental formulation of Maxwell’s Equations. Here, we shall use the Bianchi identities as the analogue (and it really is a precise mathematical analogue) of “the divergence of the curl is zero,” a *geometrical* constraint that ensures that the Gravitational Field Equations have conservation of the stress energy tensor automatically built into *their* fundamental formulation, just as Maxwell’s Field Equations have charge conservation built into their underlying structure. What is good for Maxwell is good for Einstein.

## 6 The Einstein Field Equations

*In the spring of 1913, Planck and Nernst had come to Zürich for the purpose of sounding out Einstein about his possible interest in moving to Berlin...Planck [asked him] what he was working on, and Einstein described general relativity as it was then. Planck said ‘As an older friend, I must advise you against it for in the first place you will not succeed; and even if you succeed, no one will believe you.’*

— A. Pais, writing in ‘Subtle is the Lord’

### 6.1 Formulation

We will now apply the principle of general covariance to the gravitational field itself. What is the relativistic analogue of  $\nabla^2\Phi = 4\pi G\rho$ ? We have now built up a sufficiently strong mathematical arsenal from Riemannian geometry to be able to give a satisfactory answer to this question.

We know that we must work with vectors and tensors to maintain general covariance, and that the Newton-Poisson source,  $\rho$ , is a only one component of a more general stress-energy tensor  $T_{\mu\nu}$  (in its covariant form) in relativity. We expect, therefore, that the gravitational field equations will take the form

$$G_{\mu\nu} = CT_{\mu\nu} \quad (234)$$

where  $C$  is a constant, and  $G_{\mu\nu}$  is some tensor that is comprised of  $g_{\mu\nu}$  and its second derivatives, or quadratic products of the first derivatives of  $g_{\mu\nu}$ . We guess this since i) we know that in the Newtonian limit the largest component of  $g_{\mu\nu}$  is the  $g_{00} \simeq -1 - 2\Phi/c^2$  component; ii) we need to recover the Poisson equation; and iii) we assume that we are seeking a theory of gravity that does not change its character with length scale: it has no characteristic length associated with it where the field properties change fundamentally. We need exclusively “ $1/r^2$ ” scaling on the left side.

The last condition may strike you as a bit too restrictive. *Hey, who ordered that?* Well, umm...OK, we now know this is actually wrong. It is wrong when applied to the Universe as a whole. But it is the simplest assumption that we can make that will satisfy all the basic requirements of a good theory. Let’s come back to the general relativity updates once we have version GR1.0 installed.

Next, we know that the stress energy tensor is conserved in the sense of  $T_{;\nu}^{\mu\nu} = 0$ . We also know from our work with the Bianchi identities of the previous section that this will automatically be satisfied if we take  $G_{\mu\nu}$  to be proportional to the particular linear combination

$$G_{\mu\nu} \propto R_{\mu\nu} - \frac{g_{\mu\nu}R}{2}$$

(Notice that there is no difficulty shifting indices up or down as considerations demand: our



index shifters  $g_{\mu\nu}$  and  $g^{\mu\nu}$  all have vanishing covariant derivatives and can be moved inside and outside of semi-colons.) We have therefore determined the field equations of gravity up to an overall normalisation:

$$R_{\mu\nu} - \frac{g_{\mu\nu}R}{2} = CT_{\mu\nu} \quad (235)$$

The final step is to recover the Newtonian limit. In this limit,  $T_{\mu\nu}$  is dominated by  $T_{00}$ , and  $g_{\mu\nu}$  can be replaced by  $\eta_{\alpha\beta}$  when shifting indices. The leading order derivative of  $g_{\mu\nu}$  that enters into the field equations comes from

$$g_{00} \simeq -1 - \frac{2\Phi}{c^2}$$

where  $\Phi$  is the usual Newtonian potential. In what follows, we use  $i, j, k$  to indicate spatial indices, and 0 will always be reserved for time.

The trace of equation (235) reads (raise  $\mu$ , contract with  $\nu$ ):

$$R - \frac{4 \times R}{2} = -R = CT. \quad (236)$$

Substituting this for  $R$  back in the original equation leads to

$$R_{\mu\nu} = C \left( T_{\mu\nu} - \frac{g_{\mu\nu}T}{2} \right) \equiv CS_{\mu\nu} \quad (237)$$

which defines the so-called source function, a convenient grouping we shall use later:

$$S_{\mu\nu} = T_{\mu\nu} - g_{\mu\nu}T/2. \quad (238)$$

The 00 component of (237) is

$$R_{00} = C \left( T_{00} - \frac{g_{00}T}{2} \right) \quad (239)$$

In the Newtonian limit, the trace  $T \equiv T^\mu_\mu$  is dominated by the 0 term,  $T^0_0$ , and raising and lowering of the indices is done by the  $\eta_{\mu\nu}$  weak field limit of  $g_{\mu\nu}$ .

$$R_{00} = C \left( T_{00} - \frac{\eta_{00}T^0_0}{2} \right) = C \left( T_{00} - \frac{T_{00}}{2} \right) = C \frac{T_{00}}{2} = C \frac{\rho c^2}{2}, \quad (240)$$

where  $\rho$  is the Newtonian mass density. Calculating  $R_{00}$  explicitly,

$$R_{00} = R^\nu_{0\nu 0} = \eta^{\lambda\nu} R_{\lambda 0 \nu 0} \quad (241)$$

We need only the linear part of  $R_{\lambda\mu\nu\kappa}$  in the weak field limit:

$$R_{\lambda\mu\nu\kappa} = \frac{1}{2} \left( \frac{\partial^2 g_{\lambda\nu}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\kappa \partial x^\lambda} - \frac{\partial^2 g_{\lambda\kappa}}{\partial x^\nu \partial x^\mu} + \frac{\partial^2 g_{\mu\kappa}}{\partial x^\nu \partial x^\lambda} \right), \quad (242)$$

and in the static limit with  $\mu = \kappa = 0$ , only the final term on the right side of this equation survives:

$$R_{\lambda 0 \nu 0} = \frac{1}{2} \frac{\partial^2 g_{00}}{\partial x^\nu \partial x^\lambda}. \quad (243)$$

Finally,

$$R_{00} = \eta^{\lambda\nu} R_{\lambda 0 \nu 0} = \frac{1}{2} \eta^{\lambda\nu} \frac{\partial^2 g_{00}}{\partial x^\lambda \partial x^\nu} = \frac{1}{2} \nabla^2 g_{00} = -\frac{1}{c^2} \nabla^2 \Phi = \frac{C \rho c^2}{2} \quad (244)$$

This happily agrees with the Poisson equation if  $C = -8\pi G/c^4$ . Hello Isaac Newton. As Einstein himself put it: “No fairer destiny could be allotted to any physical theory, than that it should of itself point out the way to the introduction of a more comprehensive theory, in which it lives on as a limiting case.” We therefore arrive at the *Einstein Field Equations*:

$$\boxed{G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\frac{8\pi G}{c^4} T_{\mu\nu}} \quad (245)$$

The Field Equations first appeared in Einstein’s notes on 25 November 1915, just over a hundred years ago, after an inadvertent competition with the mathematician David Hilbert, triggered by an Einstein colloquium at Göttingen. (Talk about being scooped! Hilbert actually derived the Field Equations first, by a variational method, but rightly insisted on giving Einstein full credit for the physical theory. Incidentally, in common with Einstein, Hilbert didn’t know the Bianchi identities.)

It is useful to also exhibit these equations explicitly in source function form. Contracting  $\mu$  and  $\nu$ ,

$$R = \frac{8\pi G}{c^4} T, \quad (246)$$

and the field equations become

$$R_{\mu\nu} = -\frac{8\pi G}{c^4} \left( T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right) \equiv -\frac{8\pi G}{c^4} S_{\mu\nu} \quad (247)$$

where as before,

$$S_{\mu\nu} = T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T, \quad (248)$$

a “Bianchified form” of the stress tensor. In vacuo, the Field Equations reduce to the analogue of the Laplace Equation:

$$R_{\mu\nu} = 0. \quad (249)$$

One final point. If we allow the possibility that gravity could change its form on different scales, it is always possible to add a term of the form  $\pm \Lambda g_{\mu\nu}$  to  $G_{\mu\nu}$ , where  $\Lambda$  is a constant, without violating the conservation of  $T_{\mu\nu}$  condition. This is because the covariant derivatives of  $g_{\mu\nu}$  vanish identically, so that  $T_{\mu\nu}$  is still conserved. Einstein, pursuing the consequences of a cosmological theory, realised that his field equations did not produce a static universe. This is bad, he thought, and Nature is not bad. Everyone knows the Universe is static. This is good. So he sought a source of static stabilisation, adding an offsetting positive  $\Lambda$  term to the right side of the field equations:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\frac{8\pi G}{c^4} T_{\mu\nu} + \Lambda g_{\mu\nu}, \quad (250)$$

and dubbed  $\Lambda$  the *cosmological constant*. Had he not done so, he could have made a spectacular prediction: the Universe is dynamic, a player in its own game, and must be either expanding or contracting.<sup>4</sup> With the historical discovery of an expanding universe, Einstein retracted the  $\Lambda$  term, calling it “the biggest mistake of my life.”

---

<sup>4</sup>Even within the context of straight Euclidian geometry and Newtonian dynamics, uniform expansion of an infinite space avoids the self-consistency problems associated with a static model. I’ve never understood why this simple point is not emphasised more.

Surprise. We now know that this term is, in fact, present on the largest cosmological scales, and on these scales it is no small effect. It mimics (and may well be) an energy density of the vacuum itself. It is measured to be 70% of the effective energy density in the Universe, and it certainly doesn't make the Universe static! It is to be emphasised that  $\Lambda$  must be taken into account only on the largest scales, scales over which the locally much higher baryon and dark matter inhomogeneities are lowered by effective smoothing.  $\Lambda$  is otherwise completely negligible. The so-called biggest mistake of Einstein's life was, therefore, quadratic in amplitude: one factor of error for introducing  $\Lambda$  for the wrong reason, the second factor for retracting  $\Lambda$  for the wrong reason!

Except for cosmological problems, we will always assume  $\Lambda = 0$ .

## 6.2 Coordinate ambiguities

There is no unique solution to the Field Equation because of the fact that they have been constructed to admit a new solution by a transformation of coordinates. To make this point as clear as possible, imagine that we have worked hard, solved for the metric  $g_{\mu\nu}$ , and it turns out to be plain old Minkowski space.<sup>5</sup> We have the usual time  $t$  coordinate for index 0, and let us say  $\alpha, \beta, \gamma$  for the three spatial dimensions. Even if we restrict ourselves to diagonal  $g_{\mu\nu}$ , we might have found that the four diagonal entries are  $(-1, 1, 1, 1)$  or  $(-1, 1, \alpha^2, 1)$  or  $(-1, 1, \alpha^2, \alpha^2 \sin^2 \beta)$  depending upon whether we happen to be using Cartesian  $(x, y, z)$ , cylindrical  $(R, \phi, z)$ , or spherical  $(r, \theta, \phi)$  spatial coordinate systems. The upside is that we always have the freedom to work with coordinates that simplify our equations, that make physical properties of our solutions more transparent, or that join smoothly on to a favourite flat space coordinates at large distances from a source. The downside is that in a complicated problem, it is far from easy to know the best coordinates to be using.

Coordinate freedom is particularly useful for gravitational radiation. You may remember when you studied electromagnetic radiation that the equations for the potentials (both  $\mathbf{A}$  and  $\Phi$ ) simplified considerably when a particular “gauge” was used—the Lorenz gauge. A different gauge could have been used, in which case the *potentials* would certainly have looked different, but the physical *fields* would have been just the same. The same is true for gravitational radiation. Here, coordinate transformations play this gauge role, but in a rather peculiar way: we change the components of  $g_{\mu\nu}$  as though a coordinate transformation were taking place, yet we actually keep our working coordinates the same! (Which is why we call it a gauge transformation.) What seems like an elementary blunder is actually a perfectly correct thing to do, and will be explained more fully in Chapter 7.

For the problem of determining  $g_{\mu\nu}$  around a point mass—the Schwarzschild black hole—we will choose to work with coordinates that look as much as possible like standard spherical coordinates.

## 6.3 The Schwarzschild Solution

We wish to determine the form of the metric tensor  $g_{\mu\nu}$  for the spacetime surrounding a point mass  $M$  by solving the equation  $R_{\mu\nu} = 0$ , subject to the appropriate boundary conditions.

Because the spacetime is static and spherically symmetric, we expect the invariant line element to take the form

$$-c^2 d\tau^2 = -B c^2 dt^2 + A dr^2 + C d\Omega^2 \quad (251)$$

---

<sup>5</sup>Don't smirk. If we're using awkward coordinates, it can be *very* hard to tell. You'll see.

where  $d\Omega$  is the (undistorted) solid angle,

$$d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2$$

and  $A$ ,  $B$ , and  $C$  are all functions of the radial variable. We may choose our coordinates so that  $C$  is defined to be  $r^2$  (if it is not already, do a coordinate transformation  $r'^2 = C(r)$  and then drop the  $'$ ).  $A$  and  $B$  will then be some unknown functions of  $r$  to be determined. Our metric is now in “standard form:”

$$-c^2 d\tau^2 = -B(r) c^2 dt^2 + A(r) dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \quad (252)$$

We may read off the components of  $g_{\mu\nu}$ :

$$g_{tt} = -B(r) \quad g_{rr} = A(r) \quad g_{\theta\theta} = r^2 \quad g_{\phi\phi} = r^2 \sin^2 \theta \quad (253)$$

and its inverse  $g^{\mu\nu}$ ,

$$g^{tt} = -B^{-1}(r) \quad g^{rr} = A^{-1}(r) \quad g^{\theta\theta} = r^{-2} \quad g^{\phi\phi} = r^{-2} (\sin \theta)^{-2} \quad (254)$$

The determinant of  $g_{\mu\nu}$  is  $-g$ , where

$$g = r^4 AB \sin^2 \theta \quad (255)$$

We have seen that the affine connection for a *diagonal* metric tensor will be of the form

$$\Gamma_{ab}^a = \Gamma_{ba}^a = \frac{1}{2g_{aa}} \frac{\partial g_{aa}}{\partial x^b}$$

no sum on  $a$ , with  $a = b$  permitted; or

$$\Gamma_{bb}^a = -\frac{1}{2g_{aa}} \frac{\partial g_{bb}}{\partial x^a}$$

no sum on  $a$  or  $b$ , with  $a$  and  $b$  distinct. The nonvanishing components follow straightforwardly:

$$\begin{aligned} \Gamma_{tr}^t &= \Gamma_{rt}^t = \frac{B'}{2B} \\ \Gamma_{tt}^r &= \frac{B'}{2A} \quad \Gamma_{rr}^r = \frac{A'}{2A} \quad \Gamma_{\theta\theta}^r = -\frac{r}{A} \quad \Gamma_{\phi\phi}^r = -\frac{r \sin^2 \theta}{A} \\ \Gamma_{r\theta}^\theta &= \Gamma_{\theta r}^\theta = \frac{1}{r} \quad \Gamma_{\phi\phi}^\theta = -\sin \theta \cos \theta \\ \Gamma_{\phi r}^\phi &= \Gamma_{r\phi}^\phi = \frac{1}{r} \quad \Gamma_{\phi\theta}^\phi = \Gamma_{\theta\phi}^\phi = \cot \theta \end{aligned} \quad (256)$$

where  $A' = dA/dr$ ,  $B' = dB/dr$ .

Next, we need the Ricci Tensor:

$$R_{\mu\kappa} \equiv R_{\mu\lambda\kappa}^\lambda = \frac{\partial \Gamma_{\mu\lambda}^\lambda}{\partial x^\kappa} - \frac{\partial \Gamma_{\mu\kappa}^\lambda}{\partial x^\lambda} + \Gamma_{\mu\lambda}^\eta \Gamma_{\kappa\eta}^\lambda - \Gamma_{\mu\kappa}^\eta \Gamma_{\lambda\eta}^\lambda \quad (257)$$

Remembering equation (173), this may be written

$$R_{\mu\kappa} = \frac{1}{2} \frac{\partial^2 \ln g}{\partial x^\kappa \partial x^\mu} - \frac{\partial \Gamma_{\mu\kappa}^\lambda}{\partial x^\lambda} + \Gamma_{\mu\lambda}^\eta \Gamma_{\kappa\eta}^\lambda - \frac{\Gamma_{\mu\kappa}^\eta}{2} \frac{\partial \ln g}{\partial x^\eta} \quad (258)$$

Right. First  $R_{tt}$ . Remember, static fields.

$$\begin{aligned} R_{tt} &= -\frac{\partial \Gamma_{tt}^r}{\partial r} + \Gamma_{t\lambda}^\eta \Gamma_{t\eta}^\lambda - \frac{\Gamma_{tt}^\eta}{2} \frac{\partial \ln |g|}{\partial x^\eta} \\ &= -\frac{\partial}{\partial r} \left( \frac{B'}{2A} \right) + \Gamma_{tr}^t \Gamma_{tt}^r + \Gamma_{tt}^r \Gamma_{tr}^t - \frac{\Gamma_{tt}^r}{2} \frac{\partial \ln g}{\partial r} \\ &= -\left( \frac{B''}{2A} \right) + \frac{B' A'}{2A^2} + \frac{B'^2}{4AB} + \cancel{\frac{B'^2}{4AB}} - \frac{B'}{4A} \left( \frac{A'}{A} + \cancel{\frac{B'}{B}} + \frac{4}{r} \right) \end{aligned}$$

This gives

$$R_{tt} = -\frac{B''}{2A} + \frac{B'}{4A} \left( \frac{B'}{B} + \frac{A'}{A} \right) - \frac{B'}{rA} \quad (259)$$

Next,  $R_{rr}$ :

$$\begin{aligned} R_{rr} &= \frac{1}{2} \frac{\partial^2 \ln g}{\partial r^2} - \frac{\partial \Gamma_{rr}^r}{\partial r} + \Gamma_{r\lambda}^\eta \Gamma_{r\eta}^\lambda - \frac{\Gamma_{rr}^r}{2} \frac{\partial \ln g}{\partial r} \\ &= \frac{1}{2} \frac{\partial}{\partial r} \left( \cancel{\frac{A'}{A}} + \frac{B'}{B} + \frac{4}{r} \right) - \cancel{\frac{\partial}{\partial r} \left( \frac{A'}{2A} \right)} + \Gamma_{r\lambda}^\eta \Gamma_{r\eta}^\lambda - \frac{A'}{4A} \left( \frac{A'}{A} + \frac{B'}{B} + \frac{4}{r} \right) \\ &= \frac{B''}{2B} - \frac{1}{2} \left( \frac{B'}{B} \right)^2 - \frac{2}{r^2} + (\Gamma_{rt}^t)^2 + (\Gamma_{rr}^r)^2 + (\Gamma_{r\theta}^\theta)^2 + (\Gamma_{r\phi}^\phi)^2 - \frac{1}{4} \left( \frac{A'}{A} \right)^2 - \frac{A'B'}{4AB} - \frac{A'}{rA} \\ &= \frac{B''}{2B} - \frac{1}{2} \left( \frac{B'}{B} \right)^2 - \cancel{\frac{2}{r^2}} + \frac{B'^2}{4B^2} + \cancel{\frac{A'^2}{4A^2}} + \cancel{\frac{1}{r^2}} + \cancel{\frac{1}{r^2}} - \cancel{\frac{1}{4} \left( \frac{A'}{A} \right)^2} - \frac{A'B'}{4AB} - \frac{A'}{rA} \end{aligned}$$

so that finally we arrive at

$$R_{rr} = \frac{B''}{2B} - \frac{1}{4} \frac{B'}{B} \left( \frac{A'}{A} + \frac{B'}{B} \right) - \frac{A'}{rA} \quad (260)$$

Tired? Well, here is a spoiler: all we will need for the problem at hand is in  $R_{tt}$  and  $R_{rr}$ , so you can now skip right now to the end of the section. For the true fanatics, we are just getting warmed up. On to  $R_{\theta\theta}$ —have at it!

$$\begin{aligned} R_{\theta\theta} &= \frac{\partial \Gamma_{\theta\lambda}^\lambda}{\partial \theta} - \frac{\partial \Gamma_{\theta\theta}^\lambda}{\partial x^\lambda} + \Gamma_{\theta\lambda}^\eta \Gamma_{\theta\eta}^\lambda - \Gamma_{\theta\theta}^\eta \Gamma_{\lambda\eta}^\lambda \\ &= \frac{1}{2} \frac{\partial^2 \ln g}{\partial \theta^2} - \frac{\partial \Gamma_{\theta\theta}^r}{\partial r} + \Gamma_{\theta\lambda}^\eta \Gamma_{\theta\eta}^\lambda - \Gamma_{\theta\theta}^r \Gamma_{\lambda r}^\lambda \\ &= \frac{d(\cot \theta)}{d\theta} + \frac{d}{dr} \left( \frac{r}{A} \right) + \Gamma_{\theta\lambda}^\eta \Gamma_{\theta\eta}^\lambda + \frac{r}{2A} \frac{\partial \ln g}{\partial r} \\ &= -\frac{1}{\sin^2 \theta} + \frac{1}{A} - \frac{rA'}{A^2} + \Gamma_{\theta\lambda}^r \Gamma_{\theta r}^\lambda + \Gamma_{\theta\lambda}^\theta \Gamma_{\theta\theta}^\lambda + \Gamma_{\theta\lambda}^\phi \Gamma_{\theta\phi}^\lambda + \frac{r}{2A} \left( \frac{A'}{A} + \frac{B'}{B} + \frac{4}{r} \right) \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{\sin^2 \theta} + \frac{3}{A} - \frac{rA'}{2A^2} + \Gamma_{\theta\theta}^r \Gamma_{\theta r}^\theta + \Gamma_{\theta r}^\theta \Gamma_{\theta\theta}^r + \left(\Gamma_{\theta\phi}^\phi\right)^2 + \frac{rB'}{2AB} \\
&= -\frac{1}{\sin^2 \theta} + \frac{3}{A} - \frac{rA'}{2A^2} - \frac{2}{A} + \cot^2 \theta + \frac{rB'}{2AB}
\end{aligned}$$

The trigonometric terms add to  $-1$ . We finally obtain

$$R_{\theta\theta} = -1 + \frac{1}{A} + \frac{r}{2A} \left( -\frac{A'}{A} + \frac{B'}{B} \right) \quad (261)$$

$R_{\phi\phi}$  is the last nonvanishing Ricci component. No whining now! The first term in (257) vanishes, since nothing in the metric depends on  $\phi$ . Then,

$$\begin{aligned}
R_{\phi\phi} &= -\frac{\partial \Gamma_{\phi\phi}^\lambda}{\partial x^\lambda} + \Gamma_{\phi\lambda}^\eta \Gamma_{\phi\eta}^\lambda - \frac{\Gamma_{\phi\phi}^\eta}{2} \frac{\partial \ln |g|}{\partial x^\eta} \\
&= -\frac{\partial \Gamma_{\phi\phi}^r}{\partial r} - \frac{\partial \Gamma_{\phi\phi}^\theta}{\partial \theta} + \Gamma_{\phi\lambda}^r \Gamma_{\phi r}^\lambda + \Gamma_{\phi\lambda}^\theta \Gamma_{\phi\theta}^\lambda + \Gamma_{\phi\lambda}^\phi \Gamma_{\phi\phi}^\lambda - \frac{1}{2} \Gamma_{\phi\phi}^r \frac{\partial \ln |g|}{\partial r} - \frac{1}{2} \Gamma_{\phi\phi}^\theta \frac{\partial \ln |g|}{\partial \theta} \\
&= \frac{\partial}{\partial r} \left( \frac{r \sin^2 \theta}{A} \right) + \frac{\partial}{\partial \theta} (\sin \theta \cos \theta) + \Gamma_{\phi\phi}^r \Gamma_{\phi r}^\phi + \Gamma_{\phi\phi}^\theta \Gamma_{\phi\theta}^\phi + \Gamma_{\phi r}^\phi \Gamma_{\phi\phi}^r + \Gamma_{\phi\theta}^\phi \Gamma_{\phi\phi}^\theta \\
&\quad + \frac{1}{2} \sin \theta \cos \theta \frac{\partial \ln \sin^2 \theta}{\partial \theta} + \frac{1}{2} \left( \frac{r \sin^2 \theta}{A} \right) \left( \frac{A'}{A} + \frac{B'}{B} + \frac{4}{r} \right) \\
&= \cancel{\frac{\sin^2 \theta}{A}} - \frac{rA' \sin^2 \theta}{A^2} + \cancel{\cos^2 \theta} - \sin^2 \theta - \cancel{\frac{\sin^2 \theta}{A}} - \cancel{\cos^2 \theta} - \frac{\sin^2 \theta}{A} - \cancel{\cos^2 \theta} + \cancel{\cos^2 \theta} + \frac{r \sin^2 \theta}{2A} \left( \frac{A'}{A} + \frac{B'}{B} + \frac{4}{r} \right) \\
&= \sin^2 \theta \left[ \frac{r}{2A} \left( -\frac{A'}{A} + \frac{B'}{B} \right) + \frac{1}{A} - 1 \right] = \sin^2 \theta R_{\theta\theta}
\end{aligned}$$

The fact that  $R_{\phi\phi} = \sin^2 \theta R_{\theta\theta}$  and that  $R_{\mu\nu} = 0$  if  $\mu$  and  $\nu$  are not equal are consequences of the spherical symmetry and time reversal symmetry of the problem respectively. If the first relation did not hold, or if  $R_{ij}$  did not vanish when  $i$  and  $j$  were different spatial coordinates, then an ordinary rotation of the axes would change the relative form of the tensor components, despite the spherical symmetry. This is impossible. If  $R_{ti} \equiv R_{it}$  were non-vanishing ( $i$  is again a spatial index), the coordinate transformation  $t' = -t$  would change the components of the Ricci tensor. But a static  $R_{\mu\nu}$  must be invariant to this form of time reversal coordinate change. (Why?) Note that this argument is *not* true for  $R_{tt}$ . (Why not?)

Learn to think like a mathematical physicist in this kind of a calculation, taking into account the symmetries that are present, and you will save a lot of work.

**Exercise. Self-gravitating masses in general relativity.** We are solving in this section the vacuum equations  $R_{\mu\nu} = 0$ , but it is of great interest for stellar structure and cosmology to have a set of equations for a self-gravitating spherical mass. Toward that end, we recall equation (247):

$$R_{\mu\nu} = -\frac{8\pi G}{c^4} S_{\mu\nu} \equiv -\frac{8\pi G}{c^4} \left( T_{\mu\nu} - \frac{g_{\mu\nu}}{2} T^\lambda{}_\lambda \right)$$

Let us evaluate  $S_{\mu\nu}$  for the case of an isotropic stress energy tensor of an ideal gas in its rest frame. With

$$g_{tt} = -B, \quad g_{rr} = A, \quad g_{\theta\theta} = r^2, \quad g_{\phi\phi} = r^2 \sin^2 \theta,$$

the stress-energy tensor

$$T_{\mu\nu} = P g_{\mu\nu} + (\rho + P/c^2) U_\mu U_\nu,$$

where  $U_\mu$  is the 4-velocity, show that, in addition to the trivial condition

$$U_r = U_\theta = U_\phi = 0,$$

we must have  $U_t = -c\sqrt{B}$  (remember equation [182]) and that

$$S_{tt} = \frac{B}{2}(3P + \rho c^2), \quad S_{rr} = \frac{A}{2}(\rho c^2 - P), \quad S_{\theta\theta} = \frac{r^2}{2}(\rho c^2 - P)$$

We will develop the solutions of  $R_{\mu\nu} = -8\pi G S_{\mu\nu}/c^4$  shortly.

Enough. We have more than we need to solve the problem at hand. To solve the equations  $R_{\mu\nu} = 0$  is now a rather easy task. Two components will suffice (we have only  $A$  and  $B$  to solve for after all), all others then vanish identically. In particular, work with  $R_{rr}$  and  $R_{tt}$ , both of which must separately vanish, so

$$\frac{R_{rr}}{A} + \frac{R_{tt}}{B} = -\frac{1}{rA} \left( \frac{A'}{A} + \frac{B'}{B} \right) = 0 \quad (262)$$

whence we find

$$AB = \text{constant} = 1 \quad (263)$$

where the constant must be unity since  $A$  and  $B$  go over to their Minkowski values at large distances. The condition that  $R_{tt} = 0$  is now from (259) simply

$$B'' + \frac{2B'}{r} = 0, \quad (264)$$

which means that  $B$  is a linear superposition of a constant plus another constant times  $1/r$ . But  $B$  must approach unity at large  $r$ , so the first constant is one, and we know from long ago that the next order term at large distances must be  $2\Phi/c^2$  in order to recover the Newtonian limit. Hence,

$$B = 1 - \frac{2GM}{rc^2}, \quad A = \left( 1 - \frac{2GM}{rc^2} \right)^{-1} \quad (265)$$

The *Schwarzschild Metric* for the spacetime around a point mass is exactly

$$\boxed{-c^2 d\tau^2 = - \left( 1 - \frac{2GM}{rc^2} \right) c^2 dt^2 + \left( 1 - \frac{2GM}{rc^2} \right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2} \quad (266)$$

This remarkable, simple and critically important exact solution of the Einstein Field Equation was obtained in 1916 by Karl Schwarzschild from the trenches of World War I. Tragically, Schwarzschild did not survive the war,<sup>6</sup> dying from a skin infection five months after finding his marvelous solution. He managed to communicate his result fully in a letter to Einstein. His final correspondence to Einstein was dated 22 December 1915, 28 days after the formulation of the Field Equations.

---

<sup>6</sup>The senseless WWI deaths of Karl Schwarzschild for the Germans and Oxford's Henry Moseley for the British were incalculable losses for science. Schwarzschild's son Martin, a 4-year-old at the time of his father's death, also became a great astrophysicist, developing much of the modern theory of stellar evolution.

**Exercise. The Tolman-Oppenheimer-Volkoff Equation.** Let us strike again while the iron is hot. Referring back to the previous exercise, we repeat part of our Schwarzschild calculation, but with the source terms  $S_{\mu\nu}$  retained. Form a familiar combination once again:

$$\frac{R_{rr}}{A} + \frac{R_{tt}}{B} = -\frac{1}{rA} \left( \frac{A'}{A} + \frac{B'}{B} \right) = -\frac{8\pi G}{c^4} \left( \frac{S_{tt}}{B} + \frac{S_{rr}}{A} \right) = -\frac{8\pi G}{c^4} (P + \rho c^2)$$

Show now that adding  $2R_{\theta\theta}/r^2$  eliminates the  $B$  dependence:

$$\frac{R_{rr}}{A} + \frac{R_{tt}}{B} + \frac{2R_{\theta\theta}}{r^2} = -\frac{2A'}{rA^2} - \frac{2}{r^2} + \frac{2}{Ar^2} = -\frac{16\pi G\rho}{c^2}.$$

Solve this equation for  $A$  and show that the solution with  $A(0) = 1$  at the centre of the star is

$$A(r) = \left( 1 - \frac{2G\mathcal{M}(r)}{c^2 r} \right)^{-1}, \quad \mathcal{M}(r) = \int_0^r 4\pi\rho(r') r'^2 dr'$$

Why is this the correct boundary condition?

Finally, use the equation  $R_{\theta\theta} = -8G\pi S_{\theta\theta}/c^4$  together with hydrostatic equilibrium (186) (for the term  $B'/B$  in  $R_{\theta\theta}$ ) to obtain the celebrated Tolman-Oppenheimer-Volkoff equation for the interior structure of general relativistic stars:

$$\frac{dP}{dr} = -\frac{G\mathcal{M}(r)\rho}{r^2} \left( 1 + \frac{P}{\rho c^2} \right) \left( 1 + \frac{4\pi r^3 P}{\mathcal{M}(r) c^2} \right) \left( 1 - \frac{2G\mathcal{M}(r)}{rc^2} \right)^{-1}$$

Note: this is a rather long, but completely straightforward, exercise.

Students of stellar structure will recognise the classical equation hydrostatic equilibrium equation for a Newtonian star, with three correction terms. The final factor on the right is purely geometrical, the radial curvature term  $A$  from the metric. The corrective replacement of  $\rho$  by  $\rho + P/c^2$  arises even in the special relativistic equations of motion for the inertial density; for inertial purposes  $P/c^2$  is an effective density. Finally the modification of the gravitating  $\mathcal{M}(r)$  term (to  $\mathcal{M}(r) + 4\pi r^3 P/c^2$ ) also includes a contribution from the pressure, as though an additional effective mass density  $3P(r)/c^2$  were spread throughout the interior spherical volume within  $r$ , even though  $P(r)$  is just the local pressure. Note that in massive stars, this pressure could be *radiative*!

## 6.4 The Schwarzschild Radius

It will not have escaped the reader's attention that at

$$r = \frac{2GM}{c^2} \equiv R_S \quad (267)$$

the metric becomes singular in appearance.  $R_S$  is known as the Schwarzschild radius. Numerically, normalising  $M$  to one solar mass  $M_\odot$ ,

$$R_S = 2.95 (M/M_\odot) \text{ km}, \quad (268)$$

which is *well* inside any normal star! The Schwarzschild radius is part of the external vacuum spacetime only for black holes. Indeed, it is what makes black holes black. At least it was



*thought* to be the feature that made black holes truly black, until Hawking came along in 1974 and showed us that quantum field theory changes the behaviour of black holes. But as usual, we are getting ahead of ourselves. Let us stick to classical theory.

I have been careful to write “singular in appearance” because in fact, the spacetime is perfectly well behaved at  $r = R_S$ . It is only the coordinates that become strained at this point, and these coordinates have been introduced, you will recall, so that they would be familiar to us, we few, we happy band of observers at infinity, as ordinary spherical coordinates. The curvature scalar  $R$ , for example, remains zero without so much as a ripple as we pass through  $r = R_S$ . We can see this coordinate effect staring at us if we start with the ordinary metric on the unit sphere,

$$ds^2 = d\theta^2 + \sin^2 \theta d\phi^2,$$

and change coordinates to  $x = \sin \theta$ :

$$ds^2 = \frac{dx^2}{1-x^2} + x^2 d\phi^2.$$

This looks horrible at  $x = 1$ , but in reality nothing is happening. Since  $x$  is just the distance from the z-axis to spherical surface (i.e. cylindrical radius), the “singularity” simply reflects the fact that at the equator  $x$  has reached its maximum value 1. So,  $dx$  must be zero at this point.  $x$  is just a bad coordinate at the equator. But then  $\phi$  is a bad coordinate at the poles,  $\theta = 0$  or  $\theta = \pi$ , or for that matter  $x = 0$ . Bad coordinates happen to good, decent spacetimes. Get over it.

The physical interpretation of the first two terms of the metric (266) is that the proper time interval at a fixed spatial location is given by

$$dt \left( 1 - \frac{2GM}{rc^2} \right)^{1/2} \quad (\text{proper time interval at fixed location}). \quad (269)$$

The proper radial distance interval at a fixed angular location and time is

$$dr \left( 1 - \frac{2GM}{rc^2} \right)^{-1/2} \quad (\text{proper radial distance interval at fixed time \& angle}). \quad (270)$$

**Exercise. Getting rid of the Schwarzschild coordinate singularity. A challenge problem for the adventurous student only. Make sure you want to do this before you start.** Consider the rather unusual coordinate transformation found by Martin Kruskal. Start with our standard spherical coordinates  $t, r, \theta, \phi$  and introduce new  $r'$  and  $t'$  coordinates:

$$\begin{aligned} r'^2 - c^2 t'^2 &= c^2 T^2 \left( \frac{rc^2}{2GM} - 1 \right) \exp \left( \frac{rc^2}{2GM} \right) \\ \frac{2r'ct'}{r'^2 + c^2 t'^2} &= \tanh \left( \frac{c^3 t}{2GM} \right) \end{aligned}$$

where  $T$  is an arbitrary constant. Show that the Schwarzschild metric transforms to

$$-c^2 d\tau^2 = \left( \frac{32G^3 M^3}{c^8 r T^2} \right) \exp \left( \frac{-rc^2}{2GM} \right) (c^2 dt'^2 - dr'^2) - r^2 d\Omega^2$$

where  $T$  is arbitrary constant with dimensions of time, and  $r$  is the implicit solution of our first equation for  $r'^2 - c^2 t'^2$ . The right side of this equation has a minimum of  $-c^2 T^2$  at  $r = 0$ , hence we must have

$$r'^2 > c^2(t'^2 - T^2)$$

always. When  $t' < T$  there is no problem. But when  $t' > T$  there are two distinct regions:  $r' = \pm c\sqrt{t'^2 - T^2}$ . Then the metric has a real singularity at either of these values of  $r'$  (which is just  $r = 0$ ), but still no singularity at  $r' = \pm ct'$ , the value  $r = R_S$ .

## 6.5 Schwarzschild spacetime.

### 6.5.1 Radial photon geodesic

This doesn't mean that there is nothing of interest happening at  $r = R_S$ .

For starters, the gravitational redshift recorded by an observer at infinity relative to someone at rest at location  $r$  in the Schwarzschild spacetime is given (we now know) *precisely* by

$$dt = \frac{d\tau}{(1 - 2GM/rc^2)^{1/2}} \quad (\text{Exact.}) \quad (271)$$

so that at  $r \rightarrow R_S$ , signals arrive at a distant observer's post infinitely redshifted. What does this mean?

Comfortably sitting in the Clarendon Labs, monitoring the radio signals my hardworking graduate student is sending me whilst engaged on a perfectly reasonable thesis mission to take measurements of the  $r = R_S$  tidal forces in a nearby black hole, I grow increasingly impatient. Not only are the incessant complaints becoming progressively more torpid and drawn out, the transmission frequency keeps shifting to longer and longer wavelengths, slipping out of my receiver's bandpass. Most irritating. Eventually, all contact is lost. (Typical.) I never receive any signal of any kind from within  $R_S$ .  $R_S$  is said to be the location of the *event horizon*. The singularity at  $r = 0$  is present, but completely hidden from the outside world at  $R = R_S$  within an event horizon. It is what Roger Penrose has aptly named "cosmic censorship."

The time coordinate change for light to travel from  $r_A$  to  $r_B$  following its geodesic path is given by setting

$$-(1 - 2GM/rc^2)c^2 dt^2 + dr^2/(1 - 2GM/rc^2) = 0$$

and then computing

$$t_{AB} = \int_A^B dt = \frac{1}{c} \int_{r_A}^{r_B} \frac{dr}{(1 - 2GM/rc^2)} = \frac{r_B - r_A}{c} + \frac{R_S}{c} \ln \left( \frac{r_B - R_S}{r_A - R_S} \right) \quad (272)$$

which will be recognised as the Newtonian time interval plus a logarithmic correction proportional to the Schwarzschild radius  $R_S$ . Note that our expression becomes infinite when a path endpoint includes  $R_S$ . When  $R_S$  may be considered small over the entire integration path, to leading order

$$t_{AB} \simeq \frac{r_B - r_A}{c} + \frac{R_S}{c} \ln \left( \frac{r_A}{r_B} \right) = \frac{r_B - r_A}{c} \left( 1 + \frac{R_S \ln(r_A/r_B)}{r_B - r_A} \right) \quad (273)$$

A GPS satellite orbits at an altitude of 20,200 km, and the radius of the earth is 6370 km.  $R_S$  for the earth is only 9mm! (Make a fist. Squeeze the entire earth inside it. You're not even close to making a black hole.) Then, the general relativistic correction factor is

$$\frac{R_S}{r_B - r_A} \simeq \frac{9 \times 10^{-3}}{(20,200 - 6370) \times 10^3} = 6.5 \times 10^{-10}$$

This level of accuracy, about a part in  $10^9$ , is needed for determining positions on the surface of the earth to a precision of a few meters (as when your GPS intones "Turn right onto the Lon-don Road."). How does the gravitational effect compare with the *second order* kinematic time dilation due to the satellite's motion? You should find them comparable.

### 6.5.2 Orbital equations

Formally, we wish to solve the geodesic equation for the orbits around a point mass. We write the equation not as a function of  $\tau$ , but of some other parameter  $p$ :

$$\frac{d^2 x^\lambda}{dp^2} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{dp} \frac{dx^\nu}{dp} = 0 \quad (274)$$

We will assume that  $\tau$  is just  $p$  multiplied by some constant, *but that the constant is zero for a photon*. That way, we don't have to worry anything singular happening when  $d\tau = 0$  for a photon. We use some other scalar  $dp$ , a differential that is finite for both ordinary matter and photon orbits. Don't worry about what  $p$  is for now, we'll see how that all works mathematically in a moment. If you're still bothered, just think of  $dp$  as the time you view as elapsed on your own personal Casio (whatever) wristwatch. That's a good scalar!

The orbital equations themselves are most easily derived by starting with the Euler-Lagrange Equations for the Lagrangian

$$\mathcal{L} = g_{\mu\nu} \frac{dx^\mu}{dp} \frac{dx^\nu}{dp} = -B(r)c^2 \dot{t}^2 + A(r)\dot{r}^2 + r^2 \dot{\phi}^2 \quad (275)$$

where the dot now represents  $d/dp$ , just as we learned at the end of section 3.5, page 31, using  $\tau$ . (Remember  $dp$  and  $d\tau$  are just proportional to one another.) We have fixed the orbital plane to  $\theta = \pi/2$ , so that  $d\theta/dp$  vanishes identically.  $A$  and  $B$  depend explicitly on  $r$ , and implicitly on  $p$  via  $r = r(p)$ . Recall that for large  $r$ ,  $A$  and  $B$  approach unity. The Euler-Lagrange equation for time  $t$  is simple, since there is no  $\partial\mathcal{L}/\partial t$ :

$$\frac{d}{dp} \left( \frac{\partial\mathcal{L}}{\partial\dot{t}} \right) = -2c^2 \frac{d}{dp} \left( B \frac{dt}{dp} \right) = 0, \quad (276)$$

and the same holds true analogously for the  $\phi$  equation:

$$\frac{d}{dp} \left( \frac{\partial\mathcal{L}}{\partial\dot{\phi}} \right) = 2 \frac{d}{dp} \left( r^2 \frac{d\phi}{dp} \right) = 0. \quad (277)$$

These equations are both very easy to solve. Since  $p$  can be always be multiplied by some arbitrary factor, it is handy to choose the integration constant of (276) in such a way that  $p$  goes over to the time  $t$  at large  $r$  distances. Then,  $t$  is given by the equation:

$$\frac{dt}{dp} = B^{-1}. \quad (278)$$

Moreover, general relativity conserves angular momentum for a spherical spacetime just like ordinary mechanics. The angular variable  $\phi$  is given by the differential equation:

$$r^2 \frac{d\phi}{dp} = J \quad (\text{constant.}) \quad (279)$$

We could now proceed with the formal Euler-Lagrange equation for  $r$ , but there is a much more direct option available to us. Just write down directly the line element for  $(d\tau/dp)^2$ , and use the last two results we just derived for  $dt/dp$  and  $d\phi/dp$ .

$$A \left( \frac{dr}{dp} \right)^2 + \frac{J^2}{r^2} - \frac{c^2}{B} = -c^2 \left( \frac{d\tau}{dp} \right)^2 \equiv -E \quad (\text{constant}) \quad (280)$$

where we have introduced our mysterious proportionality factor linking  $d\tau$  and  $dp$  via  $c^2 d\tau^2 = E dp^2$ , i.e.  $p$  and  $\tau$  differ only by the proportionality constant  $E$ . (Notice the flexibility we have: had we chosen a different integration constant in (278), we could now just absorb it into our definition  $E$ .) For ordinary matter,  $E > 0$ , while  $E = 0$  for photons. Far from the Schwarzschild radius of the point mass, assuming that we are still on the orbit, to leading Newtonian order the final term on the left  $-c^2/B$  dominates over the first two. With  $B \sim 1$ , we then find  $E \simeq c^2$ , which is just the rest mass energy per unit mass. The small difference between  $E$  and  $c^2$  is proportional to the Newtonian energy. Substituting for  $B$  in (280), we note more generally that the extremal (maximum and minimum) values of orbital radius  $r$  for a bound orbit (when  $\dot{r} = 0$ ) are solutions of

$$\left( 1 - \frac{2GM}{rc^2} \right) \left( \frac{J^2}{r^2} + E \right) - c^2 = 0 \quad (281)$$

for ordinary matter, and of

$$\left( 1 - \frac{2GM}{rc^2} \right) \frac{J^2}{r^2} - c^2 = 0 \quad (282)$$

for photons. (A bound photon orbit can only be circular, and the orbit is actually unstable.)

The radial equation of motion may be written for either  $dr/d\tau$ ,  $dr/dt$ , or  $dr/d\phi$  respectively (we use  $AB = 1$ ):

$$\left( \frac{dr}{d\tau} \right)^2 + Bc^2 \left( 1 + \frac{J^2}{Er^2} \right) = \frac{c^4}{E} \quad (283)$$

$$\left( \frac{dr}{dt} \right)^2 + B^3 \left( E + \frac{J^2}{r^2} \right) = B^2 c^2 \quad (284)$$

$$\left( \frac{dr}{d\phi} \right)^2 + r^2 B \left( 1 + \frac{Er^2}{J^2} \right) = \frac{c^2 r^4}{J^2} \quad (285)$$

From here on, it is only a matter of evaluating a (perhaps complicated) integral over  $r$  to obtain the explicit solution for the orbit. We can always do this task numerically. When we are looking only for small perturbations from Newtonian theory, as discussed in the next section of classical tests of general relativity, analytic progress is possible.

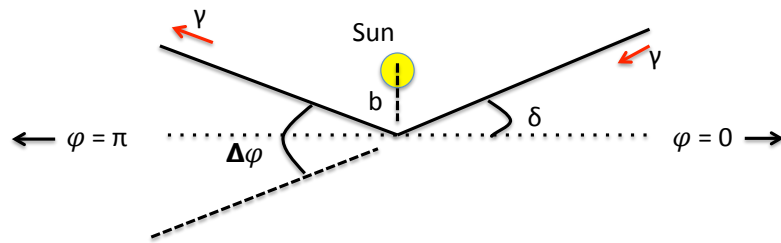


Figure 2: Bending of light by the gravitational field of the sun. In flat spacetime the photon  $\gamma$  travels the straight line from  $\varphi = 0$  to  $\varphi = \pi$  along the path  $r \sin \varphi = b$ . The presence of spacetime curvature starts the photon at  $\varphi = -\delta$  and finishes its passage at  $\varphi = \pi + \delta$ . The deflection angle is  $\Delta\varphi = 2\delta$ .

## 6.6 The deflection of light by an intervening body.

The first prediction made by General Relativity Theory that could be tested was that starlight passing by the limb of the sun would be slightly, but measurably, deflected by the gravitational field. This type of measurement can only be done, of course, when the sun is completely eclipsed by the moon. Fortunately, the timing of the appearance of Einstein's theory with an eclipse was ideal. One of the longest total solar eclipses of the 20th century occurred on 29 May 1919. The path of totality extended along a narrow strip of the earth starting in South America and continuing to central Africa. An expedition headed by Arthur Eddington observed the eclipse from the island of Principe, just off the west coast of Africa. Measurements of the angular shifts of thirteen stars confirmed not only that gravity certainly affected the propagation of light, but that it did so by an amount in much better accord with general relativity theory than with a Newtonian “corpuscular theory,” with the test mass velocity set equal to  $c$ . (The latter gives a deflection angle only half as large as GR, in essence because the  $2GM/rc^2$  terms in both the  $dt$  and  $dr$  metric coefficients contribute equally to the photon deflection, whereas in the Newtonian limit only the modification in the  $dt$  metric coefficient is retained.) This success earned Einstein press coverage that today is normally reserved for rock stars. *Everybody* knew who Albert Einstein was!

Today, not only mere deflection, but “gravitational lensing” and actual image formation (across the electromagnetic spectrum) are standard astronomical techniques to probe intervening matter in all of its forms: from small planets to huge, diffuse cosmological agglomerations of dark matter. The weak lensing caused by the presence of the latter is the target of the *Euclid* space-based telescope mission to be launched in 2021.

Let us return to the classic test. As in Newtonian dynamics, it turns out to be easier to work with  $u \equiv 1/r$ , in which case

$$\left(\frac{du}{d\phi}\right)^2 = \frac{1}{r^4} \left(\frac{dr}{d\phi}\right)^2. \quad (286)$$

Equation (285) with  $E = 0$  for a photon may be written

$$\frac{1}{r^4} \left(\frac{dr}{d\phi}\right)^2 + \frac{B}{r^2} = \frac{c^2}{J^2} = \text{constant} \quad (287)$$

In terms of  $u$ :

$$\left(\frac{du}{d\phi}\right)^2 + u^2 \left(1 - \frac{2GMu}{c^2}\right) = \frac{c^2}{J^2} \quad (288)$$

Differentiating with respect to  $\phi$  ( $du/d\phi \equiv u'$ ) leads quickly to

$$u'' + u = \frac{3GM}{c^2} u^2 \equiv 3\epsilon u^2. \quad (289)$$

We treat  $\epsilon \equiv GM/c^2$  as a small parameter. We expand  $u$  as  $u = u_0 + u_1$ , with  $u_1 = O(\epsilon u_0) \ll u_0$  (read “ $u_1$  is of order  $\epsilon$  times  $u_0$  and much smaller than  $u_0$ ”). Then, terms of order unity must obey the equation

$$u_0'' + u_0 = 0, \quad (290)$$

and the terms of order  $\epsilon$  must obey the equation

$$u_1'' + u_1 = 3\epsilon u_0^2. \quad (291)$$

To leading order ( $u = u_0$ ), nothing happens: the photon moves in a straight line. If the point of closest approach is the impact parameter  $b$ , then the equation for a straight line is  $r \sin \phi = b$ , or

$$u_0 = \frac{\sin \phi}{b} \quad (292)$$

which is the unique solution to equation (290) with boundary conditions  $r = \infty$  at  $\phi = 0$  and  $\phi = \pi$ .

At order  $\epsilon$ , there is a deflection from a straight line due to the presence of  $u_1$ :

$$u_1'' + u_1 = 3\epsilon u_0^2 = \frac{3\epsilon}{b^2} \sin^2 \phi = \frac{3\epsilon}{2b^2} (1 - \cos 2\phi) \quad (293)$$

Clearly, we need to search for solutions of the form  $u_1 = U + V \cos 2\phi$ , where  $U$  and  $V$  are constants. Substituting this into (293), we easily find  $U = 3\epsilon/2b^2$  and  $V = \epsilon/2b^2$ . Our solution is then

$$\frac{1}{r} = u_0 + u_1 = \frac{\sin \phi}{b} + \frac{3\epsilon}{2b^2} + \frac{\epsilon \cos 2\phi}{2b^2} \quad (294)$$

With  $\epsilon = 0$ , the solution describes a straight line,  $r \sin \phi = b$ . The first order effects of including  $\epsilon$  incorporate the tiny deflections from this straight line. The  $\epsilon = 0$  solution sends  $r$  off to infinity at  $\phi = 0$  and  $\phi = \pi$ . We may compute the leading order small changes to these two “infinity angles” by using  $\phi = 0$  and  $\phi = \pi$  in the correction  $\epsilon \cos 2\phi$  term. Then we find that  $r$  goes off to infinity not at  $\phi = 0$  and  $\pi$ , but at the slightly corrected values  $\phi = -\delta$  and  $\phi = \pi + \delta$  where

$$\delta = \frac{2\epsilon}{b} \quad (295)$$

In other words, there is now a total deflection angle  $\Delta\phi$  from a straight line of  $2\delta$ , or

$$\Delta\phi = \frac{4GM}{bc^2} = 1.75 \text{ arcseconds for the Sun.} \quad (296)$$

Happily, arcsecond deflections were just at the limit of reliable photographic methods of measurement in 1919. Those arcsecond deflections unleashed a truly revolutionary paradigm shift. For once, the word is not an exaggeration.

## 6.7 The advance of the perihelion of Mercury

For Einstein personally, the revolution had started earlier, even before he had his Field Equations. The vacuum form of the Field Equations is, as we know, sufficient to describe the spacetime outside the gravitational source bodies themselves. Working with the equation  $R_{\mu\nu} = 0$ , Einstein found, and on 18 November 1915 presented, the explanation of a 60-year-old astronomical puzzle: what was the cause of Mercury’s excess perihelion advance of 43'' per century? The directly measured perihelion advance is actually much larger than this, but after the interactions from all the planets are taken into account, the excess 43'' per century is an unexplained residual of 7.5% of the total. According to Einstein’s biographer A. Pais, the discovery that this precise perihelion advance emerged from general relativity was

“...by far the strongest emotional experience in Einstein’s scientific life, perhaps in all his life. Nature had spoken to him. He had to be right.”

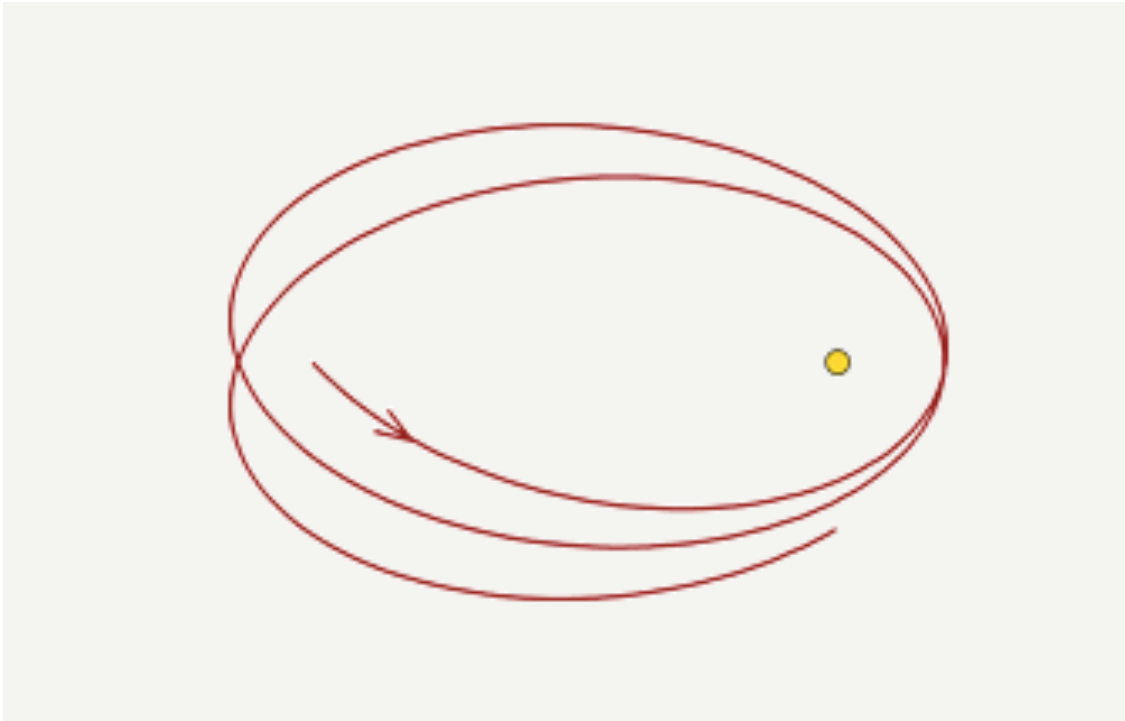


Figure 3: Departures from a  $1/r$  gravitational potential cause elliptical orbits not to close. In the case of Mercury, the perihelion advances by 43 seconds of arc per century. The effect is shown here, *greatly* exaggerated.



### 6.7.1 Newtonian orbits

Interestingly, the perihelion first-order GR calculation is not much more difficult than straight Newtonian. GR introduces a  $1/r^2$  term in the effective gravitational potential, but there is already a  $1/r^2$  term from the centrifugal term! Other corrections do not add substantively to the difficulty. We thus begin with a detailed review of the Newtonian problem, and we will play off this solution for the GR perihelion advance.

Conservation of energy is

$$\frac{v_r^2}{2} + \frac{J^2}{2r^2} - \frac{GM}{r} = \mathcal{E} \quad (297)$$

where  $J$  is the (constant) specific angular momentum  $r^2 d\phi/dt$  and  $\mathcal{E}$  is the constant energy per unit mass. (In this Newtonian case, when the two bodies have comparable masses,  $M$  is actually the sum of the individual masses, and  $r$  the relative separation of the two bodies.) This is just the low energy limit of (283), whose exact form we may write as

$$\frac{1}{2} \left( \frac{dr}{d\tau} \right)^2 + \frac{c^2}{E} \left( \frac{J^2}{2r^2} \right) - \frac{GM}{r} \left( 1 + \frac{J^2}{r^2 E} \right) = \left( \frac{c^2 - E}{2E} \right) c^2. \quad (298)$$

We now identify  $E$  with  $c^2$  to leading order, and to next order  $(c^2 - E)/2$  with  $\mathcal{E}$  (i.e. the mechanical energy above and beyond the rest mass energy). The Newtonian equation may be written

$$v_r = \frac{dr}{d\phi} \frac{d\phi}{dt} = \frac{J}{r^2} \frac{dr}{d\phi} = \pm \left( 2\mathcal{E} + \frac{2GM}{r} - \frac{J^2}{r^2} \right)^{1/2} \quad (299)$$

and thence separated:

$$\int \frac{J dr}{r^2 \left( 2\mathcal{E} + \frac{2GM}{r} - \frac{J^2}{r^2} \right)^{1/2}} = \pm \phi \quad (300)$$

With  $u = 1/r$ ,

$$\int \frac{du}{\left( \frac{2\mathcal{E}}{J^2} + \frac{2GMu}{J^2} - u^2 \right)^{1/2}} = \mp \phi \quad (301)$$

or

$$\int \frac{du}{\left[ \frac{2\mathcal{E}}{J^2} + \frac{G^2 M^2}{J^4} - \left( u - \frac{GM}{J^2} \right)^2 \right]^{1/2}} = \mp \phi \quad (302)$$

Don't be put off by all the fluff. The integral is standard trigonometric,

$$\int dy / \sqrt{A^2 - y^2} = -\cos^{-1}(y/A),$$

giving us:

$$\cos^{-1} \left[ \frac{u - \frac{GM}{J^2}}{\left( \frac{2\mathcal{E}}{J^2} + \frac{G^2 M^2}{J^4} \right)^{1/2}} \right] = \pm \phi \quad (303)$$

In terms of  $r = 1/u$  this equation unfolds and simplifies down to

$$r = \frac{J^2/GM}{1 + \epsilon \cos \phi}, \quad \epsilon^2 \equiv 1 + \frac{2\mathcal{E}J^2}{G^2M^2} \quad (304)$$

With  $\mathcal{E} < 0$  we find that  $\epsilon < 1$ , and that (304) is just the equation for a classical elliptical orbit of eccentricity  $\epsilon$ . We identify the *semilatus rectum*,

$$L = J^2/GM \quad (305)$$

the perihelion (radius of closest approach)  $r_-$  and the aphelion (radius of farthest extent)  $r_+$ ,

$$r_- = \frac{L}{1 + \epsilon}, \quad r_+ = \frac{L}{1 - \epsilon}, \quad \frac{1}{L} = \frac{1}{2} \left( \frac{1}{r_+} + \frac{1}{r_-} \right) \quad (306)$$

and the semi-major axis

$$a = \frac{1}{2}(r_+ + r_-), \quad \text{whence } L = a(1 - \epsilon^2) \quad (307)$$

Notice that the zeros of the denominator in the integral (302) occur at  $u_- = 1/r_-$  and  $u_+ = 1/r_+$ , corresponding in our arccosine function to  $\phi$  equals 0 and  $\pi$  respectively.

**Exercise.) The Shows must go on.** Show that the semi-minor axis of an ellipse is  $b = a\sqrt{1 - \epsilon^2}$ . Show that the area of an ellipse is  $\pi ab$ . Show that the total energy of a two-body bound system (masses  $m_1$  and  $m_2$ ) is  $-Gm_1m_2/2a$ , independent of  $\epsilon$ . With  $M = m_1 + m_2$ , show that the period of a two-body bound system is  $2\pi\sqrt{a^3/GM}$ , independent of  $\epsilon$ . (There is a very simple way to do the latter!)

### 6.7.2 The perihelion advance of Mercury

Equation (285) may be written in terms of  $u = 1/r$  as

$$\left( \frac{du}{d\phi} \right)^2 + \left( 1 - \frac{2GMu}{c^2} \right) \left( u^2 + \frac{E}{J^2} \right) = \frac{c^2}{J^2}. \quad (308)$$

Now differentiate with respect to  $\phi$  and simplify. The resulting equation is:

$$u'' + u = \frac{GME}{c^2J^2} + \frac{3GMu^2}{c^2} \simeq \frac{GM}{J^2} + \frac{3GMu^2}{c^2}, \quad (309)$$

since  $E$  is very close to  $c^2$  for a nonrelativistic Mercury, and the difference here is immaterial. The Newtonian limit corresponds to dropping the final term on the right side of the equation; the resulting solution is

$$u \equiv u_N = \frac{GM}{J^2}(1 + \epsilon \cos \phi) \quad \text{or} \quad r = \frac{J^2/GM}{1 + \epsilon \cos \phi} \quad (310)$$

where  $\epsilon$  is an arbitrary constant. This is just the classic equation for a conic section, with hyperbolic ( $\epsilon > 1$ ), parabolic ( $\epsilon = 1$ ) and ellipsoidal ( $\epsilon < 1$ ) solutions. For ellipses,  $\epsilon$  is the eccentricity.

As the general relativistic term  $3GMu^2/c^2$  is tiny, we are entirely justified in using the Newtonian solution for  $u^2$  in this higher order term. Writing  $u = u_N + \delta u$  with  $u_N$  given by (310), the differential equation becomes

$$\frac{d^2 \delta u}{d\phi^2} + \delta u = \frac{3GM}{c^2} u_N^2 = \frac{3(GM)^3}{c^2 J^4} (1 + 2\epsilon \cos \phi + \epsilon^2 \cos^2 \phi). \quad (311)$$

In Problem Set 2, you will be asked to solve this equation. The resulting solution for  $u = u_N + \delta u$  may be written

$$u \simeq \frac{GM}{J^2} (1 + \epsilon \cos[\phi(1 - \alpha)]) \quad (312)$$

where  $\alpha = 3(GM/Jc)^2$ . Thus, the perihelion occurs not with a  $\phi$ -period of  $2\pi$ , but with a slightly longer period of

$$\frac{2\pi}{1 - \alpha} \simeq 2\pi + 2\pi\alpha, \quad (313)$$

i.e. an advance of the perihelion by an amount

$$\Delta\phi = 2\pi\alpha = 6\pi \left( \frac{GM}{Jc} \right)^2 = 6\pi \left( \frac{GM}{c^2 L} \right) = 2.783 \times 10^{-6} \left( \frac{10^{10} \text{m}}{L} \right) \quad (314)$$

in units of radians per orbit. With  $L = 5.546 \times 10^{10}$  m, the measured semilatus rectum for Mercury's orbit, this value of  $\Delta\phi$  works out to be precisely 43 seconds of arc per century. (There are 415.2 orbits of Mercury per century.) Einstein confided to a colleague that when he found that his result (314) agreed so precisely with observations, he felt as though something inside him actually snapped...

From the discovery in 1915, until the 1982 gravitational radiation measurement of the binary pulsar 1913+16, the accord with the Mercury perihelion advance was general relativity's greatest observational success.

## 6.8 Shapiro delay: the fourth protocol

For many years, the experimental foundation of general relativity consisted of the three tests we have described that were first proposed by Einstein: the gravitational red shift, the bending of light by gravitational fields, and the advance of Mercury's perihelion. In 1964, nearly a decade after Einstein's passing, a fourth test was proposed: the time delay by radio radar signals when passing near the sun in the inner solar system. The idea, proposed and carried out by Irwin Shapiro, is that a radio signal is sent from earth, bounces off Mercury, and returns. One does the experiment when Mercury is at its closest point to the earth, then repeats the experiment when the planet is on the far side of its orbit. There should be an additional delay of the pulses when Mercury is on the far side of the sun because of the traversal of the radio waves across the sun's Schwarzschild geometry. It is this delay that is measured.

Recall equation (284), using the "ordinary" time parameter  $t$  for an observer at infinity, with  $E = 0$  for radio waves:

$$\left( \frac{dr}{dt} \right)^2 + \frac{B^2 J^2}{A r^2} = \frac{Bc^2}{A} \quad (315)$$

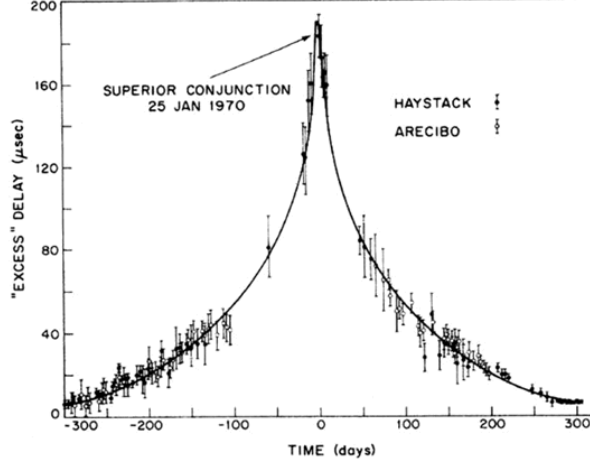


Figure 4: Radar echo delay from Venus as a function of time, fit with general relativistic prediction.

It is convenient to evaluate the constant  $J$  in terms of  $r_0$ , the point of closest approach to the sun. With  $dr/dt = 0$ , we easily find

$$J^2 = \frac{r_0^2 c^2}{B_0} \quad (316)$$

where  $B_0 \equiv B(r_0)$ . The differential equation then separates and we find that the time  $t(r, r_0)$  to traverse from  $r_0$  to  $r$  (or vice-versa) is

$$t(r, r_0) = \frac{1}{c} \int_{r_0}^r \frac{A dr}{\left(1 - \frac{B}{B_0} \frac{r_0^2}{r^2}\right)^{1/2}}, \quad (317)$$

where we have made use of  $AB = 1$ . Expanding to first order in  $GM/c^2 r$  with  $B = 1 - 2GM/c^2 r$ :

$$1 - \frac{B}{B_0} \frac{r_0^2}{r^2} \simeq 1 - \left[1 + \frac{2GM}{c^2} \left(\frac{1}{r_0} - \frac{1}{r}\right)\right] \frac{r_0^2}{r^2}. \quad (318)$$

This may now be rewritten as:

$$1 - \frac{B}{B_0} \frac{r_0^2}{r^2} \simeq \left(1 - \frac{r_0^2}{r^2}\right) \left(1 - \frac{2GM r_0}{c^2 r(r + r_0)}\right) \quad (319)$$

Using this in our time integral for  $t(r_0, r)$  and expanding,

$$t(r_0, r) = \frac{1}{c} \int_{r_0}^r dr \left(1 - \frac{r_0^2}{r^2}\right)^{-1/2} \left(1 + \frac{2GM}{rc^2} + \frac{GM r_0}{c^2 r(r + r_0)}\right) \quad (320)$$

The required integrals are

$$\frac{1}{c} \int_{r_0}^r \frac{r dr}{(r^2 - r_0^2)^{1/2}} = \frac{1}{c} (r^2 - r_0^2)^{1/2} \quad (321)$$

$$\frac{2GM}{c^3} \int_{r_0}^r \frac{dr}{(r^2 - r_0^2)^{1/2}} = \frac{2GM}{c^3} \cosh^{-1} \left( \frac{r}{r_0} \right) = \frac{2GM}{c^3} \ln \left( \frac{r}{r_0} + \sqrt{\frac{r^2}{r_0^2} - 1} \right) \quad (322)$$

$$\frac{GM r_0}{c^3} \int_{r_0}^r \frac{dr}{(r + r_0)(r^2 - r_0^2)^{1/2}} = \frac{GM}{c^3} \sqrt{\frac{r - r_0}{r + r_0}} \quad (323)$$

Thus,

$$t(r, r_0) = \frac{1}{c} (r^2 - r_0^2)^{1/2} + \frac{2GM}{c^3} \ln \left( \frac{r}{r_0} + \sqrt{\frac{r^2}{r_0^2} - 1} \right) + \frac{GM}{c^3} \sqrt{\frac{r - r_0}{r + r_0}} \quad (324)$$

We are interested in  $2t(r_1, r_0) \pm 2t(r_2, r_0)$  for the path from the earth at  $r_1$ , reflected from the planet (at  $r_2$ ), and back. The  $\pm$  sign depends upon whether the signal passes through  $r_0$  while enroute to the planet, i.e. on whether the planet is on the far side or the near side of the sun.

It may seem straightforward to plug in values appropriate to the earth's radial location and the planet's (either Mercury or Venus, in fact), compute the “expected Newtonian time” for transit (a sum of the first terms) and then measure the actual time for comparison with our formula. In practise, to know what the delay is, we have to know what the Newtonian transit time is to fantastic accuracy! In fact, the way this is done is to treat the problem not as a measurement of a single delay time, but as an entire function of time given by our solution (324) with  $r = r(t)$ . Figure (3) shows such a fit near the passage of superior conjunction (i.e. the far side orbital near the sun in sky projection), in excellent agreement with theory. Exactly how the parameterisation is carried out would take us too far afield; there is some discussion in W72 pp. 202–207, and an abundance of topical information on the internet under “Shapiro delay.”

Modern applications of the Shapiro delay use pulsars as signal probes, whose time passage properties are altered by the presence of gravitational waves, a topic for the next chapter.

*They are not objective, and (like absolute velocity) are not detectable by any conceivable experiment. They are merely sinuosities in the co-ordinate system, and the only speed of propagation relevant to them is “the speed of thought.”*

— *A. S. Eddington writing in 1922 of Einstein’s suspicions.*

*On September 14, 2015, at 09:50:45 UTC the two detectors of the Laser Interferometer Gravitational Wave Observatory simultaneously observed a transient gravitational wave signal. The signal sweeps upwards from 35 to 250 Hz with a peak gravitational wave strain of  $1 \times 10^{-21}$ . It matches the waveform predicted by general relativity for the inspiral and merger of a pair of black holes and the ringdown of the resulting single black hole.*

— *B. P. Abbott et al., 2016, Physical Review Letters, 116, 061102*

## 7 Gravitational Radiation

Gravity is spoken in the three languages. First, there is traditional Newtonian potential theory, the language used by most practising astrophysicists. Then, there is the language of Einstein’s General Relativity Theory, the language of Riemannian geometry that we have been studying. Finally, there is the language of quantum field theory: gravity is a theory of the exchange of spin 2 particles, gravitons, much as electromagnetism is a theory arising from the exchange of spin 1 photons. Just as the starting point of quantum electrodynamics is the radiation theory of Maxwell, the starting point of quantum gravity must be a classical radiation theory of gravity. Unlike quantum electrodynamics, the most accurate physical theory ever created, there is no quantum theory of gravity at present, and there is not even a consensus approach. Quantum gravity is therefore very much an active area of ongoing research. For the theorist, this is reason enough to study the theory of gravitational radiation in general relativity. But there are good reasons for the practical astrophysicist to get involved. In February 2016, the first detection of gravitational waves was announced. The event signal had been received and recorded on September 14, 2015, and is denoted G[ravitational]W[ave]150914. The detection was so clean, and matched the wave form predictions of general relativity in such detail, there can be no doubt that the detection was genuine. A new way to probe the most impenetrable parts of the Universe is at hand.

The theory of general relativity in the limit when  $g_{\mu\nu}$  is very close to  $\eta_{\mu\nu}$  is a classical theory of gravitational radiation (and not just Newtonian theory), in the same way that Maxwellian Electrodynamics is a classical electromagnetic radiation theory. The field equations for the small difference tensor  $g_{\mu\nu} - \eta_{\mu\nu}$  become, in the weak field limit, a set of rather ordinary looking wave equations with source terms—much like Maxwell’s Equations. The principal difference is that electrodynamics is sourced by a vector quantity (the usual vector potential  $\mathbf{A}$  with the potential  $\Phi$  combine to form a 4-vector), whereas gravitational fields in general relativity are sourced by a tensor quantity  $T_{\mu\nu}$ . This becomes a major difference when we relax the condition that the gravity field be weak: the gravitational radiation itself makes a contribution to its own source, something electromagnetic radiation cannot do. But this is not completely unprecedented in wave theories. We have seen this sort of thing before, in a purely classical context: sound waves can themselves generate acoustical disturbances, and one of the consequences is a shock wave, or sonic boom. While a few somewhat pathological mathematical solutions for exact gravitational radiation waves are known, in general people either work in the weak field limit or resort to numerical solutions of the field equations. Even with powerful computers, however, precise numerical solutions of the field equations for astrophysically interesting problems—like merging black holes—have long been a major technical challenge. In the last decade, a practical mathematical breakthrough has occurred, and it is now possible to compute highly accurate wave forms for these kinds of problems, with important predictions for the new generation of gravitational wave detectors.

As we have noted, astrophysicists now have perhaps the most important reason of all to understand gravitational radiation: we are on the verge of what will surely be a golden age of gravitational wave astronomy. That gravitational radiation truly exists was established in the years following 1974, when a close binary system (7.75 hour period) with a neutron star and a pulsar (PSR 1913+16) was discovered and followed-up by Hulse and Taylor. So much orbital information could be extracted from this remarkable system that it was possible to predict, then measure, the rate of orbital decay (the gradual shortening of the period of the pulsar’s decaying orbit) caused by the energy loss in gravitational radiation. Though tiny in any practical sense, the period change was large enough to be cleanly measured. General relativity turned out to be exactly correct (Taylor & Weisberg, ApJ, 1982, 253, 908), and the 1993 Nobel Prize in Physics was duly awarded to Hulse and Taylor for this historical achievement.

The September 2015 gravitational wave detection pushed back the envelope dramatically. It established that i) the direct reception and quantitative analysis of gravitational waves is technically feasible and will soon become a widely-used probe of the universe; ii) black holes exist beyond any doubt whatsoever, this truly is the proverbial “smoking-gun”; iii) the full dynamical content of strong field general relativity, on time and length scales characteristic of stellar systems, is correct. This achievement is an historical milestone in physics. Some have speculated that its impact on astronomy will rival Galileo’s introduction of the telescope. Perhaps Hertz’s 1887 detection of electromagnetic radiation in the lab is another, more apt, comparison. (Commercial exploitation of gravity waves is probably some ways off. Maybe it will be licenced someday as a revenue source.)

There may be more to come. In the near future, it is anticipated that extremely delicate pulsar timing experiments, in which arrival times of pulses are measured to fantastic precision, will detect distortions in space. In essence, this is a measure of the Shapiro delay, caused neither by the Sun nor by a star, but by the passage of a large scale gravitational wave between us and the pulsar probes!

The subject of gravitational radiation is complicated and computationally intensive. Even the basics will involve some effort on your part. I hope you will agree that the effort is well rewarded.

## 7.1 The linearised gravitational wave equation

*Summary: The linearised gravitational wave equation of general relativity takes on the form of a standard wave equation when written in what are known as “harmonic” coordinates. These coordinates are very important for the study of gravitational radiation.*

We assume that the metric is close to Minkowski space. Let us introduce the quantity  $h_{\mu\nu}$ , the (small) departure in the metric tensor  $g_{\mu\nu}$  from its Minkowski  $\eta_{\mu\nu}$  limit:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad (325)$$

To leading order, when we raise and lower indices we may do so with  $\eta_{\mu\nu}$ . But be careful with  $g^{\mu\nu}$  itself. Don't just lower the indices in the above equation willy-nilly! Instead, note that

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} \quad (326)$$

to ensure  $g_{\mu\nu}g^{\nu\kappa} = \delta_{\mu}^{\kappa}$ . (You can raise the index of  $g$  with  $\eta$  only when approximating  $g^{\mu\nu}$  as its leading order value, which is  $\eta^{\mu\nu}$ .) Note that

$$\eta^{\mu\nu}h_{\nu\kappa} = h^{\mu}_{\kappa}, \quad \eta^{\mu\nu}\frac{\partial}{\partial x^{\nu}} = \frac{\partial}{\partial x_{\mu}}, \quad (327)$$

so that we can slide dummy indices “up-down” as follows:

$$\frac{\partial h_{\mu\nu}}{\partial x_{\mu}} = \eta_{\mu\rho}\frac{\partial h^{\rho}_{\nu}}{\partial x_{\mu}} = \frac{\partial h^{\rho}_{\nu}}{\partial x^{\rho}} \equiv \frac{\partial h^{\mu}_{\nu}}{\partial x^{\mu}} \quad (328)$$

The story begins with the Einstein Field Equations cast in a form in which the “linearised Ricci tensor” is isolated on the left side of our working equation. Specifically, we write

$$R_{\mu\nu} = R_{\mu\nu}^{(1)} + R_{\mu\nu}^{(2)} + \dots \text{etc.} \quad (329)$$

and

$$G_{\mu\nu}^{(1)} = R_{\mu\nu}^{(1)} - \eta_{\mu\nu}\frac{R^{(1)}}{2} \quad (330)$$

where  $R_{\mu\nu}^{(1)}$  is all the Ricci tensor terms linear in  $h_{\mu\nu}$ ,  $R_{\mu\nu}^{(2)}$  all terms quadratic in  $h_{\mu\nu}$ , and so forth. The linearised affine connection is

$$\Gamma_{\mu\nu}^{\lambda} = \frac{1}{2}\eta^{\lambda\rho}\left(\frac{\partial h_{\rho\nu}}{\partial x^{\mu}} + \frac{\partial h_{\rho\mu}}{\partial x^{\nu}} - \frac{\partial h_{\mu\nu}}{\partial x^{\rho}}\right) = \frac{1}{2}\left(\frac{\partial h^{\lambda}_{\nu}}{\partial x^{\mu}} + \frac{\partial h^{\lambda}_{\mu}}{\partial x^{\nu}} - \frac{\partial h_{\mu\nu}}{\partial x_{\lambda}}\right). \quad (331)$$

In terms of  $h_{\mu\nu}$  and  $h = h^{\mu}_{\mu}$ , from equation (219) on page 51, we explicitly find

$$R_{\mu\nu}^{(1)} = \frac{1}{2}\left(\frac{\partial^2 h}{\partial x^{\mu}\partial x^{\nu}} - \frac{\partial^2 h^{\lambda}_{\mu}}{\partial x^{\nu}\partial x^{\lambda}} - \frac{\partial^2 h^{\lambda}_{\nu}}{\partial x^{\mu}\partial x^{\lambda}} + \square h_{\mu\nu}\right) \quad (332)$$

where

$$\square \equiv \frac{\partial^2}{\partial x^{\lambda}\partial x_{\lambda}} = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \quad (333)$$



is the d'Alembertian (clearly a Lorentz invariant), making a most welcome appearance into the proceedings. Contracting  $\mu$  with  $\nu$ , we find that

$$R^{(1)} = \square h - \frac{\partial^2 h^{\mu\nu}}{\partial x^\mu \partial x^\nu} \quad (334)$$

where we have made use of

$$\frac{\partial h_\mu^\lambda}{\partial x_\mu} = \frac{\partial h^{\lambda\mu}}{\partial x^\mu}.$$

Assembling  $G_{\mu\nu}^{(1)}$ , we find

$$G_{\mu\nu}^{(1)} = \frac{1}{2} \left[ \frac{\partial^2 h}{\partial x^\mu \partial x^\nu} - \frac{\partial^2 h_\mu^\lambda}{\partial x^\lambda \partial x^\nu} - \frac{\partial^2 h_\nu^\lambda}{\partial x^\lambda \partial x^\mu} + \square h_{\mu\nu} - \eta_{\mu\nu} \left( \square h - \frac{\partial^2 h^{\lambda\rho}}{\partial x^\lambda \partial x^\rho} \right) \right]. \quad (335)$$

The full, nonlinear Field Equations may then formally be written

$$G_{\mu\nu}^{(1)} = - \left( \frac{8\pi G T_{\mu\nu}}{c^4} + G_{\mu\nu} - G_{\mu\nu}^{(1)} \right) \equiv - \frac{8\pi G (T_{\mu\nu} + \tau_{\mu\nu})}{c^4}, \quad (336)$$

where

$$\tau_{\mu\nu} = \frac{c^4}{8\pi G} (G_{\mu\nu} - G_{\mu\nu}^{(1)}) \simeq \frac{c^4}{8\pi G} \left( R_{\mu\nu}^{(2)} - \eta_{\mu\nu} \frac{R^{(2)}}{2} \right) \quad (337)$$

Though composed of geometrical terms, the quantity  $\tau_{\mu\nu}$  is written on the right side of the equation with the stress energy tensor  $T_{\mu\nu}$ , and is interpreted as the *stress energy contribution of the gravitational radiation itself*. We shall have more to say on this in section 7.4. In linear theory,  $\tau_{\mu\nu}$  is neglected in comparison with the ordinary matter  $T_{\mu\nu}$ .

*Exercise.* Show that, in terms of the source function  $S_{\mu\nu} = T_{\mu\nu} - \eta_{\mu\nu} T/2$ , the linear field equation is

$$\frac{\partial^2 h}{\partial x^\mu \partial x^\nu} - \frac{\partial^2 h_\mu^\lambda}{\partial x^\lambda \partial x^\nu} - \frac{\partial^2 h_\nu^\lambda}{\partial x^\lambda \partial x^\mu} + \square h_{\mu\nu} = - \frac{16\pi G}{c^4} S_{\mu\nu}$$

Recover the static Poisson equation limit, as per our more general treatment in Chapter 6.

This is all a bit disappointing to behold. Even the linearised Field Equations look to be a mess! But then, you may have forgotten that the raw Maxwell wave equations for the potentials are no present, either. You will permit me to remind you. Here are the equations for the scalar potential  $\Phi$  and vector potential  $\mathbf{A}$ :

$$\nabla^2 \Phi + \frac{1}{c} \frac{\partial}{\partial t} (\nabla \cdot \mathbf{A}) = -4\pi \rho \quad (338)$$

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla \left( \nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \Phi}{\partial t} \right) = -\frac{4\pi}{c} \mathbf{J} \quad (339)$$

(Note: I have used esu units, which are much more natural for relativity. Here  $\rho$  is the electric charge density.) Do the following exercise!

*Exercise.* In covariant notation, with  $A^\alpha = (\Phi, \mathbf{A})$  and  $J^\alpha = (\rho, \mathbf{J}/c)$  representing respectively the potential and source term 4-vectors, the original general equations look a bit more presentable. The only contravariant 4-vectors that we can form which are second order in

the derivatives of  $A^\alpha$  are  $\square A^\alpha$  and  $\partial^\alpha \partial_\beta A^\beta$ . Show that if  $\partial_\alpha J^\alpha = 0$  identically, then our equation relating  $A^\alpha$  to  $J^\alpha$  must be of the form

$$\square A^\alpha - \partial^\alpha \partial_\beta A^\beta = C J^\alpha$$

where  $C$  is a constant to be determined, and that this equation remains unchanged when the transformation  $A^\alpha \rightarrow A^\alpha + \partial^\alpha \Lambda$  is made. This property is known as gauge-invariance. We will shortly see something very analogous in general relativity. In the meantime, how do we determine  $C$ ?

Do you remember the way forward from here? Work in the ‘‘Lorenz gauge,’’ which we are always free to do:

$$\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \Phi}{\partial t} = 0 \quad (340)$$

In covariant 4-vector language, this is simply a vanishing divergence condition,  $\partial_\alpha A^\alpha = 0$ . Then, the dynamical equations simplify:

$$\nabla^2 \Phi - \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2} = \square \Phi = -4\pi \rho \quad (341)$$

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = \square \mathbf{A} = -\frac{4\pi}{c} \mathbf{J} \quad (342)$$

This is much nicer. Physically transparent Lorentz-invariant wave equations emerge. Might something similar happen for the Einstein Field Equations?

That the answer might be YES is supported by noticing that  $G_{\mu\nu}^{(1)}$  can be written entirely in terms of the Bianchi-like quantity

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{\eta_{\mu\nu} h}{2}, \quad \text{or} \quad \bar{h}_\nu^\mu = h_\nu^\mu - \frac{\delta_\nu^\mu h}{2}. \quad (343)$$

Using this in (335), the *linearised* Field Equation becomes

$$2G_{\mu\nu}^{(1)} = \square \bar{h}_{\mu\nu} - \frac{\partial^2 \bar{h}_\mu^\lambda}{\partial x^\nu \partial x^\lambda} - \frac{\partial^2 \bar{h}_\nu^\lambda}{\partial x^\mu \partial x^\lambda} + \eta_{\mu\nu} \frac{\partial^2 \bar{h}^{\lambda\rho}}{\partial x^\lambda \partial x^\rho} = -\frac{16\pi G T_{\mu\nu}}{c^4}. \quad (344)$$

(It is easiest to verify this by starting with (344), substituting with (343), and showing that this leads to (335).)

Interesting. Except for  $\square \bar{h}_{\mu\nu}$ , every term in this equation involves the divergence  $\partial_\mu \bar{h}^\mu_\nu$  or  $\partial_\mu \bar{h}^{\mu\nu}$ . Hmmm. Shades of Maxwell’s  $\partial_\alpha A^\alpha$ . In the Maxwell case, the freedom of gauge invariance allowed us to pick the gauge in which  $\partial_\alpha A^\alpha = 0$ . Does equation (344) have a gauge invariance that will allow us to do the same for gravitational radiation, so that we can set these  $\bar{h}$ -divergences to zero?

It does. Go back to equation (335) and on the right side, change  $h_{\mu\nu}$  to  $h'_{\mu\nu}$ , where

$$h'_{\mu\nu} = h_{\mu\nu} - \frac{\partial \xi_\nu}{\partial x^\mu} - \frac{\partial \xi_\mu}{\partial x^\nu}, \quad h' = h - 2 \frac{\partial \xi^\mu}{\partial x^\mu} \quad (345)$$

and the  $\xi_\mu$  represent any vector function. You will find that the form of the equation is completely unchanged, i.e., the  $\xi_\mu$  terms cancel out identically! This is a true gauge invariance.

*Exercise.* Show that under a gauge transformation,

$$\bar{h}'_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{\partial \xi_\nu}{\partial x^\mu} - \frac{\partial \xi_\mu}{\partial x^\nu} + \eta_{\mu\nu} \frac{\partial \xi^\rho}{\partial x^\rho}, \quad \frac{\partial \bar{h}'^\mu_\nu}{\partial x^\mu} = \frac{\partial \bar{h}^\mu_\nu}{\partial x^\mu} - \square \xi_\nu$$

What is happening here is that an infinitesimal coordinate transformation itself is *acting to leading order* as a gauge transformation. If

$$x'^\mu = x^\mu + \xi^\mu(x), \quad \text{or} \quad x^\mu = x'^\mu - \xi^\mu(x') \quad (\text{to leading order}), \quad (346)$$

then

$$g'_{\mu\nu} = \eta'_{\mu\nu} + h'_{\mu\nu} = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma} = \left( \delta^\rho_\mu - \frac{\partial \xi^\rho}{\partial x^\mu} \right) \left( \delta^\sigma_\nu - \frac{\partial \xi^\sigma}{\partial x^\nu} \right) (\eta_{\rho\sigma} + h_{\rho\sigma}) \quad (347)$$

Notice that for the derivatives of  $\xi^\rho$ , we may use  $x$  instead of  $x'$ , since the difference is quadratic in  $\xi$ . With  $\eta'$  identical to  $\eta$ , we must have to leading order in  $\xi$

$$h'_{\mu\nu} = h_{\mu\nu} - \frac{\partial \xi_\nu}{\partial x^\mu} - \frac{\partial \xi_\mu}{\partial x^\nu}, \quad (348)$$

as before. Though closely related, don't confuse *general covariance* under coordinate transformations with this gauge transformation. Unlike general covariance, the gauge transformation is implemented without actually changing the coordinates! We keep the same  $x$ 's and add a group of certain functional derivatives to the  $h_{\mu\nu}$ , analogous to adding a gradient  $\nabla\Phi$  to  $\mathbf{A}$  in Maxwell's theory. We find that the equations remain identical, just as we would find if we took  $\nabla \times (\mathbf{A} + \nabla\Lambda)$  in the Maxwell case.

Pause for a moment. In general relativity, don't we actually need to change the coordinates when we...well, when we change the coordinates? What is going on here? Keeping the coordinates is not an option, is it? Change the  $h^{\mu\nu}$  tensor components as though we were changing coordinates, but then leave the coordinates untouched? Why should that work? The answer is that the additional terms that we pick up when we elect to do the full coordinate transformation are of *higher order* than the purely linear (in  $\partial\xi$ ) terms that come from merely changing the *components* of  $\bar{h}_{\mu\nu}$ . Remember that  $\bar{h}_{\mu\nu}$  is itself already infinitesimal! The additional terms that we are ignoring are of quadratic order, the product of  $\partial\xi$  and  $\partial h$ , and all  $h$ 's are small. The terms that we retain are of linear order,  $\partial\xi$  alone. This is the right and proper thing to do. We are, after all, only working to linear order in our formulation of the problem.

Understanding the gauge invariant properties of the gravitational wave equation was very challenging in the early days of the subject. The opening "speed-of-thought" quotation of this chapter by Eddington is taken somewhat out of context. What he really said in his famous paper (Eddington A.S. 1922 *Proc. Roy. Soc. A*, **102**, 716, 268) is the following:

"Weyl has classified plane gravitational waves into three types, viz.: (1) longitudinal-longitudinal; (2) longitudinal-transverse; (3) transverse-transverse. The present investigation leads to the conclusion that transverse-transverse waves are propagated with the speed of light *in all systems of co-ordinates*. Waves of the first and second types have no fixed velocity—a result which rouses suspicion as to their objective existence. Einstein had also become suspicious of these waves (in so far as they occur in his special co-ordinate system) for another reason, because he found they convey no energy. They are not objective and (like absolute velocity) are not detectable by any conceivable experiment. They are merely sinuosities in the co-ordinate system, and the only speed of propagation relevant to them is the 'speed of thought.' "

The final line of this quotation is often taken to be dismissive of the entire notion of gravitational radiation, which it clearly is not. Rather, it is directed toward those solutions which we would now say are gauge-dependent (either of the first two types of waves, which involve at least one longitudinal component) and those which are gauge-independent (the third category, which is completely transverse). Physical solutions must ultimately be gauge independent. Matters would have been quite clear to anyone who bothered to examine the components of the Riemann curvature tensor! The first two types of waves would have produced an identically zero  $R^\lambda_{\mu\nu\kappa}$ . They produce no curvature; they are indeed “merely sinuositities in the co-ordinate system,” and they are unphysical.

Back to our problem. Just as the Lorenz gauge  $\partial_\alpha A^\alpha = 0$  was useful in the case of Maxwell’s equations, so now is the so-called *harmonic gauge*:

$$\frac{\partial \bar{h}^\mu_\nu}{\partial x^\mu} = \frac{\partial h^\mu_\nu}{\partial x^\mu} - \frac{1}{2} \frac{\partial h}{\partial x^\nu} = 0 \quad (349)$$

In this gauge, the Field Equations (344) take the “wave-equation” form

$$\boxed{\square \bar{h}_{\mu\nu} = -\frac{16\pi G T_{\mu\nu}}{c^4}} \quad (350)$$

How can we be sure that, even with our gauge freedom, we can find the right  $\xi^\mu$  to get into a harmonic gauge and ensure the emergence of (350)? Well, if we have been unfortunate enough to be working in a gauge in which equation (349) is not satisfied, then form  $h'_{\mu\nu}$  à la equation (348) and demand that in our new gauge,  $\partial h'^\mu_\nu / \partial x^\mu = (1/2) \partial h' / \partial x^\nu$ . We find that this implies

$$\square \xi_\nu = \frac{\partial \bar{h}^\mu_\nu}{\partial x^\mu}, \quad (351)$$

a wave equation for  $\xi_\nu$  identical in form to (350). For this equation, a solution certainly exists. Indeed, our experience with electrodynamics has taught us that the solution to the fundamental radiation equation (350) takes the form

$$\bar{h}_{\mu\nu}(\mathbf{r}, t) = \frac{4G}{c^4} \int \frac{T_{\mu\nu}(\mathbf{r}', t - R/c)}{R} d^3r', \quad R \equiv |\mathbf{r} - \mathbf{r}'| \quad (352)$$

and hence a similar solution exists for (351). The  $\bar{h}_{\mu\nu}$ , like their electrodynamic counterparts, are determined at time  $t$  and location  $\mathbf{r}$  by a source intergration over  $\mathbf{r}'$  taken at the retarded times  $t' \equiv t - R/c$ . In other words, disturbances in the gravitational field travel at a finite speed, the speed of light  $c$ .

*Exercise.* Show that for a source with motions near the speed of light, like merging black holes,  $\bar{h}_{\mu\nu}$  (or  $h_{\mu\nu}$  for that matter) is of order  $R_S/r$ , where  $R_S$  is the Schwarzschild radius based on the total mass of the system in question and  $r$  is the distance to the source. You want to know how big  $h_{\mu\nu}$  is going to be in your detector when black holes merge? Count the number of expected Schwarzschild radii to the source and take the reciprocal. With  $M_\odot^{tot}$  equal to the total mass measured in solar masses, show that  $h_{\mu\nu} \sim 3M_\odot^{tot}/r_{km}$ , measuring  $r$  in km. We are pushing our weak field approximation here, but to this order it works fine. We’ll give a sharper estimate shortly.

### 7.1.1 Come to think of it...

You may not have actually seen the solution (352) before, or maybe, you know, you just need a little reminding. It is important. Let’s derive it.

Consider the equation

$$-\frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} + \nabla^2 \Psi = -4\pi f(\mathbf{r}, t) \quad (353)$$

We specialise to the Green's function solution

$$-\frac{1}{c^2} \frac{\partial^2 G}{\partial t^2} + \nabla^2 G = -4\pi \delta(\mathbf{r}) \delta(t) \quad (354)$$

Of course, our particular choice of origin is immaterial, as is our zero of time, so that we shall replace  $\mathbf{r}$  by  $\mathbf{R} \equiv \mathbf{r} - \mathbf{r}'$  ( $R \equiv |\mathbf{R}|$ ), and  $t$  by  $\tau \equiv t - t'$  at the end of the calculation, with the primed values being fiducial reference points. The form of the solution we find here will still be valid with the shifts of space and time origins.

Fourier transform (354) by integrating over  $\int e^{i\omega t} dt$  and denote the fourier transform of  $G$  by  $\tilde{G}$ :

$$k^2 \tilde{G} + \nabla^2 \tilde{G} = -4\pi \delta(\mathbf{r}) \quad (355)$$

where  $k^2 = \omega^2/c^2$ . Clearly  $\tilde{G}$  is a function only of  $r$ , hence the solution to the homogeneous equation away from the origin,

$$\frac{d^2(r\tilde{G})}{dr^2} + k^2(r\tilde{G}) = 0,$$

is easily found to be  $\tilde{G} = e^{\pm ikr}/r$ . Although we haven't done anything to deserve this favour, the singular delta function behaviour is actually *already* included here! This may be seen by taking the static limit  $k \rightarrow 0$ , in which case we recover the correctly normalised  $1/r$  potential of a point charge at the origin. The back transform gives

$$G = \frac{1}{2\pi r} \int_{-\infty}^{\infty} e^{\pm ikr - i\omega t} d\omega = \frac{1}{2\pi r} \int_{-\infty}^{\infty} e^{-i\omega(t \mp r/c)} d\omega \quad (356)$$

which we recognise as a Dirac delta function (remember  $\omega/k = c$ ):

$$G = \frac{\delta(t \mp r/c)}{r} \rightarrow \frac{\delta(t - r/c)}{r} \rightarrow \frac{\delta(\tau - R/c)}{R}. \quad (357)$$

We have selected the *retarded* time solution  $t - r/c$  as a nod to causality, and moved thence to  $(\tau, R)$  variables for an arbitrary time and space origin. We see directly that a flash at  $t = t'$ , located at  $\mathbf{r} = \mathbf{r}'$ , produces an effect at a time  $R/c$  later, at a distance  $R$  from the flash. The general solution constructed from our Green's function is

$$\Psi(\mathbf{r}, t) = \int \frac{f(\mathbf{r}', t')}{R} \delta(t - t' - R/c) dt' d\mathbf{r}' = \int \frac{f(\mathbf{r}', t')}{R} d\mathbf{r}' \quad (358)$$

where in the final integral we have set  $t' = t - R/c$ , the retarded time. Remember that  $t'$  depends on both  $\mathbf{r}$  and  $\mathbf{r}'$ .

## 7.2 Plane waves

*Summary: Linear gravitational radiation comes in only two modes of plane wave polarisation. In what are known as "transverse-traceless" (TT) coordinates, a plane wave travelling in the*

$z$  direction has one mode with  $h_{xx}$  and  $h_{yy} = -h_{xx}$  and all others zero. The other mode has  $h_{xy} = h_{yx}$  and all others zero. In contrast to harmonic coordinates, which are generally available, the TT gauge is available only for plane wave solutions. These coordinates are the most widely used for studying linear gravitational radiation.

To understand more fully the solution (352), consider the problem in which  $T_{\mu\nu}$  has an oscillatory time dependence,  $e^{-i\omega t'}$ . Since we are dealing with a linear theory, this isn't particularly restrictive, since any well-behaved time dependence can be represented by a Fourier sum. The source, say a binary star system, occupies a finite volume. We seek the solution for  $\bar{h}_{\mu\nu}$  at distances huge compared with the scale of the source itself, i.e.  $r \gg r'$ . Then,

$$R \simeq r - \mathbf{e}_r \cdot \mathbf{r}' \quad (359)$$

where  $\mathbf{e}_r$  is a unit vector in the  $\mathbf{r}$  direction, and

$$\bar{h}_{\mu\nu}(\mathbf{r}, t) \simeq \frac{e^{i(kr - \omega t)}}{r} \frac{4G}{c^4} \int T_{\mu\nu}(\mathbf{r}') \exp(-i\mathbf{k} \cdot \mathbf{r}') d^3r' \quad (360)$$

with  $\mathbf{k} = (\omega/c)\mathbf{e}_r$  the wavenumber in the radial direction. Since  $r$  is huge, this has the asymptotic form of a plane wave. Hence,  $\bar{h}_{\mu\nu}$  and thus  $h_{\mu\nu}$  itself have the form of simple plane waves, travelling in the radial direction, at large distances from the source generating them. These waves turn out to have some remarkable polarisation properties, which we now discuss.

### 7.2.1 The transverse-traceless (TT) gauge

Consider a traveling plane wave for  $h_{\mu\nu}$ , orienting our  $z$  axis along  $\mathbf{k}$ , so that

$$k^0 = \omega/c, \quad k^1 = 0, \quad k^2 = 0, \quad k^3 = \omega/c \quad \text{and} \quad k_0 = -\omega/c, \quad k_i = k^i \quad (361)$$

where as usual we raise and lower indices with  $\eta_{\mu\nu}$ , or its numerically identical dual  $\eta^{\mu\nu}$ .

Then  $h_{\mu\nu}$  takes the form

$$h_{\mu\nu} = e_{\mu\nu} a \exp(ik_\rho x^\rho) \quad (362)$$

where  $a$  is an amplitude and  $e_{\mu\nu} = e_{\nu\mu}$  a polarisation tensor, again with the  $\eta$ 's raising and lowering subscripts. Thus

$$e_{ij} = e_j^i = e^{ij} \quad (363)$$

$$e^{0i} = -e_i^0 = e_i^0 = -e_{0i} \quad (364)$$

$$e^{00} = e_{00} = -e_0^0 \quad (365)$$

The harmonic constraint

$$\frac{\partial h_\nu^\mu}{\partial x^\mu} = \frac{1}{2} \frac{\partial h_\mu^\mu}{\partial x^\nu} \quad (366)$$

implies

$$k_\mu e_\nu^\mu = k_\nu e_\mu^\mu / 2. \quad (367)$$

This leads to several linear dependencies between the  $e_{\mu\nu}$ . For example, when  $\nu = 0$  this means

$$k_0 e_0^0 + k_3 e_0^3 = k_0 (e_i^i + e_0^0) / 2, \quad (368)$$

or

$$-(e_{00} + e_{30}) = (e_{ii} - e_{00}) / 2. \quad (369)$$

When  $\nu = j$  (a spatial index),

$$k_0 e_j^0 + k_3 e_j^3 = k_j (e_{ii} - e_{00})/2. \quad (370)$$

The  $j = 1$  and  $j = 2$  cases reduce to

$$e_{01} + e_{31} = e_{02} + e_{32} = 0, \quad (371)$$

while  $j = 3$  yields

$$e_{03} + e_{33} = (e_{ii} - e_{00})/2 = -(e_{00} + e_{03}). \quad (372)$$

Equations (371) and the first=last equality of (372) yield

$$e_{01} = -e_{31}, \quad e_{02} = -e_{32}, \quad e_{03} = -(e_{00} + e_{33})/2. \quad (373)$$

Using the above expression for  $e_{03}$  in the first=second equality of (372) then gives

$$e_{22} = -e_{11}. \quad (374)$$

Of the 10 independent components of the symmetric  $e_{\mu\nu}$  the harmonic condition (366) thus enables us to express  $e_{0i}$  and  $e_{22}$  in terms of  $e_{3i}$ ,  $e_{00}$ , and  $e_{11}$ . These latter 5 components plus a sixth,  $e_{12}$ , remain unconstrained for the moment.

But wait! We have not yet used the gauge freedom of equation (348) *within* the harmonic constraint. We can still continue to eliminate components of  $e_{\mu\nu}$ . In particular, let us choose

$$\xi_\mu(x) = i\epsilon_\mu \exp(ik_\rho x^\rho) \quad (375)$$

where the  $\epsilon_\mu$  are four constants to be chosen. This satisfies  $\square \xi_\mu = 0$ , and therefore does not change the harmonic coordinate condition,  $\partial_\mu \bar{h}_\nu^\mu = 0$ . Then, following the prescription of (348), we generate a new, but physically equivalent polarisation tensor,

$$e'_{\mu\nu} = e_{\mu\nu} + k_\mu \epsilon_\nu + k_\nu \epsilon_\mu. \quad (376)$$

Now, by choosing the  $\epsilon_\mu$  appropriately, we can actually eliminate *all* of the  $e'_{\mu\nu}$  except for  $e'_{11}$ ,  $e'_{22} = -e'_{11}$ , and  $e'_{12}$ . In particular, using (376),

$$e'_{11} = e_{11}, \quad e'_{12} = e_{12} \quad (377)$$

both remain *unchanged*. But, with  $k = \omega/c$ ,

$$e'_{13} = e_{13} + k\epsilon_1, \quad e'_{23} = e_{23} + k\epsilon_2, \quad e'_{33} = e_{33} + 2k\epsilon_3, \quad e'_{00} = e_{00} - 2k\epsilon_0, \quad (378)$$

so that these four components may be set to zero by a simple choice of the  $\epsilon_\mu$ . When working with plane waves we may always choose this gauge, which is *transverse* (since the only  $e_{ij}$  components that are present are transverse to the  $z$  direction of propagation) and traceless (since  $e_{11} = -e_{22}$ ). Oddly enough, this gauge is named the transverse-traceless (TT) gauge. *Notice that in the TT gauge,  $h_{\mu\nu}$  vanishes if any of its indices are 0, whether raised or lowered, and gravitational waves propagate to leading order only as distortions in space, not time.*

We have thus seen that there are only two independent modes of a gravitational wave. The first, with  $e_{11}$  and  $e_{22} = -e_{11}$  present, is known as the “+” mode. The second, with  $e_{12} = e_{21}$  present, is the cross or “ $\times$ ” mode. All gravitational radiation is a superposition of these two modes.

### 7.3 The quadrupole formula

*Summary:*  $\bar{h}^{ij}$  is given by a very simple formula in harmonic coordinates (not just the TT gauge), and is directly proportional to the second time derivative of the classical moment of inertia tensor,  $I^{ij}$ . The latter requires use of only the  $T^{00}$  component of the stress tensor.

In the limit of large  $r$  (“compact source approximation”), equation (352) is:

$$\bar{h}^{\mu\nu}(\mathbf{r}, t) = \frac{4G}{rc^4} \int T^{\mu\nu}(\mathbf{r}', t') d^3r', \quad (379)$$

where  $t' = t - r/c$  is the retarded time. Moreover, for the TT gauge, we are interested in the spatial  $ij$  components of this equation, since all time indices vanish. (Also, because  $\bar{h}_{\mu\nu}$  is traceless, we need not distinguish between  $h$  and  $\bar{h}$ .) The integral over  $T_{ij}$  may be cast in a very convenient form as follows.

$$0 = \int \frac{\partial(x'^j T^{ik})}{\partial x'^k} d^3r' = \int \left( \frac{\partial T^{ik}}{\partial x'^k} \right) x'^j d^3r' + \int T^{ij} d^3r', \quad (380)$$

where the first equality follows because the first integral reduces to a surface integration of  $T^{ik}$  at infinity, where it is presumed to vanish. Thus

$$\int T^{ij} d^3r' = - \int \left( \frac{\partial T^{ik}}{\partial x'^k} \right) x'^j d^3r' = \int \left( \frac{\partial T^{i0}}{\partial x'^0} \right) x'^j d^3r' = \frac{1}{c} \frac{d}{dt'} \int T^{i0} x'^j d^3r' \quad (381)$$

where the second equality uses the conservation of  $T^{\mu\nu}$ . Remember that  $t'$  is the retarded time. As  $T_{ij}$  is symmetric in its indices,

$$\frac{d}{dt'} \int T^{i0} x'^j d^3r' = \frac{d}{dt'} \int T^{j0} x'^i d^3r' \quad (382)$$

Continuing in this same spirit,

$$0 = \int \frac{\partial(T^{0k} x'^i x'^j)}{\partial x'^k} d^3r' = \int \left( \frac{\partial T^{0k}}{\partial x'^k} \right) x'^i x'^j d^3r' + \int (T^{0i} x'^j + T^{0j} x'^i) d^3r' \quad (383)$$

Using exactly the same reasoning as before,

$$\int (T^{0i} x'^j + T^{0j} x'^i) d^3r' = \frac{1}{c} \frac{d}{dt'} \int T^{00} x'^i x'^j d^3r' \quad (384)$$

Differentiating with respect to  $t'$ , and using (382) and (381) gives an elegant result:

$$\int T^{ij} d^3r' = \frac{1}{2c^2} \frac{d^2}{dt'^2} \int T^{00} x'^i x'^j d^3r'. \quad (385)$$

Inserting this in (379), we obtain the *quadrupole formula* for gravitational radiation:

$$\boxed{\bar{h}^{ij} = \frac{2G}{c^6 r} \frac{d^2 I^{ij}}{dt'^2}} \quad (386)$$



where  $I^{ij}$  is the *quadrupole-moment tensor* of the energy density:

$$I^{ij} = \int T^{00} x'^i x'^j d^3 r' \quad (387)$$

To estimate this numerically, we write

$$\frac{d^2 I^{ij}}{dt'^2} \sim M a^2 c^2 \omega^2 \quad (388)$$

where  $M$  is the characteristic mass of the rotating system,  $a$  an internal separation, and  $\omega$  a characteristic frequency, an orbital frequency for a binary say. Then

$$\bar{h}^{ij} \sim \frac{2GMa^2\omega^2}{c^4 r} \simeq 7 \times 10^{-22} (M/M_\odot) (a_{11}^2 \omega_7^2 / r_{100}) \quad (389)$$

where  $M/M_\odot$  is the mass in solar masses,  $a_{11}$  the separation in units of  $10^{11}$  cm (about a separation of one solar radius),  $\omega_7$  the frequency associated with a 7 hour orbital period (similar to PSR193+16) and  $r_{100}$  the distance in units of 100 parsecs, some  $3 \times 10^{20}$  cm. A typical rather large  $h$  one might expect at earth from a local astronomical source is then of order  $10^{-21}$ .

What about the LIGO source, GW150914? How does our formula work in this case? The distance in this case is cosmological, not local, with  $r = 1.2 \times 10^{22}$  km, or in astronomical parlance, about 400 megaparsecs (Mpc). In this case, we write (389) as

$$\bar{h}^{ij} \sim \frac{2GMa^2\omega^2}{c^4 r} = \left( \frac{2.9532}{r_{km}} \right) \left( \frac{M}{M_\odot} \right) \left( \frac{a\omega}{c} \right)^2 \simeq 1 \times 10^{-22} \frac{M/M_\odot}{r_{Gpc}} \left( \frac{a\omega}{c} \right)^2, \quad (390)$$

since  $2GM_\odot/c^2$  is just the Sun's Schwarzschild radius. (One Gpc =  $10^3$  Mpc =  $3.0856 \times 10^{22}$  km.) The point is that  $(a\omega/c)^2$  is a number not very different from 1 for a relativistic source, perhaps 0.1 or so. Plugging in numbers with  $M/M_\odot = 60$  and  $(a\omega/c)^2 = 0.1$ , we find  $\bar{h}_{ij} = 1.5 \times 10^{-21}$ , just about as observed at peak amplitude.

*Exercise.* Prove that  $\bar{h}^{ij}$  given by (386) is an *exact* solution of  $\square \bar{h}^{ij} = 0$ , for any  $r$ , even if  $r$  is not large.

## 7.4 Radiated Energy

*Overall Section Summary:* The energy flux at infinity carried off by gravitational radiation is given in the  $TT$  gauge is given by  $-(c^4/32\pi G) \partial_t \bar{h}^{\mu\nu} \partial_i \bar{h}_{\mu\nu}$ . This is needed to understand how the orbits of binary systems evolve due to the emission of gravitational radiation.

### 7.4.1 A useful toy problem

We have yet to make the link between  $h_{\mu\nu}$  and the actual energy flux that is carried off by these time varying metric coefficients. Relating metric coefficients to energy is not trivial. To see how to do this, start with a simpler toy problem. Imagine that the wave equation for general relativity looked like this:

$$-\frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2} + \nabla^2 \Phi = 4\pi G \rho \quad (391)$$

This is what a relativistic theory would look like if the source  $\rho$  were just a simple scalar quantity, instead of a component of a stress tensor. Then, if we multiply by  $(1/4\pi G)\partial\Phi/\partial t$ , integrate  $(\partial\Phi/\partial t)\nabla^2\Phi$  by parts and regroup, this leads to

$$-\frac{1}{8\pi G}\frac{\partial}{\partial t}\left[\left(\frac{\partial\Phi}{\partial t}\right)^2 + |\nabla\Phi|^2\right] + \nabla\cdot\left(\frac{1}{4\pi G}\frac{\partial\Phi}{\partial t}\nabla\Phi\right) = \rho\frac{\partial\Phi}{\partial t}. \quad (392)$$

But

$$\rho\frac{\partial\Phi}{\partial t} = \frac{\partial(\rho\Phi)}{\partial t} - \Phi\frac{\partial\rho}{\partial t} = \frac{\partial(\rho\Phi)}{\partial t} + \Phi\nabla\cdot(\rho\mathbf{v}) = \frac{\partial(\rho\Phi)}{\partial t} + \nabla\cdot(\rho\mathbf{v}\Phi) - \rho\mathbf{v}\cdot\nabla\Phi \quad (393)$$

where  $\mathbf{v}$  is the velocity and the mass conservation equation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0$$

has been used in the second “=” sign from the left. Combining (392) and (393), and then rearranging the terms a bit leads to

$$\frac{\partial}{\partial t}\left[\rho\Phi + \frac{1}{8\pi G}\left(\left(\frac{\partial\Phi}{\partial t}\right)^2 + |\nabla\Phi|^2\right)\right] + \nabla\cdot\left(\rho\mathbf{v}\Phi - \frac{1}{4\pi G}\frac{\partial\Phi}{\partial t}\nabla\Phi\right) = \rho\mathbf{v}\cdot\nabla\Phi \quad (394)$$

The right side is just minus the rate at which work is being done on the sources per unit volume. (The force per unit volume, you recall, is  $-\rho\nabla\Phi$ .) For the usual case of interest when the source  $\rho$  vanishes outside a certain radius, the left side may then be readily interpreted as a far-field wave energy density of

$$\mathcal{E} = [(\partial_t\Phi)^2 + |\nabla\Phi|^2]/8\pi G \quad (\text{Scalar gravity energy density}) \quad (395)$$

and a wave energy flux of

$$\mathcal{F} = -(\partial_t\Phi)\nabla\Phi/4\pi G \quad (\text{Scalar gravity energy flux}) \quad (396)$$

(Is the sign of the flux sensible for outgoing waves?) The question we raise here is whether an analogous method might work on the more involved linear wave equation of tensorial general relativity. The answer is YES, but we have to set things up properly. And, needless to say, it is a bit more messy index-wise!

#### 7.4.2 A conserved energy flux for linearised gravity

*Summary: We identify the energy flux for gravitational radiation in the TT gauge by showing that its 4-divergence is equal to the rate of doing work by a suitable Newtonian potential on the material sources.*

Start with equation (344):

$$\square\bar{h}_{\mu\nu} - \frac{\partial^2\bar{h}_\mu^\lambda}{\partial x^\nu\partial x^\lambda} - \frac{\partial^2\bar{h}_\nu^\lambda}{\partial x^\mu\partial x^\lambda} + \eta_{\mu\nu}\frac{\partial^2\bar{h}^{\lambda\rho}}{\partial x^\lambda\partial x^\rho} = -\frac{16\pi GT_{\mu\nu}}{c^4}. \quad (397)$$

To facilitate the construction of a conserved flux, we need to rewrite the final term on the left side of this equation. Contract on  $\mu\nu$ : the first term on the left becomes  $\square\bar{h}$ , the second

and third each become  $-\partial^2 \bar{h}^{\lambda\rho}/\partial x^\lambda \partial x^\rho$ , while the final contraction turns  $\eta_{\mu\nu}$  into a factor of 4. (Why?) This leads us to

$$\square \bar{h} + 2 \frac{\partial^2 \bar{h}^{\lambda\rho}}{\partial x^\lambda \partial x^\rho} = -\kappa T \quad (398)$$

where we have written  $\kappa = 16\pi G/c^4$ . Using this to substitute for  $\partial_\lambda \partial_\rho \bar{h}^{\lambda\rho}$ , we may recast our original equation as

$$\square \bar{h}_{\mu\nu} - \frac{\partial^2 \bar{h}_\mu^\lambda}{\partial x^\nu \partial x^\lambda} - \frac{\partial^2 \bar{h}_\nu^\lambda}{\partial x^\mu \partial x^\lambda} - \frac{\eta_{\mu\nu}}{2} \square \bar{h} = -\kappa S_{\mu\nu} \quad (399)$$

where we have introduced the source function

$$S_{\mu\nu} = T_{\mu\nu} - \frac{\eta_{\mu\nu} T}{2} \quad (400)$$

Now we may use the same technique we employed for our toy problem. Multiply (399) by  $\partial \bar{h}^{\mu\nu}/\partial x^\sigma$ , summing over  $\mu$  and  $\nu$  as usual but keeping  $\sigma$  free. We want to fold everything on the left side into a divergence. The first term on the left becomes

$$\begin{aligned} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma} \square \bar{h}_{\mu\nu} &= \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma} \frac{\partial^2 \bar{h}_{\mu\nu}}{\partial x_\rho \partial x^\rho} = \frac{\partial}{\partial x_\rho} \left( \frac{\partial \bar{h}_{\mu\nu}}{\partial x^\rho} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma} \right) - \frac{\partial \bar{h}_{\mu\nu}}{\partial x^\rho} \frac{\partial^2 \bar{h}^{\mu\nu}}{\partial x_\rho \partial x^\sigma} \\ &= \frac{\partial}{\partial x_\rho} \left( \frac{\partial \bar{h}_{\mu\nu}}{\partial x^\rho} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma} \right) - \frac{\partial \bar{h}_{\mu\nu}}{\partial x^\rho} \frac{\partial}{\partial x^\sigma} \frac{\partial \bar{h}^{\mu\nu}}{\partial x_\rho} = \frac{\partial}{\partial x_\rho} \left( \frac{\partial \bar{h}_{\mu\nu}}{\partial x^\rho} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma} \right) - \frac{\partial}{\partial x^\sigma} \left( \frac{1}{2} \frac{\partial \bar{h}_{\mu\nu}}{\partial x^\rho} \frac{\partial \bar{h}^{\mu\nu}}{\partial x_\rho} \right) \end{aligned} \quad (401)$$

Do you see why the final equality is valid for the  $\partial/\partial x^\sigma$  exact derivative? For this purpose, it doesn't matter which group of  $\mu\nu$ 's (or  $\rho$ 's) on the  $\bar{h}$ 's (or the  $\partial/\partial x$ ) is the up and which is down!

Now that you've seen the tricks of the trade, you should be able to juggle the indices with me and recast every single term on the left as an exact divergence. The second term on the left is

$$-\frac{\partial^2 \bar{h}_\mu^\lambda}{\partial x^\nu \partial x^\lambda} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma} = -\frac{\partial^2 \bar{h}^{\lambda\mu}}{\partial x_\nu \partial x^\lambda} \frac{\partial \bar{h}_{\mu\nu}}{\partial x^\sigma} = -\frac{\partial}{\partial x_\nu} \left( \frac{\partial \bar{h}^{\lambda\mu}}{\partial x^\lambda} \frac{\partial \bar{h}_{\mu\nu}}{\partial x^\sigma} \right) + \frac{\partial \bar{h}^{\lambda\mu}}{\partial x^\lambda} \frac{\partial^2 \bar{h}_{\mu\nu}}{\partial x^\sigma \partial x_\nu}.$$

Replacing dummy-index  $\nu$  with dummy-index  $\rho$  in the first group on the right, and recognising an exact derivative in the second group on the right,

$$-\frac{\partial^2 \bar{h}_\mu^\lambda}{\partial x^\nu \partial x^\lambda} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma} = -\frac{\partial}{\partial x_\rho} \left( \frac{\partial \bar{h}^{\lambda\mu}}{\partial x^\lambda} \frac{\partial \bar{h}_{\mu\rho}}{\partial x^\sigma} \right) + \frac{1}{2} \frac{\partial}{\partial x^\sigma} \left( \frac{\partial \bar{h}^{\lambda\mu}}{\partial x^\lambda} \frac{\partial \bar{h}_{\mu\nu}}{\partial x_\nu} \right). \quad (402)$$

The third term of our equation is

$$-\frac{\partial^2 \bar{h}_\nu^\lambda}{\partial x^\mu \partial x^\lambda} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma}.$$

But this is exactly the same as what we've just done: simply interchange the dummy indices  $\mu$  and  $\nu$  and remember that  $\bar{h}^{\mu\nu}$  is symmetric in  $\mu\nu$ . No need to do any more here. The fourth and final term of the left side of equation is

$$-\frac{1}{2} \frac{\partial \bar{h}}{\partial x^\sigma} \frac{\partial^2 \bar{h}}{\partial x^\rho \partial x_\rho} = -\frac{1}{2} \frac{\partial}{\partial x_\rho} \left( \frac{\partial \bar{h}}{\partial x^\rho} \frac{\partial \bar{h}}{\partial x^\sigma} \right) + \frac{1}{4} \frac{\partial}{\partial x^\sigma} \left( \frac{\partial \bar{h}}{\partial x^\rho} \frac{\partial \bar{h}}{\partial x_\rho} \right). \quad (403)$$

Thus, after dividing our fundamental equation by  $2\kappa$ , the left side of equation (399) takes on a nice compact form, and we find

$$\frac{\partial \mathcal{U}_{\rho\sigma}}{\partial x_\rho} = -\frac{1}{2} S_{\mu\nu} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma}, \quad (404)$$

where

$$\mathcal{U}_{\rho\sigma} = \mathcal{T}_{\rho\sigma} + \eta_{\rho\sigma} \mathcal{S}. \quad (405)$$

$\mathcal{S}$  is the scalar density:

$$\mathcal{S} = -\left(\frac{1}{4\kappa} \frac{\partial \bar{h}_{\mu\nu}}{\partial x^\rho} \frac{\partial \bar{h}^{\mu\nu}}{\partial x_\rho}\right) + \frac{1}{2\kappa} \left(\frac{\partial \bar{h}^{\lambda\mu}}{\partial x^\lambda} \frac{\partial \bar{h}_{\mu\nu}}{\partial x_\nu}\right) + \frac{1}{8\kappa} \left(\frac{\partial \bar{h}}{\partial x^\rho} \frac{\partial \bar{h}}{\partial x_\rho}\right), \quad (406)$$

and  $\mathcal{T}_{\rho\sigma}$  is a flux tensor:

$$\mathcal{T}_{\rho\sigma} = \frac{1}{2\kappa} \left(\frac{\partial \bar{h}_{\mu\nu}}{\partial x^\rho} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma}\right) - \frac{1}{\kappa} \left(\frac{\partial \bar{h}^{\lambda\mu}}{\partial x^\lambda} \frac{\partial \bar{h}_{\mu\rho}}{\partial x^\sigma}\right) - \frac{1}{4\kappa} \left(\frac{\partial \bar{h}}{\partial x^\rho} \frac{\partial \bar{h}}{\partial x^\sigma}\right). \quad (407)$$

Lots of terms, but in the right gauge they almost all vanish! Note first that by working with plane waves in harmonic coordinates,  $\partial \bar{h}^{\lambda\mu}/\partial x^\lambda = 0$ , some terms disappear and  $\mathcal{U}_{\rho\sigma}$  becomes manifestly symmetric in  $\rho\sigma$ . (This symmetry must then always be true, even if it is not blatant: the tensor  $\mathcal{U}_{\rho\sigma} - \mathcal{U}_{\sigma\rho}$  is gauge invariant and vanishes for the harmonic gauge, so it vanishes for all gauges.) Remembering as well that  $k_\rho k^\rho = 0$  for the TT gauge, matters dramatically simplify (almost everything vanishes) and we find the elegant result

$$\mathcal{U}_{\rho\sigma} = \frac{c^4}{32\pi G} \left(\frac{\partial \bar{h}_{\mu\nu}}{\partial x^\rho} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma}\right) \quad (\text{TT gauge}). \quad (408)$$

Why did we choose to divide by  $2\kappa$  for our overall normalisation constant? Why not just  $\kappa$ , or for that matter,  $4\kappa$ ? It is the right side of our energy equation that answers this. This is

$$-\frac{1}{2} S_{\mu\nu} \frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma} = -\frac{1}{2} \left(T_{\mu\nu} - \frac{\eta_{\mu\nu}}{2} T\right) \left(\frac{\partial h^{\mu\nu}}{\partial x^\sigma} - \frac{\eta^{\mu\nu}}{2} \frac{\partial h}{\partial x^\sigma}\right) = -\frac{1}{2} T_{\mu\nu} \frac{\partial h^{\mu\nu}}{\partial x^\sigma}. \quad (409)$$

Choose  $\sigma = 0$ , the time component. We work in the Newtonian limit  $h^{00} \simeq -2\Phi/c^2$ , where  $\Phi$  is a Newtonian gravitational potential. In the  $\mu\nu$  summation on the right side of the equation, we are then dominated by the 00 components of both  $h^{\mu\nu}$  and  $T_{\mu\nu}$ . Now, we are about to do a number of integration by parts. We will always ignore an exact derivative! Why? Because the exact derivative of a periodic function (and everything here *is* periodic) must oscillate away to zero on average. It is not a sustained energy source for the waves. But in general the products of the periodic functions do not oscillate to zero; for example the average of  $\cos^2(\omega t) = 1/2$ . This is a sustained source. Thus we keep these product terms, only if they are not an exact derivative. Using the right arrow  $\rightarrow$  to mean “integrate by parts and ignore the pure derivatives” (as inconsequential for wave losses), we perform the following manipulations on the right side of equation (409):

$$-\frac{1}{2} T_{00} \frac{\partial h^{00}}{\partial x^0} \rightarrow \frac{1}{2} \frac{\partial T_{00}}{\partial x^0} h^{00} = -\frac{1}{2} \frac{\partial T^{0i}}{\partial x^i} h^{00} \rightarrow \frac{1}{2} T^{0i} \frac{\partial h^{00}}{\partial x^i} \simeq -\rho \frac{\mathbf{v}}{c} \cdot \nabla \Phi, \quad (410)$$

where the first equality follows from  $\partial_\nu T^{0\nu} = 0$ . We have arrived on the right at an expression for the rate at which the effective Newtonian potential does net work on the matter. Why

is that  $1/c$  there? Don't worry, it cancels out with the same factor on the left (flux) side of the original equation. What about the sign of this? This expression is *negative* if the force  $-\rho\nabla\Phi$  is oppositely directed to the velocity, so that the source is losing energy by generating outgoing waves. Our harmonic gauge expression (408) for  $\mathcal{U}_{0i}$  is also negative for an *outward* flowing wave that is a function of the argument  $(r - ct)$ ,  $r$  being spherical radius and  $t$  time. By contrast,  $\mathcal{U}^{0i}$  would be positive, as befits an outward moving wave energy.

The fact that division by  $2\kappa$  produces a source corresponding to the rate at which work is done on the Newtonian sources (when  $\sigma = 0$ ) means that our overall normalisation is indeed correct. The  $\sigma = 0$  energy flux of (408) is the true energy flux of gravitational radiation in the weak field limit:

$$\mathcal{F}^i = \mathcal{F}_i = c\mathcal{U}^{i0} = -c\mathcal{U}_{i0} = -\frac{c^4}{32\pi G} \left( \frac{\partial \bar{h}_{\mu\nu}}{\partial x^i} \frac{\partial \bar{h}^{\mu\nu}}{\partial t} \right) \quad (\text{TT gauge}). \quad (411)$$

## 7.5 The energy loss formula for gravitational waves

*Summary:* Einstein's classic formula for the total gravitational energy radiated by a source is  $G/(5c^9) \ddot{J}_{ij} \ddot{J}_{ij}$ , where  $J_{ij}$  is the traceless form of the source's moment of inertia tensor:  $J_{ij} = I_{ij} - \delta_{ij} I_{kk}/3$ .

Our next step is to evaluate the transverse and traceless components of  $h_{ij}$ , denoted  $h_{ij}^{TT}$ , in terms of the transverse and traceless components of  $I_{ij}$ . Begin with the traceless component, denoted  $J_{ij}$ :

$$J_{ij} = I_{ij} - \frac{\delta_{ij}}{3} I \quad (412)$$

where  $I$  is the trace of  $I_{ij}$ . Next, we address the transverse property. The projection of a vector  $\mathbf{v}$  onto a plane perpendicular to a unit direction vector  $\mathbf{n}$  is accomplished simply by removing the component of  $\mathbf{v}$  along  $\mathbf{n}$ . Denoting the resulting projected vector as  $\mathbf{w}$ ,

$$\mathbf{w} = \mathbf{v} - (\mathbf{n} \cdot \mathbf{v})\mathbf{n} \quad (413)$$

or

$$w_j = (\delta_{ij} - n_i n_j) v_i \equiv P_{ij} v_i \quad (414)$$

where we have introduced the projection tensor

$$P_{ij} = \delta_{ij} - n_i n_j,$$

with the easily shown properties

$$n_i P_{ij} = n_j P_{ij} = 0, \quad P_{ij} P_{jk} = P_{ik}, \quad P_{ii} = 2. \quad (415)$$

Notice that  $P_{ij} P_{jk} = P_{ik}$  simply states that if you project something that is already projected, nothing is changed!

Projecting tensor components presents no difficulties,

$$w_{ij} = P_{ik} P_{jl} v_{kl} \rightarrow n_i w_{ij} = n_j w_{ij} = 0, \quad (416)$$

nor does the extraction of a projected tensor that is both traceless and transverse:

$$w_{ij}^{TT} = \left( P_{ik} P_{jl} - \frac{1}{2} P_{ij} P_{kl} \right) v_{kl} \rightarrow w_{ii}^{TT} = (P_{ik} P_{il} - P_{kl}) v_{kl} = (P_{kl} - P_{kl}) v_{kl} = 0. \quad (417)$$

Using the properties of  $P_{ij}$  given in (415), it is a straightforward calculation to show that the analogue of  $P_{ij}P_{jk} = P_{ik}$  for the the traceless transverse operator is:

$$(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl})(P_{im}P_{jn} - \frac{1}{2}P_{ij}P_{mn}) = (P_{mk}P_{nl} - \frac{1}{2}P_{kl}P_{mn}). \quad (418)$$

An informal way to say this is that if you take the transverse-traceless part of a tensor that is already transverse and traceless, you don't change anything! Knowing this simple mathematical identity in advance will save a great deal of calculational effort.

Let us define

$$J_{ij}^{TT} = \left( P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl} \right) J_{kl}. \quad (419)$$

Now do the following exercise.

*Exercise.* Prove that  $J_{ij}^{TT}$  is also given by

$$J_{ij}^{TT} = \left( P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl} \right) I_{kl},$$

so that nothing is changed by using only the traceless form of  $I_{ij}$  in our calculations,  $J_{ij}$ . Yet another ploy to save computational effort.

Notice now that  $(J_{ij} - J_{ij}^{TT})J_{ij}^{TT}$  is the contraction of the part of  $J_{ij}$  that has no transverse component with the part that is completely transverse. This ought to vanish, if there is any justice. Happily, it does:

$$(J_{ij} - J_{ij}^{TT})J_{ij}^{TT} = J_{ij}J_{ij}^{TT} - J_{kl}(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl})(P_{im}P_{jn} - \frac{1}{2}P_{ij}P_{mn})J_{mn} \quad (420)$$

Following (418), the right side of this is just

$$J_{ij}J_{ij}^{TT} - J_{kl}(P_{mk}P_{nl} - \frac{1}{2}P_{kl}P_{mn})J_{mn} = J_{ij}J_{ij}^{TT} - J_{kl}J_{kl}^{TT} = 0 \quad (421)$$

This will come in handy in a moment.

Next, we write down the traceless-transverse part of the quadrupole formula:

$$h_{ij}^{TT} = \frac{2G}{c^6 r} \frac{d^2 J_{ij}^{TT}}{dt'^2}. \quad (422)$$

Recalling that  $t' = t - r/c$  and the  $J^{TT}$ 's are functions of  $t'$  (not  $t$ !),

$$\frac{\partial h_{ij}^{TT}}{\partial t} = \frac{2G}{c^6 r} \frac{d^3 J_{ij}^{TT}}{dt'^3}, \quad \frac{\partial h_{ij}^{TT}}{\partial r} = -\frac{2G}{c^7 r} \frac{d^3 J_{ij}^{TT}}{dt'^3} \quad (423)$$

where, in the second expression we retain only the dominant term in  $1/r$ . The radial flux of gravitational waves is then given by (411):

$$\mathcal{F}_r = \frac{G}{8\pi r^2 c^9} \frac{d^3 J_{ij}^{TT}}{dt'^3} \frac{d^3 J_{ij}^{TT}}{dt'^3} \quad (424)$$

The  $1/c^9$  dependence ultimately translates into a  $1/c^5$  dependence for Newtonian sources, since each of the  $J$ 's carries a  $c^2$  factor.

The final step is to write out  $J_{ij}^{TT}$  in terms of the  $J_{ij}$  via the TT projection operator. It is here that the fact that  $J_{ij}$  is traceless is a computational help.

$$\left(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl}\right) = (\delta_{ik} - n_i n_k)(\delta_{jl} - n_j n_l) - \frac{1}{2}(\delta_{ij} - n_i n_j)(\delta_{kl} - n_k n_l) \quad (425)$$

With  $J_{kl}$  traceless, we may work this through to find

$$J_{ij}^{TT} = \left(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl}\right) J_{kl} = J_{ij} + \frac{1}{2}(\delta_{ij} + n_i n_j)n_k n_l J_{kl} - n_i n_k J_{jk} - n_j n_k J_{ik} \quad (426)$$

Now,

$$\ddot{J}_{ij}^{TT} \ddot{J}_{ij}^{TT} = [\ddot{J}_{ij} + (\ddot{J}_{ij}^{TT} - \ddot{J}_{ij})] \ddot{J}_{ij}^{TT}. \quad (427)$$

But we've seen in (421) that

$$(\ddot{J}_{ij}^{TT} - \ddot{J}_{ij}) \ddot{J}_{ij}^{TT} = 0,$$

so we are left with

$$\begin{aligned} \ddot{J}_{ij}^{TT} \ddot{J}_{ij}^{TT} &= \ddot{J}_{ij} \ddot{J}_{ij}^{TT} \equiv \ddot{J}_{ij} (\ddot{J}_{ij} + \frac{1}{2}(\delta_{ij} + n_i n_j)n_k n_l \ddot{J}_{kl} - n_i n_k \ddot{J}_{jk} - n_j n_k \ddot{J}_{ik}) = \\ &\ddot{J}_{ij} \ddot{J}_{ij} - 2n_j n_k \ddot{J}_{ij} \ddot{J}_{ik} + \frac{1}{2}n_i n_j n_k n_l \ddot{J}_{ij} \ddot{J}_{kl} \end{aligned} \quad (428)$$

Remember, the vector  $\mathbf{n}$  is just the unit vector pointing in the direction of spherical angles  $\theta$  and  $\phi$ . In Cartesian coordinates:

$$\mathbf{n} = (n_x, n_y, n_z) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta).$$

The total gravitational wave luminosity is an integration of the distribution (428) over all solid angles,

$$L_{GW} = \int r^2 \mathcal{F}_r d\Omega = \frac{G}{8\pi c^9} \int \ddot{J}_{ij}^{TT} \ddot{J}_{ij}^{TT} d\Omega. \quad (429)$$

To evaluate this, you will need

$$\int n_i n_j d\Omega = \frac{4\pi}{3} \delta_{ij}. \quad (430)$$

This is pretty simple: if the two components  $n_i$  and  $n_j$  of the unit normal vector are not the same, the integral vanishes by symmetry (e.g. the average of  $xy$  over a sphere is zero). That means the integral is proportional to a delta function, say  $C\delta_{ij}$ . To get the constant of proportionality  $C$ , just take the trace of both sides:  $\int d\Omega = 4\pi = 3C$ , so  $C = 4\pi/3$ . Much more scary-looking is the other identity you'll need:

$$\int n_i n_j n_k n_l d\Omega = \frac{4\pi}{15} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{kj}). \quad (431)$$

For example, the combination  $n_x n_x n_y n_z$  would be  $\sin^3 \theta \cos \theta \cos^2 \phi \sin \phi$ . But keep calm and think. The only way the integral *cannot* vanish is if two of the indices agree with one another and the remaining two indices also agree with one another. Maybe the second agreeing pair is just the same as the first, maybe not. (The  $n_x n_x n_y n_z$  example provided clearly vanishes when integrated over solid angles, as you can easily show.) This pairwise index agreement

rule is precisely what the additive combination of delta functions ensures, symmetrically summed over the three different ways by which the pairwise index agreement can occur. To get the  $4\pi/15$  factor, simply set  $i = j$  and  $k = l$ , then sum. The integral on the left is then trivially  $\int n_i n_i n_l n_l d\Omega \equiv \int d\Omega = 4\pi$ . The combination of delta functions is  $9 + 3 + 3 = 15$ . Hence the normalisation factor  $4\pi/15$ . Putting this all together via (424), (428), (430) and (431), remembering  $J_{ii} = 0$ , and carrying out the angular integral, the total gravitational luminosity is thus given by

$$L_{GW} = \frac{G}{8\pi c^9} \times \left[ 4\pi - 2 \times \frac{4\pi}{3} + \frac{1}{2} \times \frac{4\pi}{15} \times (0 + 1 + 1) \right] \ddot{J}_{ij} \ddot{J}_{ij},$$

which amounts to:

$$L_{GW} = \frac{G}{5c^9} \ddot{J}_{ij} \ddot{J}_{ij} = \frac{G}{5c^9} \left( \ddot{I}_{ij} \ddot{I}_{ij} - \frac{1}{3} \ddot{I}_{ii} \ddot{I}_{jj} \right). \quad (432)$$

The calculation has not been easy. But the endpoint is a beautifully simple formula, a classical result first derived by Albert Einstein<sup>7</sup> in 1918. Indirect verification of this formula came in about 1980 with the analysis of the binary pulsar, some 25 years after Einstein's death. Direct verification with LIGO came in 2015, almost exactly one century after the epochal relativity paper.

## 7.6 Gravitational radiation from binary stars

*Summary: This section contains more advanced material not on the syllabus, and is optional. Later, we will make use of the results for a circular orbit (equation 447 below), but you will be provided with the formula.*

In W72, the detection of gravitational radiation looms as only a very distant possibility, and rightly so. The section covering this topic devotes its attention to the possibility that rapidly rotating neutron stars might, just might, be a good source. Alas, for this to occur the neutron star would have to possess a sizeable and rapidly varying quadrupole moment, and this neutron stars do not seem to possess. Neutron stars are nearly exact spheres, even when rotating rapidly as pulsars. They are in essence perfectly axisymmetric; were they to have any quadrupole moment, it would hardly change with time.

The possibility that *Keplerian orbits* might be interesting for measuring gravitational radiation effects is never mentioned in W72. Certainly ordinary orbits involving ordinary stars are not a promising source. But compact objects (white dwarfs, neutron stars or black holes) in very close binaries, with orbital periods measured in hours, were discovered within two years of the book's publication, and these turn out to be *extremely* interesting. They are the central focus of modern day gravitational wave research. As we have noted earlier, the first confirmation of the existence of gravitational radiation came from the binary pulsar system 1913+16, in which the change in the orbital period from the conversion of mechanical energy into gravitational waves was inferred via the changing interval of the arrival times of the pulsar's signal. The amplitude of the gravitational waves was itself well below the threshold of direct detection at the time, and still would be today at the emitted frequency. Over long enough time scales, a tight binary of compact objects, black holes in the most spectacular manifestation, will lose enough energy through gravitational radiation that the

---

<sup>7</sup>Actually, Einstein found a coefficient of  $1/10$ , not  $1/5$ . Eddington put matters right a few years later. Tricky business, this gravitational radiation.



resulting inspiral goes all the way to completion, and the system either coalesces or explodes. Predictions suggest that there are enough merging binaries in the universe to produce a rather high detection rate: several per year at a minimum. LIGO has already published its first detection, and given how quickly it was found when the threshold detector upgrade was made, there are grounds for optimism for more to come<sup>8</sup>. The final frenzied seconds before coalescence will produce detectable gravitational wave signatures rich in physical content at frequencies for which LIGO has been designed. Black hole merger waveforms can now be determined by high precision numerical calculation (only relatively recently: see, e.g., F. Pretorius 2005, Phys. Rev. Lett. 95, 121101).

Let us apply equation (432) to the case of two point masses in a classical Keplerian orbit. There is of course no contradiction between assuming a classical Newtonian orbit and calculating its relativistic gravitational energy loss. We are working in the regime in which the losses exert only the most tiny change on the orbit over one period, and the masses, though very close to one another by ordinary astronomical standards, are still separated by a distance well beyond their respective Schwarzschild radii. (The numerical calculations make no such restriction, faithfully calculating the orbital evolution all the way down to the final merger of two black holes into one.)

The Newtonian orbital elements are defined on page 74. The separation  $r$  of the two bodies is given as a function of azimuth  $\phi$  as

$$r = \frac{L}{1 + \epsilon \cos \phi} \quad (433)$$

where  $L$  is the semilatus rectum and  $\epsilon$  is the orbital eccentricity. With  $M$  being the total mass of the individual objects,  $M = m_1 + m_2$ ,  $l$  the constant specific angular momentum (we forego  $J$  for angular momentum to avoid confusion with  $J_{ij}$ ), and  $a$  is the semi-major axis, we have

$$r^2 \frac{d\phi}{dt} = l, \quad L = \frac{l^2}{GM} = a(1 - \epsilon^2) \quad (434)$$

and thus

$$\frac{d\phi}{dt} = \left( \frac{GM}{a^3(1 - \epsilon^2)^3} \right)^{1/2} (1 + \epsilon \cos \phi)^2 \quad \frac{dr}{dt} = \left( \frac{GM}{a(1 - \epsilon^2)} \right)^{1/2} \epsilon \sin \phi \quad (435)$$

The distance of each body from the center-of-mass is denoted  $r_1$  and  $r_2$ . Writing these as vector quantities,

$$\mathbf{r}_1 = \frac{m_2 \mathbf{r}}{M}, \quad \mathbf{r}_2 = -\frac{m_1 \mathbf{r}}{M}. \quad (436)$$

Thus, the  $xy$  coordinates in the orbital plane are

$$\mathbf{r}_1 = \frac{m_2 r}{M} (\cos \phi, \sin \phi), \quad \mathbf{r}_2 = \frac{m_1 r}{M} (-\cos \phi, -\sin \phi). \quad (437)$$

The nonvanishing moment tensors  $I_{ij}$  are then

$$I_{xx} = \frac{m_1 m_2^2 + m_1^2 m_2}{M^2} r^2 \cos^2 \phi = \mu r^2 \cos^2 \phi \quad (438)$$

$$I_{yy} = \mu r^2 \sin^2 \phi \quad (439)$$

---

<sup>8</sup>Update September 2017: yes indeed. Several confirmed sources are now being studied, including a merger of neutron stars with baryons flying around everywhere.

$$I_{xy} = I_{yx} = \mu r^2 \sin \phi \cos \phi \quad (440)$$

$$I_{ii} = I_{xx} + I_{yy} = \mu r^2 \quad (441)$$

where  $\mu$  is the reduced mass  $m_1 m_2 / M$ . It is a now lengthy, but utterly straightforward task to differentiate each of these moments three times. If you work through these, begin with the relatively easy  $\epsilon = 0$  case when reproducing the formulae below, though I present the results for finite  $\epsilon$  here:

$$\frac{d^3 I_{xx}}{dt^3} = \alpha (1 + \epsilon \cos \phi)^2 (2 \sin 2\phi + 3\epsilon \sin \phi \cos^2 \phi), \quad (442)$$

$$\frac{d^3 I_{yy}}{dt^3} = -\alpha (1 + \epsilon \cos \phi)^2 [2 \sin 2\phi + \epsilon \sin \phi (1 + 3 \cos^2 \phi)], \quad (443)$$

$$\frac{d^3 I_{xy}}{dt^3} = \frac{d^3 I_{yx}}{dt^3} = -\alpha (1 + \epsilon \cos \phi)^2 [2 \cos 2\phi - \epsilon \cos \phi (1 - 3 \cos^2 \phi)], \quad (444)$$

where

$$\alpha^2 \equiv \frac{4G^3 m_1^2 m_2^2 M}{a^5 (1 - \epsilon^2)^5} \quad (445)$$

Equation (432) yields, after some assembling:

$$L_{GW} = \frac{32}{5} \frac{G^4}{c^5} \frac{m_1^2 m_2^2 M}{a^5 (1 - \epsilon^2)^5} (1 + \epsilon \cos \phi)^4 \left[ (1 + \epsilon \cos \phi)^2 + \frac{\epsilon^2}{12} \sin^2 \phi \right] \quad (446)$$

Our final step is to average  $L_{GW}$  over an orbit. This is not simply an integral over  $d\phi/2\pi$ . We must integrate over time, i.e., over  $dt = d\phi/\dot{\phi}$ , and then divide by the orbital period to produce a *time* average. The answer is

$$\langle L_{GW} \rangle = \frac{32}{5} \frac{G^4}{c^5} \frac{m_1^2 m_2^2 M}{a^5} f(\epsilon) = 1.00 \times 10^{25} m_{\odot 1}^2 m_{\odot 2}^2 M_{\odot} (a_{\odot})^{-5} f(\epsilon) \text{ Watts}, \quad (447)$$

where

$$f(\epsilon) = \frac{1 + (73/24)\epsilon^2 + (37/96)\epsilon^4}{(1 - \epsilon^2)^{7/2}} \quad (448)$$

and  $\odot$  indicates solar units of mass ( $1.99 \times 10^{30}$  kg) and length (one solar radius is  $6.955 \times 10^8$  m). (Peters and Mathews 1963). Something that is of order  $v/c$  is considered to be Newtonian, like a standard kinematic Doppler shift. At order  $v^2/c^2$ , we are in the regime of relativity (e.g. gravitational redshift), and measurements are traditionally quite difficult. Equation (447) tells us that  $L_{GW}$  is a number of order  $E\Omega(v/c)^5$ , where  $E$  is the Keplerian orbital energy,  $\Omega$  the orbital frequency, and  $v$  a typical orbital velocity. No wonder gravitational radiation is so hard to measure!

*Exercise.* Show that following the procedure described above, the time-averaged luminosity  $\langle L_{GW} \rangle_{\text{time}}$  is given by the expression

$$\langle L_{GW} \rangle_{\text{time}} = \frac{32}{5} \frac{G^4}{c^5} \frac{m_1^2 m_2^2 M}{a^5 (1 - \epsilon^2)^{7/2}} \left\langle (1 + \epsilon \cos \phi)^2 \left[ (1 + \epsilon \cos \phi)^2 + \frac{\epsilon^2}{12} \sin^2 \phi \right] \right\rangle_{\text{angle}},$$

where the  $\langle \rangle$  angular average on the right is over  $2\pi$  angles in  $\phi$ . Use the fact that the angular average of  $\cos^2 \phi$  is  $1/2$  and the average of  $\cos^4 \phi$  is  $3/8$  to derive equation (447).

Equations (447) and (448) give the famous gravitational wave energy loss formula for a classical Keplerian orbit. Notice the dramatic effect of finite eccentricity via the  $f(\epsilon)$  function. The first binary pulsar to be discovered, PSR1913+16, has an eccentricity of about 0.62, and thus an enhancement of its gravitational wave energy loss that is boosted by more than an order of magnitude relative to a circular orbit.

This whole problem must have seemed like an utter flight of fancy in 1963, when the calculation was first completed: the concept of a neutron star was barely credible and not taken seriously; the notion of pulsar timing was simply beyond conceptualisation. A lesson, perhaps, that no good calculation of an interesting physics problem ever goes to waste.

*Exercise.* When we studied Schwarzschild orbits, there was an exercise to show that the total Newtonian orbital energy of a bound two body system is  $-Gm_1m_2/2a$  and that the system period is  $2\pi\sqrt{a^3/GM}$ , independent in both cases of the eccentricity. Use these results to show that the orbital period change due to the loss of gravitational radiation is given by

$$\dot{P} = -\frac{192\pi}{5} \left( \frac{m_1m_2}{M^2} \right) \left( \frac{GM}{ac^2} \right)^{5/2} f(\epsilon)$$

with  $M = m_1 + m_2$  as before. This  $\dot{P}$  is a directly measurable quantity.

*Exercise.* Now that you're an expert in the the two-body gravitational radiation problem, let's move on to three! Show that three equal masses revolving around their common centre-of-mass emit no quadrupole gravitational radiation.

## 7.7 Detection of gravitational radiation

### 7.7.1 Preliminary comments

The history of gravitational radiation has been somewhat checkered. Albert Einstein himself stumbled several times, both conceptually and computationally. Arguments of fundamental principle persisted through the early 1960's; technical arguments still go on.

At the core of the early controversy was the question of whether gravitational radiation existed at all! The now classic Peters and Mathews paper of 1963 begins with a disclaimer that they are assuming that the “standard interpretation” of the theory is correct. The confusion concerned whether the oscillatory behaviour of  $h_{\mu\nu}$  potentials were just some sort of mathematical coordinate effect, devoid of any actual physical consequences. Here is an example of how you might get into trouble. We calculate the affine connection  $\Gamma_{\nu\lambda}^\mu$  and apply the geodesic equation,

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma_{\nu\lambda}^\mu \frac{dx^\nu}{d\tau} \frac{dx^\lambda}{d\tau} = 0 \quad (449)$$

and ask what happens to a particle initially at rest with  $dx^\nu/d\tau = (-c, \mathbf{0})$ . The subsequent evolution of the spatial velocity components is then

$$\frac{d^2x^i}{d\tau^2} + \Gamma_{00}^i c^2 = 0 \quad (450)$$

But equation (331) clearly shows that  $\Gamma_{00}^i = 0$  since any  $h$  with a zero index vanishes for our TT plane waves. The particle evidently remains at rest. Is there really no effect of gravitational radiation on ordinary matter?!

Coordinates, coordinates, coordinates. The point, once again, is that coordinates by themselves mean nothing, any more than does the statement “My house is located at the vector (2, 1.3).” By now we should have learned this lesson. We picked our gauge to make life simple, and we have indeed found a very simple coordinate system that remains frozen to the individual particles! That does not mean that the particles don’t physically budge. The *proper* spatial separation between two particles with coordinate separation  $dx^i$  is  $ds^2 = (\eta_{ij} + h_{ij})dx^i dx^j$ , and *that* separation surely is not constant because  $h_{11}$ ,  $h_{22}$ , and  $h_{12} = h_{21}$  are wiggling even while the  $dx^i$  are fixed. Indeed, to first order in  $h_{ij}$ , we may write

$$ds^2 = \eta_{ij}(dx^i + h_{ik}dx^k/2)(dx^j + h_{jm}dx^m/2).$$

This makes the physical interpretation easy: the passing wave increments the initially undisturbed spatial interval  $dx^i$  by an amount  $h_{ik}dx^k/2$ . It was Richard Feynman who in 1955 seems to have given the simplest and most convincing argument for the existence of gravitational waves. If the separation is between two beads on a rigid stick and the beads are free to slide, they will oscillate with the tidal force of the wave. If there is now a tiny bit of stickiness, the beads will heat the stick. Where did that energy come from? It could only be the wave. The “sticky bead argument” became iconic in the relativity community.

Recall that the two independent states of linear polarisation of a gravitational wave are the + and  $\times$  modes, “plus” and “cross.” They behave similarly, just rotated by  $45^\circ$ . The + wave ( $e_{11} = -e_{22}$ ) causes a prolate extension along the vertical part of the plus sign as it passes, then squeezes down along the vertical axis and oblate outward along the horizontal axis, then once again squeezes inward from oblate to prolate once again. The  $\times$  mode ( $e_{21} = e_{12}$ ) shows the same oscillation pattern along a rotation pattern rotated by  $45^\circ$ . (An excellent animation is shown in the Wikipedia article “Gravitational Waves.”) These are the actual true physical distortions caused by the tidal force of the gravitational wave.

In 1968, in the midst of what had been intensively theoretical investigations and debate surrounding the nature of gravitational radiation, a physicist named Joseph Weber calmly announced that he had detected gravitational radiation experimentally in his basement lab, coming in prodigious amounts from the centre of the Milk Way Galaxy, thank you very much. His technique was to use what are now called “Weber bars”, giant cylinders of aluminium fitted with special piezoelectric devices that can convert tiny mechanical oscillations into electrical signals. The gravitational waves distorted these great big bars by a tiny, tiny amount, and the signals were picked up. At least that was the idea. The dimensionless relative strain  $\delta l/l$  of a bar of length  $l$  due to passing wave would be of order  $h_{ij}$ , or  $10^{-21}$  by our optimistic estimate. To make a long, rather sad story very short, Weber was in error in several different ways, and ultimately his experiment was completely discredited. Yet his legacy was not wholly negative. The possibility of actually detecting gravitational waves hadn’t been really taken very seriously up to this point. Post Weber, the idea gradually took hold in the physics establishment. People asked themselves how one might actually go about detecting these signals and what might be learnt from an unambiguous detection. So gravitational radiation detection became part of the mainstream, with leading figures in relativity getting directly involved. The detection of gravitational radiation is not a task for a lone researcher, however clever, working in the basement of university building, any more than was finding the Higgs boson. Substantial resources of the National Science Foundation in the US and an international research team numbering in the thousands were needed for the construction and testing of viable gravitational wave receptors. (In the UK, the University of Glasgow played a critical design role.) Almost fifty years after Weber, the LIGO facility has at last cleanly detected the exquisitely gentle tensorial strains of gravitational waves at the level of  $h \sim 10^{-21}$ . The LIGO mirrors may not have crack’d from side-to-side, but they did flutter a bit in the gravitational breeze. This truly borders on magic: if the effective length of LIGO’s interferometer arm is taken to be  $l = 10$  km, then  $\delta l$  is  $10^{-15}$  cm, one percent of the radius of a proton!

The next exercise is strongly recommended.

**Exercise. Weaker than weak interactions.** Imagine a gravitational detector of two identical masses  $m$  separated by a distance  $l$  symmetrically about the origin along the  $x$ -axis. Along comes a plane wave gravitational wave front, propagating along the  $z$ -axis, with  $h_{xx} = -h_{yy} = A_{xx} \cos(kz - \omega t)$  and no other components. The masses vibrate in response. Show that, to linear order in  $A_{xx}$ ,

$$\ddot{I}_{xx} = \frac{1}{2} m c^4 \omega^3 l^2 A_{xx} \sin \omega t,$$

that there are no other  $\ddot{I}_{ij}$ , and that the masses radiate an average gravitational wave luminosity of

$$\langle L_{GW} \rangle = \frac{G}{60c^5} m^2 \omega^6 l^4 A_{xx}^2$$

Next, show that the average energy flux for our incoming plane wave radiation is, from equation (411),

$$\mathcal{F} = \frac{c^3 \omega^2 A_{xx}^2}{32\pi G}.$$

The cross section for gravitational interaction (dimensions of area) is defined to be the ratio of the average luminosity to the average incoming flux. Why is this a good definition for the cross section? Show that this ratio is

$$\sigma = \frac{8\pi G^2 m^2 \omega^4 l^4}{15c^8} = \frac{2\pi}{15} R_S^2 \left( \frac{\omega l}{c} \right)^4$$

where  $R_S = 2Gm/c^2$  is the Schwarzschild radius of each mass. Evaluate this numerically for  $m = 10\text{kg}$ ,  $l = 10\text{m}$ ,  $\omega = 20 \text{ rad s}^{-1}$  (motivated by GW150914). Compare this with a typical weak interaction cross section of  $10^{-48}\text{m}^2$ . Just how weak is gravitational scattering?

### 7.7.2 Indirect methods: orbital energy loss in binary pulsars

In 1974, a remarkable binary system was discovered by Hulse and Taylor (1975, *ApJ Letters*, 195, L51). One of the stars was a pulsar with a pulse period of 59 milliseconds, i.e., a neutron star that rotates about 17 times a second! The orbital period was 7.75 hours. This is a very tight binary, with a separation not very different from the radius of the Sun. The non-pulsar was not actually seen, only inferred, but the small separation between the two stars together with the absence of any eclipse of the pulsar's beeping signal suggested that the companion was also a compact star. (If the orbital plane were close to being in the plane of the sky to avoid seeing these eclipses, then the pulsar pulses would likewise show no line-of-sight Doppler shifts, in stark contradiction to observations.)

What made this yet more extraordinary is that pulsars are among the most accurate clocks in the Universe, until recently more accurate than any earthbound atomic clock. The most accurately measured pulsar has a pulse period known to 17 significant figures! Indeed, pulsars, which are generally calibrated by ensemble averages of large numbers of atomic clocks, are themselves directly used as standard clocks<sup>9</sup>. Nature obligingly placed its most accurate clock in the middle of a neutron star binary, just the system for which such

<sup>9</sup>Since 2011, a bank of six pulsars, observed from Gdansk Poland, has been monitored continuously as a timekeeping device.

fantastically precise timing is a requirement. This, then, was the ultimate general relativity laboratory.

Classic nonrelativistic binary observation techniques allow one to determine five parameters from observations of the pulsar: the semimajor axis  $a$  projected against the plane of the sky ( $a \sin i$ ), the eccentricity  $e$ , the orbital period  $P$ , and two technical parameters related to the periastron (the point of closest separation). The angular position of the periastron within the orbit (“orbital phase”), and an absolute time reference point for when the periastron occurs.

Relativistic effects, something new and beyond a standard Keplerian orbit analysis, give two more parameters. The first is the *advance* of the perihelion (exactly analogous to Mercury), which in the case of PSR 1913+16 is  $4.2^\circ$  per year. (Compare with Mercury’s trifling 43 arc seconds per century!) The second is the *second order* ( $\sim v^2/c^2$ ) Doppler shift of the pulse period. This arises from both the gravitational redshift of the combined system as well as rotational kinematics. These seven parameters now allow a complete determination of the individual masses and all of the classical orbital parameters of the system, a neat achievement in itself. The masses of the neutron stars are  $1.4414 M_\odot$  and  $1.3867 M_\odot$ , remarkably similar to one another and remarkably similar to the Chandrasekhar mass  $1.42 M_\odot$ <sup>10</sup>. (The digits in the neutron stars’ masses are all significant!) More importantly, there is a third relativistic effect also present, and therefore the problem is over-constrained. That is to say, it is possible to make a *prediction*. The orbital period changes very slowly with time, shortening in duration due to the gradual approach of the two bodies. This “inspiral” is caused by the loss of orbital energy, carried off by gravitational radiation, equation (447). Thus, by monitoring the precise arrival times of the pulsar signals emanating from this slowly decaying orbit, the *existence* of gravitational radiation could be quantitatively confirmed and Einstein’s quadrupole formula verified, even though the radiation itself was not directly observable.

Figure [5] shows the results of many years of observations. The dots are the cumulative change in the time of periastron, due to the ever more rapid orbital period as gravitational radiation losses cause the neutron stars to inspiral. Without the radiation losses, there would still be a periastron advance of course, but the time *between* periastrons would not change: the interval from one periastron to the next would just be a bit longer than an orbital period. It is the accumulated time shift between periastron intervals that is the signature of actual energy loss. The solid line in figure [5] is *not* a fit to the data. It is the prediction of general relativity of what the cumulative change in the “epoch of periastron” (as it is called) should be, according to the energy loss formula of Peters and Mathews, (447). This beautiful precision fit leaves no doubt whatsoever that the quadrupole radiation formula of Einstein is correct. For this achievement, Hulse and Taylor won a well-deserved Nobel Prize in 1993. (It must be just a coincidence that 1992 is about the time that the data points seem to become more sparse.)

Direct detection of gravitational waves is a very recent phenomenon. There are two types of gravitational wave detectors currently in operation. The first is based on a classic 19th century laboratory apparatus, a Michelson interferometer. The second makes use of pulsar emissions—specifically the arrival times of pulses—as a probe of the  $h_{\mu\nu}$  induced by gravitational waves as these waves propagate across our line of site to the pulsar. The interferometer detectors are designed for wave frequencies from  $\sim 10$  Hz to 1000’s of Hz. These are now up and running! By contrast, the pulsar measurements are sensitive to frequencies of tens to hundreds of *micro* Hz—a very different range, measuring physical processes on very different scales. This technique has yet to be demonstrated. The high frequency interferometers measure the gravitational radiation from stellar-mass black holes

---

<sup>10</sup>This is the upper limit to the mass of a white dwarf star. If the mass exceeds this value, it collapses to either a neutron star or black hole, but cannot remain a white dwarf.

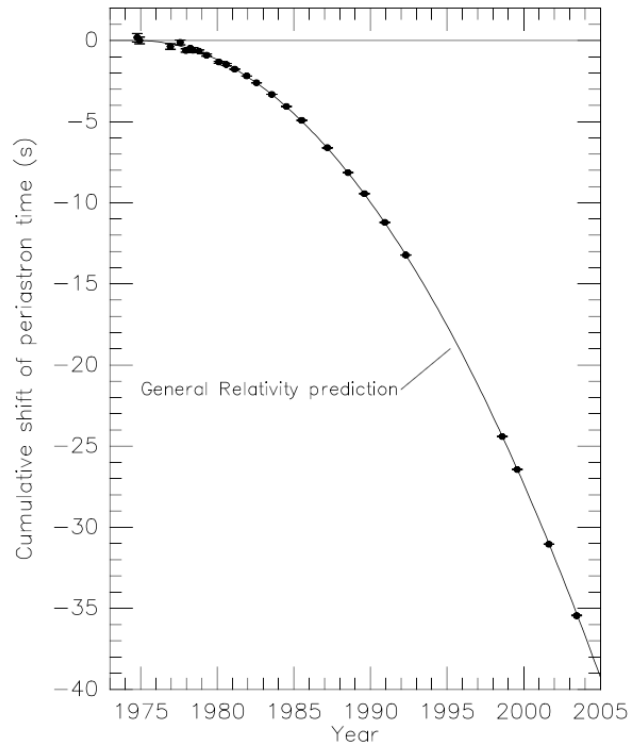


Figure 5: The cumulative change in the periastron event (“epoch”) caused by the inspiral of the pulsar PSR1913+16. The dots are the data, the curve is the *prediction*, not the best fit! This prediction is confirmed to better than a fraction of a percent.

or neutron star binaries merging together. The low frequency pulsar timing will also in principle measure black holes merging, but with masses of order  $10^9$  solar masses. These are the masses of black holes that live in the cores of active galaxies.

### 7.7.3 Direct methods: LIGO

LIGO, or Laser Interferometer Gravitational-Wave Observatory, detects gravitational waves as described in figure (6). In the absence of a wave, the arms are set to destructively interfere, so that no light whatsoever reaches the detector. The idea is that a gravitational wave passes through the apparatus, each period of oscillation slightly squeezing one arm, slightly extending the other. With coherent laser light traversing each arm, when this light re-superposes at the centre, the phase will become *ever* so slightly out of precise cancellation, and therefore photons will appear in the detector. In practise, the light makes many passages back and forth along a 4 km arm before analysis. The development of increased sensitivity comes from engineering greater and greater numbers of reflections, and thus a greater effective path length. There are two such interferometers, one in Livingston, Louisiana, the other in Hanford, Washington, a separation of 3000 km. Both must show a simultaneous wave passage (“simultaneous” after taking into account an offset of 10 milliseconds for speed of light travel time!) for the signal to be verified.

This is a highly simplified description, of course. All kinds of ingenious amplification and noise suppression techniques go into this project, which is designed to measure induced strains at the incredible level (as of March 2017) of  $10^{-23}$ . This detection is only possible at all because we measure not the flux of radiation, which would have a  $1/r^2$  dependence with distance  $r$  to the source, but the amplitude  $h_{ij}$ , which has a  $1/r$  dependence.

Figure (7) shows a match of an accurate numerical simulation to the processed LIGO event GW150914. I have overlaid three measured wave periods  $P_1$ ,  $P_2$ , and  $P_3$ , with each of their respective lengths given in seconds. (These were measured with a plastic ruler directly from the diagram!) The total duration of these three periods is 0.086 s. Throughout this time the black holes are separated by a distance in excess of  $4 R_S$ , so we are barely at the limit for which we can trust Newtonian orbit theory. Let’s give it a try for a circular orbit. (Circularity is not unexpected for the final throes of coalescence.)

Using the zero eccentricity orbital period decrease formula from the final exercise of §7.6, but remembering that the orbital period  $P$  is twice the gravitational wave period  $P_{GW}$ ,

$$\dot{P}_{GW} = -\frac{96\pi}{5} \left( \frac{m_1 m_2}{M^2} \right) \left( \frac{GM}{ac^2} \right)^{5/2}$$

We eliminate the semi-major axis  $a$  in favour of the measured period  $P_{GW}$ ,

$$P^2 = \frac{4\pi^2 a^3}{GM}, \quad \text{whence } P_{GW}^2 = \frac{\pi^2 a^3}{GM}$$

This gives

$$\dot{P}_{GW} = -\frac{96\pi^{8/3}}{5c^5} \left( \frac{GM_c}{P_{GW}} \right)^{5/3} \quad (451)$$

where we have introduced what is known as the “chirp mass”  $M_c$ ,

$$M_c = \frac{(m_1 m_2)^{3/5}}{M^{1/5}} \quad (452)$$



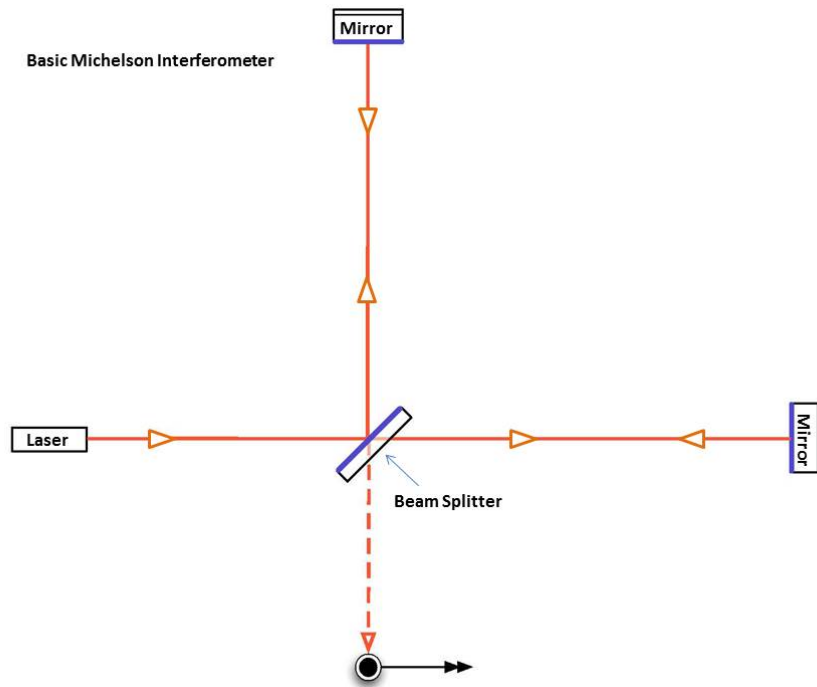


Figure 6: A schematic interferometer. Coherent light enters from the laser at the left. Half is deflected  $45^\circ$  upward by the beam splitter, half continues on. The two halves reflect from the mirrors. The beams re-superpose at the splitter, interfere, and are passed to a detector at the bottom. If the path lengths are identical or differ by an integral number of wavelengths they interfere constructively; if they differ by an odd number of half-wavelengths they cancel one another. In “null” mode, the two arms are set to destructively interfere so that no light whatsoever reaches the detector. A passing gravity wave just barely offsets this precise destructive interference and causes laser photons to appear in the detector.

The chirp mass (so-named because if the gravitational wave were audible at these same frequencies, it would indeed sound like a chirp!) is the above combination of  $m_1$  and  $m_2$ , which is directly measurable from  $P_{GW}$  and its derivative. It is easy to show (try it!) that  $M = m_1 + m_2$  is a minimum when  $m_1 = m_2$ , in which case

$$m_1 = m_2 \simeq 1.15M_c.$$

Putting numbers in (451), we find

$$M_{c\odot} = -5.522 \times 10^3 P_{GW} \dot{P}_{GW}^{3/5}, \quad (453)$$

where  $M_{c\odot}$  is the chirp mass in solar masses and  $P_{GW}$  is measured in seconds. From the GW150914 data, we estimate

$$\dot{P}_{GW} \simeq \frac{P_3 - P_1}{P_1 + P_2 + P_3} = \frac{-0.0057}{0.086} = -0.0663,$$

and for  $P_{GW}$  we use the midvalue  $P_2 = 0.0283$ . This yields

$$M_{c\odot} \simeq 30.7 \quad (454)$$

compared with “ $M_c \simeq 30M_\odot$ ” in Abbot et al. (2016). I’m sure this remarkable level of agreement is somewhat (but not entirely!) fortuitous. Even in this, its simplest presentation, the wave form presents a wealth of information. The “equal mass” coalescing black hole system comprises two  $35M_\odot$  black holes, and certainly at that mass a compact object can only be a black hole!

The two masses need not be equal of course, so is it possible that this is something other than a coalescing black hole binary? We can quickly rule out any other possibility, without a sophisticated analysis. It cannot be any combination of white dwarfs or neutron stars, because the chirp mass is too big. Could it be, say, a black hole plus a neutron star? With a fixed observed  $M_c = 30M_\odot$ , and a neutron star of at most  $\sim 2M_\odot$ , the black hole would have to be some  $1700M_\odot$ . So? Well, then the Schwarzschild radius would have to be very large, and coalescence would have occurred at a separation distance too large for any of the observed high frequencies to be generated! There are frequencies present toward the end of the wave form event in excess of 75 Hz. This is completely incompatible with a black hole mass of this magnitude.

A sophisticated analysis using accurate first principle numerical simulations of gravitational wave from coalescing black holes tells an interesting history, one rather well-captured by even our naive efforts. Using a detailed match to the waveform, the following can be deduced. The system lies at a distance of some 400 Mpc, with significant uncertainties here of order 40%. At these distances, the wave form needs to be corrected for cosmological expansion effects, and the masses in the source rest frame are  $36M_\odot$  and  $29M_\odot$ , with  $\pm 15\%$  uncertainties. The final mass,  $62M_\odot$  is definitely less than the sum of the two,  $65M_\odot$ : some  $3M_\odot c^2$  worth of energy has disappeared in gravitational waves! A release of  $5 \times 10^{47}$  J is, I believe, the largest explosion of any kind ever recorded. A billion years later, some of that energy, in the form of ripples in space itself, tickles the interferometer arms in Louisiana and Washington. It is, I believe, at  $10^{-15}$  cm, the smallest amplitude mechanical motion ever recorded.

What a story.

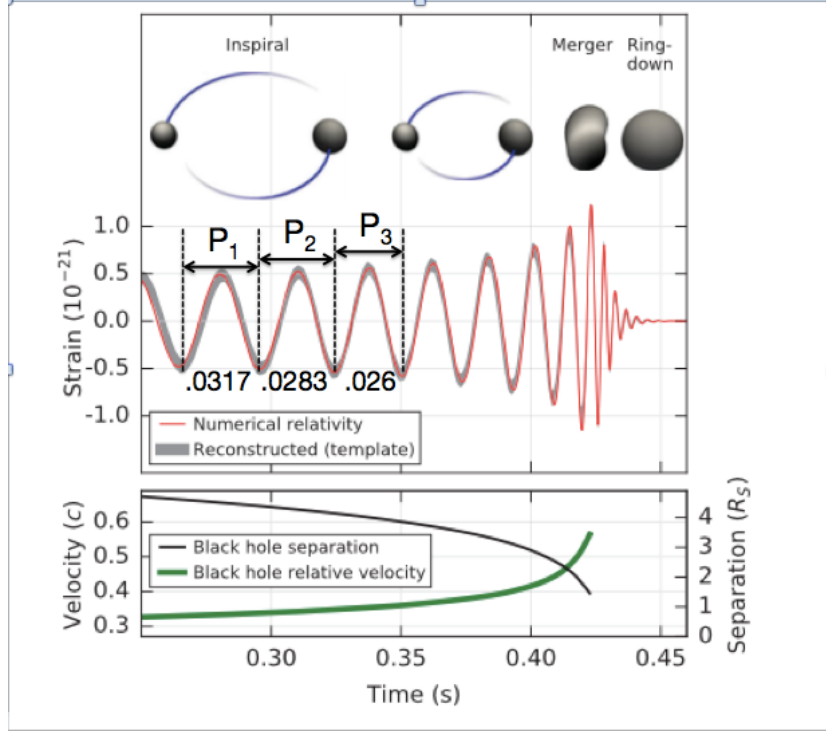


Figure 7: From Abbot et al. (2016). The upper diagram is a schematic rendering of the black hole inspiral process, from slowly evolution in a quasi-Newtonian regime, to a strongly interacting regime, followed by a coalescence and “ring-down,” as the emergent single black hole settles down to its final, nonradiating geometry. The middle figure is the gravitational wave strain, overlaid with three identified periods discussed in the text. The final bottom plot shows the separation of the system and the relative velocity as a function of time, from inspriial just up to the moment of coalescence.

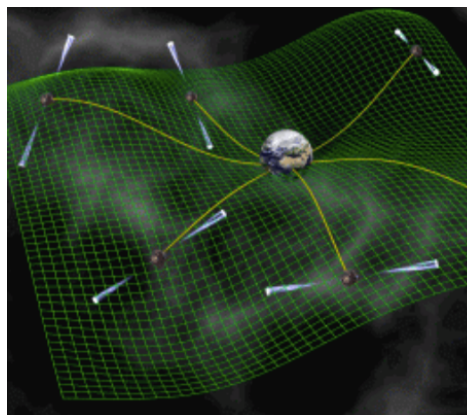


Figure 8: A schematic view of a gravitational wave passing through an array of pulsar probes.

#### 7.7.4 Direct methods: Pulsar timing array

Pulsars are, as we have noted, fantastically precise clocks. Within the pulsar cohort, those with millisecond periods are the most accurate of all. The period of PSR1937+21 is known to be 1.5578064688197945 milliseconds, an accuracy of one part in  $10^{17}$ . (That is equivalent to knowing the age of the Universe to within a few seconds!) One can then predict the arrival time of a pulse to this level of accuracy as well. By constraining variations in pulse arrival times from a single pulsar, we can set an upper limit to amount of gravitational radiation that the signal has traversed. But we don't just have one pulsar. So why settle for one pulsar and mere constraints? We know of many pulsars, distributed more or less uniformly through the galaxy. If the arrival times from this "pulsar timing array" (PTA) were correlated with one another in a mathematically calculable manner, this would be a direct indication of the the deformation of space caused by the passage of a gravitational wave. This technique is sensitive to very long wavelength gravitational radiation, light-years in extent. This is very difficult to do because all other sources introducing a spurious correlation must be scrupulously eliminated. LIGO too has noise issues, but unlike pulsar blips propagating through the interstellar medium, LIGO's signal is very clean and all hardware is accessible. Thus, PTA has its share of skeptics. At the time of this writing, there are only upper limits from the PTA measurements.

*Despite its name, the big bang theory is not really a theory of a bang at all. It is really only a theory of the aftermath of a bang.*

— Alan Guth

## 8 Cosmology

### 8.1 Introduction

*Summary. When carefully applied, local Newtonian mechanics in a background of outwardly-expanding matter provides a surprisingly good (indeed, unreasonably good) description of much of basic cosmology. Expansion is essential for mathematical self-consistency. Photons propagating in such an expanding space are observed to be redshifted relative to their rest frame: the wavelength increases, as the photon propagates, in proportion to the size of the Universe.*

#### 8.1.1 Newtonian cosmology

The subject of the origin of the Universe is irresistible to the scientist and layperson alike. What went bang? Where did the Universe come from? What happened along the way? Where are we headed? The theory of general relativity, with its rigorous mathematical formulation of the large-scale geometry of spacetime, provides both the conceptual and technical apparatus to understand the structure and evolution of the Universe. We are fortunate to live in an era in which many precise answers to these great questions are at hand. Moreover, while we need general relativity to put ourselves on a truly firm footing, we can get quite far using very simple ideas and hardly any relativity at all! Not only *can*, we absolutely *should* begin this way. Let us start with some very Newtonian dynamics and see what there is to see. Then, knowing a bit of what to expect and where we are headed, we will be in a much better position to revisit “the problem of the Universe” on a fully relativistic basis.

A plausible but naive model of the Universe might be one in which space is ordinary static Euclidian space, and the galaxies fill up this space uniformly (on average) everywhere. Putting aside the question of the origin of such a structure (let’s say it has existed for all time) and the problem that the cumulative light received at any location would be infinite (“Olber’s paradox”—that’s tougher to get around: let’s say maybe we turned on the galaxies at some finite time in the past<sup>11</sup>), the static Euclidian model is not even mathematically self-consistent.

Consider the analysis from figure [9]. In the figure on the left, we note that there are two observers, one at the centre of the sphere 1, the other at the centre of sphere  $r_2$ . Each calculates the expected acceleration at the location of the big black dot, which is a point on the surface of each of the spheres. Our model universe is spherically symmetric about  $r_1$ , but it is also spherically symmetric about  $r_2$ ! Hence the following conundrum:

---

<sup>11</sup>See W72, pp. 611-13.

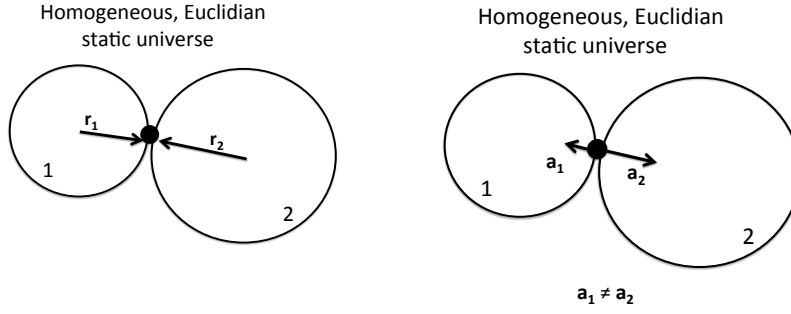


Figure 9: In a static homogeneous Euclidian universe, an observer at the center of sphere 1 would calculate a different gravitational acceleration for the dot than an observer at the centre of sphere 2, i.e.  $\mathbf{a}_1 \neq \mathbf{a}_2$ . But if we take into account the actual relative acceleration of the two observers, each considers the other to be in a *noninertial* frame, and the calculation becomes self-consistent with this non-inertial “fictitious” force. (See text.)

The observer at the centre of sphere 1 ignores the effect of the spherically symmetric mass exterior to the black point and concludes that the acceleration at the dot’s location is

$$a_1 = \frac{GM(\text{within } r_1)}{r_1^2} = \frac{4\pi G\rho r_1}{3}, \quad (455)$$

directed toward the centre of the  $r_1$  sphere. (Here  $\rho$  is meant to be the average uniform mass density of the Universe.) But the observer at the origin of the  $r_2$  sphere claims, by identical reasoning, that the acceleration must be  $a_2 = 4\pi G\rho r_2/3$  directed toward the centre of  $r_2$ ! This shown on the right in figure [9]. Both cannot be correct.

What if the Universe is dynamically active? Then we must put in the gravitational acceleration, in the form of a noninertial reference frame, from the very start of the calculation. If the observers at the centres of  $r_1$  and  $r_2$  are actually accelerating relative to one another, there is no reason to expect that their separate calculations for the black dot acceleration to agree, because the observers are not part of the same inertial frame! Can we make this picture self-consistent somehow for any two  $r_1$  and  $r_2$  observers? Yes. If the Universe exhibits a relative acceleration between two observers that is proportional to the vector difference  $\mathbf{r}_2 - \mathbf{r}_1$  between the two observers’ positions, all is well.

Here is how it works. The observer at the centre of sphere 1 measures the acceleration of the black dot to be  $-4\pi G\rho \mathbf{r}_1/3$  as above, with  $\mathbf{r}_1$  pointing from the centre of circle 1 to the surface dot. The same circle 1 observer finds that the acceleration of the centre of sphere

2 is  $-4\pi G\rho(\mathbf{r}_1 - \mathbf{r}_2)/3$ , where  $\mathbf{r}_2$  is the position vector oriented from the sphere 2 centre toward the big dot. Defined this way, these particular  $\mathbf{r}_1$  and  $\mathbf{r}_2$  vectors are colinear. Thus, the person at the centre of sphere 1 finds that the acceleration of the big dot, *as measured by an observer moving in the (noninertial) centre of sphere 2 frame* is the sphere 1 acceleration of  $-4\pi G\rho\mathbf{r}_1/3$  minus the acceleration of the observer at the centre of sphere 2:

$$-\frac{4\pi G\rho\mathbf{r}_1}{3} - \frac{-4\pi G\rho(\mathbf{r}_1 - \mathbf{r}_2)}{3} = -\frac{4\pi G\rho\mathbf{r}_2}{3} \quad (456)$$

Lo and behold, this is the acceleration that the observer at the centre of sphere 2 would find self-consistently in the privacy of his own study, without worrying about what anyone else thinks might be going on. A Euclidian, “linearly accelerating” universe is therefore perfectly self-consistent. A dynamically active, expanding universe is required. The expansion process itself is essentially Newtonian, not, as originally thought at the time of its discovery, a mysterious effect of general relativity. Naively, the rate of expansion naively ought to be slowing, since this is what gravity does: an object thrown from the surface of the earth slows down as its distance from the surface increases. This is what astronomer’s assumed for 70 years. As we shall soon see however, our Universe is a bit more devious than that. There is still some mystery here beyond the realm of the purely Newtonian!

### 8.1.2 The dynamical equation of motion

A simple way to describe the internal acceleration of the Universe is to begin with the spatially homogeneous but time-dependent relative expansion between two locations. The separation between two arbitrary points separated by a distance  $r(t)$  may be written

$$r(t) = R(t)l \quad (457)$$

where  $l$  is a comoving coordinate that labels a fixed radial distance from us in the space—fixed in the sense of being fixed to the expanding space, like latitude and longitude would be on the surface of an inflating globe. If we take  $l$  to have dimensions of length, then  $R(t)$  is a dimensionless function of time alone. It is a scale factor that embodies the dynamical behaviour of the Universe. The velocity  $v = dr/dt \equiv \dot{r}$  of a “fixed” point expanding with space is then

$$v(t) = \dot{R}l = (\dot{R}/R)r. \quad (458)$$

We should emphasise the vector character of this relationship:

$$\mathbf{v}(t) = (\dot{R}/R)\mathbf{r} \quad (459)$$

where  $\mathbf{r}$  is a vector pointing outward from our arbitrarily chosen origin. Then, the acceleration is

$$\mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = (\ddot{R}/R)\mathbf{r}. \quad (460)$$

(Why didn’t we differentiate  $\mathbf{r}(t)/R(t)$ ?) But we already *know* the relative acceleration between two points, because we know Newtonian physics. Equation [455] tells us that the gravitational acceleration is

$$\mathbf{a}(t) = -\frac{4\pi G\rho}{3}\mathbf{r}. \quad (461)$$

Hence,

$$-\frac{4\pi G\rho}{3} = \frac{\ddot{R}}{R}. \quad (462)$$

Notice how  $l$  disappears: this is an equation for the scale factor  $R$ , and the scale factor depends not on where you are, but on *when* you are. Next, multiply this equation by  $\dot{R}R$  and integrate, assuming that ordinary mass is conserved in the ordinary way, i.e.  $\rho R^3$  is constant. We then obtain

$$\dot{R}^2 - \frac{8\pi G \rho R^2}{3} = 2E, \quad (463)$$

where  $E$  is an energy-like integration constant. This, in a simple, apparently naive Euclidian-Newtonian approach, would be our fundamental dynamical cosmological equation for the evolution of the Universe. Amazingly, providing that we are prepared to allow the mass density  $\rho$  is to be upgraded to an energy density (divided by  $c^2$ ) that includes *all* contributions (in particular radiation and vacuum energy), this innocent little equation turns out to be far more general: it is *exactly* correct in full relativity theory! More on this anon.

### 8.1.3 Cosmological redshift

The expansion of the Universe leads to a very important kinematic effect known as the cosmological redshift. Since the Universe is expanding, a travelling photon is constantly overtaking sources that are moving away from it. If a photon has a wavelength  $\lambda$  at some location  $r$ , when the photon passes an observer a distance  $dr = cdt$  away, moving at a relative velocity  $\dot{R}dr/R$ , this observer measures a Doppler change in wavelength  $d\lambda$  determined by equation (459):

$$\frac{d\lambda}{\lambda} = \frac{v}{c} = \frac{\dot{R}}{cR}dr = \frac{\dot{R}}{R}dt, \quad (464)$$

or in other words

$$\frac{1}{\lambda} \frac{d\lambda}{dt} \equiv \frac{\dot{\lambda}}{\lambda} = \frac{\dot{R}}{R}. \quad (465)$$

Solving for  $\lambda$ , we find that it is linearly proportional to  $R$ . It is as though the wavelength stretches out with the rest of the Universe. This is a very general kinematic result, a property of any cosmological model that is symmetrically expanding. (Even if there is spatial curvature, the “differential proper distance” is always  $cdt$  for a travelling photon.)

This result means that the frequency of a photon goes *down* as  $1/R$ . But the frequency of a photon is proportional to its energy, hence the radiation temperature  $T_\gamma \propto 1/R$ , and the energy density in radiation (fourth power in  $T_\gamma$ ) decreases as  $1/R^4$ . The total energy in a volume  $R^3$  thus decreases as  $1/R$ ; energy is *not* conserved in the expansion of the Universe! By contrast, the *entropy* density in the radiation is proportional to  $T_\gamma^3$ , and  $(T_\gamma R)^3$  is therefore constant. The radiation entropy, which is just proportional to the number of photons, is conserved. Photons are neither created nor destroyed by the act of expansion. This makes perfect sense.

We are free, and it is customary, to choose our coordinates in such a way that the current value of  $R$  is 1, with  $R$  becoming smaller and smaller as we go back in time. If a photon is emitted with a wavelength  $\lambda_e$  at some time  $t$  in the past, the wavelength ( $\lambda_0$ ) we would now measure at time  $t_0$  is formally expressed as

$$\lambda_0 = \lambda_e(1 + z) \quad (466)$$

where  $z$  is defined by this equation and known as the redshift parameter. Therefore,

$$\frac{\lambda_0}{\lambda_e} = 1 + z = \frac{R(t_0)}{R(t)} = \frac{1}{R(t)}. \quad (467)$$



The advantage to using  $z$ , as opposed to the more geometrical quantity  $R$ , is that  $z$  is directly observed by astronomers. But the two quantities are mathematically equivalent via this completely general equation,  $R = 1/(1 + z)$ . In any symmetric cosmological model, if you measure a redshift of 2, it has come from a time when the Universe had one-third of its current size.

## 8.2 Cosmology models for the impatient

*Summary.* The concept of a spacetime metric is essential, even in Newtonian cosmology. A metric allows us to describe the paths of photons as they move through an expanding space, and to determine from which comoving coordinate location  $l$  they came for a given measured redshift  $z$ . Because different cosmologies produce different functions for  $l(z)$ , this quantity is a key observational tool which in principle allows us to determine what kind of Universe we live in.

### 8.2.1 The large-scale spacetime metric

Because the Newtonian approach works remarkably well, here is a brief reminder as to why we really do need relativity in our study of the large scale structure of the Universe.

First, we require a Riemannian metric structure to ensure that the speed of light is a universal constant  $c$ , especially when traversing a dynamically evolving spacetime background. It is easy to see what the form of this metric must be in a simple model of an expanding Euclidian Universe. Symmetry demands that time must flow identically for all observers comoving with the universal expansion, and we can always choose time to be a linear function of the time coordinate. Space is uniformly expanding at the same rate everywhere. So if space itself is Euclidian, the spacetime metric practically leaps out of the page,

$$-c^2 d\tau^2 = -c^2 dt^2 + R^2(dx^2 + dy^2 + dz^2) \quad (468)$$

where we have used the usual  $(x, y, z)$  Cartesian coordinates, and  $R$  satisfies equation (463). Note that  $x, y, z$  are *comoving* with the expansion, in essence Cartesian coordinate of length  $l$ ,

$$dx^2 + dy^2 + dz^2 = dl^2.$$

In particular, for a photon heading directly towards us, along a line of sight to a distant source,

$$R \frac{dl}{dt} = -c \quad (469)$$

where  $dl$  should be thought of as the change of radial comoving coordinate induced by the photon's passage. A mechanical analogue is an ant, moving at a constant velocity  $c$ , crawling along the surface of an expanding sphere from the pole to the equator. In this case, think of  $dl$  as the change in the ant's latitude.

Second, we must allow for the observational possibility of non-Euclidian spatial structure. Now equation (468) really does appear to be the true form of the spacetime metric for our Universe. Space happens to be very nearly, or perhaps even precisely, Euclidian. But as a purely mathematical point, this need not be the case even if we demand perfect symmetry, any more than the requirement of a perfectly symmetric two-dimensional surface demands a Euclidian plane. We could preserve our global *maximal spatial symmetry* and be in a curved space, like the surface of sphere. This surface, in common with a flat plane, is symmetric about every point, but its spatial properties are clearly distorted relative to a plane. (E.g.

the Pythagorean theorem does not hold, and the sum of the angles of a triangle exceeds 180 degrees.) The case of a two-dimensional spherical surface is readily grasped because we can easily embed it in three dimensions to form a mental image. The surface is finite in area and is said to be *positively* curved. There is also a perfectly viable, flaring, *negatively* curved two-dimensional surface. A saddle begins to capture its essence, but not quite, because the curvature of a saddle is not uniform. The case of a uniformly negatively curved two-dimensional surface cannot actually be embedded in only three dimensions, so it is very hard to picture in your mind's eye! (At least it is for me.) There are positively and negatively curved fully *three-dimensional* spaces as well, all perfectly symmetric, which are logically possible alternative structures for the space of our Universe. They just happen not to fit the data. It is fortunate for us that the real Universe seems also to be mathematically the simplest. For now we confine our attention to this expanding, good old “flat” Euclidian space.

Third, we need relativity in the form of the Birkhoff theorem to justify properly the argument neglecting exterior contributions from outside the arbitrarily chosen spheres we used in section 8.1. The Newtonian description strictly can't be applied to an infinite system, whereas nothing prevents us from using Birkhoff's theorem applied to an unbounded symmetric spacetime.

Fourth, we need a relativistically valid argument to arrive at equation (463). Nothing in the Newtonian derivation even hints at this level of generality. We shall return to this carefully in section 8.3.

Finally, we need relativity theory to relate the energy constant  $E$  to the geometry of our space. For now, we restrict ourselves to the case  $E = 0$ , which will turn out to be the only solution consistent with the adoption of a flat Euclidian spatial geometry, our own Universe.

### 8.2.2 The Einstein-de Sitter universe: a useful toy model

Consider equation (463) for the case  $E = 0$  in the presence of ordinary matter, which means that  $\rho R^3$  is a constant. Remember that we are free to choose coordinates in which  $R = R_0 = 1$  at the present time  $t = t_0$ . We may then choose the constant  $\rho R^3$  to be equal to  $\rho(t_0) = \rho_{M0}$ , the mass density at the current epoch. Equation (463) becomes

$$R^{1/2} \dot{R} = \left( \frac{8\pi G \rho_{M0}}{3} \right)^{1/2}. \quad (470)$$

Integrating,

$$\frac{2}{3} R^{3/2} = \left( \frac{8\pi G \rho_{M0}}{3} \right)^{1/2} t, \quad (471)$$

where the integration constant has been set to zero under the assumption that  $R$  was very small at early times. This leads to

$$R = \left( \frac{3H_0 t}{2} \right)^{2/3}, \quad (472)$$

where we have introduced an important structural constant of the Universe,  $H_0$ , the value of  $\dot{R}/R$  at the current time  $t_0$ . (The current growth rate.) This is known as the Hubble constant,

$$H_0 = \dot{R}(t_0) = \left( \frac{8\pi G \rho_{M0}}{3} \right)^{1/2} \rightarrow \dot{R} = H_0 R^{-1/2}. \quad (473)$$

More generally, the Hubble parameter is defined by

$$H(t) = \frac{\dot{R}}{R} \quad (474)$$

for any time  $t$ . The solution (472) of the scale factor  $R$  is known, for historical reasons, as the Einstein-de Sitter model.

*Exercise.* Show that  $H(t) = H_0(1+z)^{3/2}$  for our simple model  $R = (t/t_0)^{2/3}$ .

The Hubble constant is in principle something that we may observe directly, “simply” by measuring the distances to nearby galaxies as well as their redshift, and then using equation (459) locally. As a problem in astronomy, this is hardly simple! On the contrary, it is a very difficult task because we cannot be certain that the motion of these nearby galaxies comes from cosmological expansion alone. But astronomers have persevered, and the bottom line is that *the measured value of  $H_0$  and the measured value of the density of ordinary matter  $\rho_{M0}$  do not satisfy (473) in our Universe.* There is just not enough ordinary matter  $\rho_{M0}$  to account for the measured  $H_0$ . On the other hand, equation (463) does seem to be precisely valid, with  $E = 0$ . As the energy density of radiation in the contemporary Universe is much less than  $3H_0^2/8\pi G$ , how is all this possible?

The answer is stunning. While the energy density in the Universe of *ordinary* matter most certainly dominates over radiation, there is now strong evidence of an additional energy density associated with the vacuum of spacetime itself! This energy density,  $\rho_V$ , is the dominant energy density of the real Universe on cosmological scales, though not at present overwhelmingly so:  $\rho_V$  is about 69% of the total energy budget whereas matter (the familiar baryons plus whatever “dark matter” is) comes in at about 31%. However, because  $\rho_V$  remains constant as the Universe expands, at later times vacuum energy will dominate the expansion:  $\rho_M$  drops off as  $1/R^3$ , and the constant  $\rho_V$  takes up the slack.

What does this imply for the expansion behaviour of the Universe? With an effective vacuum Hubble parameter of

$$H_V \equiv \left( \frac{8\pi G \rho_V}{3} \right)^{1/2}, \quad (475)$$

equation (463) at later times takes the form

$$\dot{R} = H_V R \quad (476)$$

or

$$R \propto \exp(H_V t). \quad (477)$$

The Universe will expand exponentially! In other words, rather than gravity slowly decelerating the expansion by the Newtonian mutual attractive force, the vacuum energy density will actively drive an ever more vigorous repulsive force. The converse of this is that Universe was expanding more slowly in the past than in the present. It is this particular discovery which has led to our current understanding of the remarkable expansion dynamics of the Universe. The Nobel Prize in Physics was awarded to Perlmutter, Schmidt and Riess in 2011 for their use of distant supernovae as a tool for unravelling the dynamics of the Universe from early to modern times. We currently live in the epoch where exponential expansion is taking over.

Equations (467), (469) and (473) may be combined to answer the following question. If we measure a photon of redshift  $z$ , from what value of  $l$  did it originate? This is an important question because it provides the link between astronomical observations and the geometry of the Universe. With  $R(t) = (t/t_0)^{2/3} = (3H_0 t/2)^{2/3}$ , equation (469) may be integrated

over the path of the photon from its emission at  $l$  at time  $t$ :

$$\int_0^l dl' = -c \int_{t_0}^t \frac{dt'}{R} = c \int_R^1 \frac{dR}{\dot{R}R} = \frac{2c}{H_0} (1 - \sqrt{R}) = \frac{2c}{H_0} \left(1 - \frac{1}{\sqrt{1+z}}\right) \equiv l(z) \quad (478)$$

Note that as  $z \rightarrow \infty$ ,  $l \rightarrow 2c/H_0$ , a constant. The most distant photons—and therefore the most distant regions of the Universe that may causally influence us—come from  $l = 2c/H_0$ . This quantity is known as the *horizon*, denoted  $l_H$ . Beyond the horizon, we can detect—and be influenced by—nothing. This particular value of  $l_H = 2c/H_0$  is associated with the Einstein–de Sitter model, but the existence of a horizon is a general feature of many cosmologies and will generally be a multiple of  $c/H_0$ . Do not confuse this cosmological horizon, sometimes called a “particle horizon” with the *event horizon* of a black hole. The particle horizon is notionally outward, the cosmological radial extent over which causality may be exerted. The event horizon is notionally inward, the radius of a surface enveloping a blackhole, within which no causal contact with the outside world is possible.

Physical distances are given not by  $l$ , but by  $Rl$ . Since  $R = 1$  currently,  $2c/H_0$  is the current physical scale of the horizon as well. There is nothing special about *now*, however. We could be doing this analysis at any time  $t$ , and the physical size of the horizon at time  $t$  in the E-de S model would then be  $2c/H(t) = 2c/[H_0(1+z)^{3/2}]$ . So here is another interesting question. The furthest back in time that we can see is to a redshift of about  $z = 1500$ . At higher redshifts than this, the Universe was completely opaque to photons. Just as in viewing the Sun we can only see to its opaque photosphere, not into its interior, we can only see back in cosmological time to a time when the Universe itself became similarly opaque. What would be the subtended *angular* size of the horizon  $\Theta_H$  at  $z = 1500$ , as we measure it today on the sky? This is an important question because we would not expect the Universe to be very smooth or regular or correlated in any way on angular scales bigger than this. (The E-dS model doesn’t actually hold during the radiation dominated phase, but it will serve to make our point; greater numerical accuracy doesn’t help the problem!) The angular size of the horizon at a redshift  $z$  is given by the following expression:

$$\Theta_H(z) = \frac{2c}{H_0(1+z)^{3/2}} \times \frac{1}{R(z)l(z)} = \frac{2c}{H_0(1+z)^{1/2}l(z)} = \frac{1}{\sqrt{1+z}-1} \quad (479)$$

The first equality sets the horizon angle equal to the actual physical size of the horizon at redshift  $z$ , divided by  $d = R(z)l(z)$ , the distance to redshift  $z$  *at the earlier time corresponding to redshift  $z$*  (not now!). This  $d$  is the relevant distance to the photon sources at the moment of the radiation emission. The Universe was a smaller place then, and we cannot, we must not, use the *current* proper distance, which would be  $R_0l(z) = l(z)$ . If you plug in  $z = 1500$  into (479) and convert  $\Theta_H$  to degrees, you’ll find  $\Theta_H = 1.5^\circ$ , about 3 times the diameter of the full moon. But the Universe looks very regular on much, much larger angular scales, indeed it is regular over the entire sky. Even if our model is only crude, it highlights an important problem. Simply put, how does the Universe know about itself in a global sense, given that it takes signals, even signals travelling at the speed of light, so long to cross it? What we have here, ladies and gentleman, is a failure to communicate. We will see later in this course how modern cosmology can address this puzzle.

By the way, the fact that the Universe was “on top of us” at early times has another surprising consequence. Assuming that the average physical size of a galaxy isn’t changing very much with time, if we calculate the average *angular* size of a galaxy, we find that at low redshift, all is normal: the more distant galaxies appear to be smaller. But then, at progressively higher redshifts, the galaxies appear to be growing larger on the sky! Why? Because at large  $z$ , the Universe was, well, on top of us. The “distance-to-redshift- $z$ ” formula is (for E-dS models):

$$l(z)R(z) = \frac{2c}{H_0(1+z)^{3/2}} (\sqrt{1+z} - 1) \quad (480)$$

At low  $z$ , this is increasing linearly with  $z$ , which is intuitive: bigger redshift, more distant. But at large  $z$  this *declines* as  $1/z$ , since the Universe really was a smaller place.

*Exercise.* At what redshift would the average galaxy appear to be smallest in this model? (Answer:  $z = 5/4$ .)

### 8.3 The Friedman-Robertson-Walker Metric

*Summary.* The form of the metric for an expanding space time that is homogeneous and isotropic about every point, but allows for curvature to be present, is known as the Friedman-Robertson-Walker (FRW) metric. A uniformly expanding Euclidian space is a particular instance of the FRW metric, one that seems to describe our Universe. The form of the FRW can be found with simple mathematics, from the notion of a curved spherical surface in multidimensional Euclidian spaces. The FRW is required to establish the relationship between the redshift of a photon and its comoving coordinate of origin.

We begin with an important notational shift. The variable  $r$  will henceforth denote a comoving radial coordinate. To avoid any confusion, we will no longer use  $l$  as a comoving coordinate.

The Friedman-Robertson-Walker, or FRW, metric is the most general metric of a homogeneous isotropic cosmology. To understand its mathematical form, let us start simply. The ordinary metric for a planar two-dimensional space (“2-space”) without curvature may be written in cylindrical coordinates as

$$ds^2 = d\varpi^2 + \varpi^2 d\phi^2 \quad (481)$$

where the radial  $\varpi$  and angular  $\phi$  polar coordinates are related to ordinary Cartesian  $x$  and  $y$  coordinates by the familiar formulae:

$$x = \varpi \cos \phi, \quad y = \varpi \sin \phi \quad (482)$$

As we have noted, this flat 2-space is not the most general globally symmetric 2-space possible. The space could, for example, be distorted like the 2-surface of a sphere, yet retain the symmetry of every point being mathematically equivalent. The metric for a spherical surface is

$$ds^2 = a^2 d\theta^2 + a^2 \sin^2 \theta d\phi^2, \quad (483)$$

where  $a$  is the radius of the sphere. We know how to relate cylindrical polar and spherical coordinates: set  $\varpi = a \sin \theta$ . This may be viewed as a purely formal transformation of coordinates, but in our mind’s eye we can picture  $\theta$  in simple geometrical terms as the colatitude angle measured downward from the  $z$  axis, so that this  $\varpi$  is actually the same cylindrical radius  $\varpi$  seen in equation (481). But don’t expect to get back to (481) from (483) via this simple coordinate change  $\sin \theta = \varpi/a$ . Instead, we find that the spherical surface metric becomes:

$$ds^2 = \frac{d\varpi^2}{1 - \varpi^2/a^2} + \varpi^2 d\phi^2, \quad (484)$$

an entirely different space from the flat planar surface (481). The space of (484) is the same original spherical surface space we started with by any other name; the metric resembles equation (481) for a plane only in the limit  $a \rightarrow \infty$ . Changing coordinates does not change the geometry. In particular, a coordinate change cannot alter any curvature scalar, which remains invariant to coordinate transformation.

Let's stretch ourselves a bit and consider at a formal level the closely related two-dimensional metric

$$ds^2 = \frac{d\varpi^2}{1 + \varpi^2/a^2} + \varpi^2 d\phi^2, \quad (485)$$

which upon substituting  $\sinh \chi = \varpi/a$  reverts to

$$ds^2 = a^2 d\chi^2 + a^2 \sinh^2 \chi d\phi^2 \quad (486)$$

the fundamental symmetry properties of the metric are unaffected by the  $\pm\varpi^2/a^2$  sign flip. The flip in denominator sign simply changes the sign of the constant curvature from positive (convex, think sphere) to negative (flaring, think saddle). This characteristic form of the radial metric tensor component,  $g_{\varpi\varpi} = 1/(1 \pm \varpi^2/a^2)$ , will reappear when we go from curved 2-space to curved 3-space.

### 8.3.1 Maximally symmetric 3-spaces

It is best to begin with the conclusion, which ought not to surprise you. The most general form of the three dimensional metric tensor that is maximally symmetric—homogeneous and isotropic about every point—takes the form

$$-c^2 d\tau^2 = -c^2 dt^2 + R^2(t) \left( \frac{dr^2}{1 - r^2/a^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right) \quad (487)$$

where we allow ourselves the liberty of taking the curvature constant  $a^2$  to be either positive or negative.

The derivation of the metric for the three-dimensional hypersurface of a four-dimensional sphere is not difficult. The surface of such a “4-sphere” (Cartesian coordinates  $w, x, y, z$ ) is given by

$$w^2 + x^2 + y^2 + z^2 = a^2 = \text{constant}. \quad (488)$$

Thus, on the three-dimensional surface (“3-surface”), a small change in  $w^2$  is restricted to satisfy:

$$d(w^2) = -d(x^2 + y^2 + z^2) \equiv -d(r^2), \text{ whence } dw = -r dr/w \quad (489)$$

with  $r^2 \equiv x^2 + y^2 + z^2$ . Therefore,

$$(dw)^2 = \frac{r^2(dr)^2}{w^2} = \frac{r^2(dr)^2}{a^2 - r^2} \quad (490)$$

The line element in Cartesian 4-space is

$$ds^2 = dx^2 + dy^2 + dz^2 + dw^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 + dw^2. \quad (491)$$

Adding together  $dr^2$  and  $(dw)^2$  as given by (490), there follows immediately the desired line element restricted to the 3-surface of a 4-sphere, analogous to our expression for the ordinary spherical surface metric (483):

$$ds^2 = \frac{dr^2}{1 - r^2/a^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2, \quad (492)$$



just as we see in equation (487). An alternative form for (492) is sometimes useful without the singular denominator. Set  $r = a \sin \chi$ . Then

$$ds^2 = a^2 d\chi^2 + a^2 \sin^2 \chi d\theta^2 + a^2 \sin^2 \chi \sin^2 \theta d\phi^2 \quad (493)$$

*Exercise.* Show that the 3-surface of a 4-sphere with radius  $a$  has a “volume”  $\pi^2 a^3$ .

*Exercise.* Would you care to hazard a guess as to what the line element of the 4-surface of a 5-sphere looks like either in the form of (492) or (493)?

If for the moment we restrict ourselves to thinking of  $a^2$  as positive, then the corresponding negatively curved (saddle-like) 3-surface has line elements

$$ds^2 = \frac{dr^2}{1 + r^2/a^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2, \quad (494)$$

and, with  $r = a \sinh \chi$ , the alternative form is

$$ds^2 = a^2 d\chi^2 + a^2 \sinh^2 \chi d\theta^2 + a^2 \sinh^2 \chi \sin^2 \theta d\phi^2. \quad (495)$$

For our Universe,  $a$  appears to be immeasurably large (but with  $\chi$  small and  $a\chi$  finite). In other words, space is neither positively nor negatively curved. Space is basically flat.

It is customary in some textbooks to set  $a$  to be a positive constant, use  $r' = r/a$  as a rescaled radial variable, and then absorb the factor of  $a$  into the definition of what you mean by the scale factor  $R(t)$ . In effect, one introduces a new scale factor  $R' = aR$ . This eliminates  $a$  entirely from the metric! If you do that however, you give up on the convenience of setting the current value of  $R'$  equal to 1. Dropping the  $'$  on  $r'$  and  $R'$ , this general FRW metric, with  $a$  absorbed, is written:

$$-c^2 d\tau^2 = -c^2 dt^2 + R^2(t) \left( \frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right), \quad (496)$$

where the constant  $k$  is  $+1$  for a positively curved space,  $-1$  for a negatively curved space, and  $0$  for a flat space. This particular line element with

$$g_{00} = -1, \quad g_{rr} = R^2(t)/(1 - kr^2), \quad g_{\theta\theta} = R^2(t)r^2, \quad g_{\phi\phi} = R^2(t)r^2 \sin^2 \theta. \quad (497)$$

is often what people mean when they write “the FRW metric.” Because  $a$  is absent in this formulation, this can be mathematically convenient. In my view, the form (487) is more physically transparent.

Be careful! In (496),  $R$  has dimensions of length, and  $r$  is *dimensionless*. In (487),  $R$  is dimensionless and it is  $r$  that carries dimensions of length. For  $k = 0$ , we may stick to  $R$  being dimensionless and  $r$  carrying dimensions of length, as there is no need to “absorb  $a$ .” Remember that in all cases  $r$  is a comoving coordinate.

A few final words on practical uses of the FRW metric. I will use the metric in the form (487). In flat Euclidian space, the integrated comoving distance to coordinate  $r$  is just  $r$  itself. (The *proper* distance is then found by multiplying by the scale factor  $R(t)$  at the appropriate time.) Otherwise, the comoving radial distance associated with coordinate  $r$  is

$$\int_0^r \frac{dr'}{\sqrt{1 - r'^2/a^2}} = a \sin^{-1}(r/a) \quad (a^2 > 0), \quad \text{or} \quad a \sinh^{-1}(r/a) \quad (a^2 < 0).$$

(In either case above,  $a$  is regarded as a positive absolute value or norm.)

By way of a specific example, let us rework equation (478) in our current notation for the case of positive curvature. The comoving coordinate of a photon arriving with redshift  $z$  when the scale factor was  $R$  is given by

$$a \sin^{-1}(r/a) = c \int_t^{t_0} \frac{dt'}{R(t')} = c \int_R^1 \frac{dR}{\dot{R}R} = c \int_0^z \frac{dz}{(1+z)\dot{R}(z)} \equiv a \sin^{-1}(r[z]/a)$$

We've been a bit casual with notation to avoid a clutter, using  $R$  as dummy variable inside the integral and also as an integration limit. (The same with  $z$ .) But the meaning should be clear. Note that we have used  $R = 1/(1+z)$ , so that the current value of  $R$  is 1, which appropriate for the the form of the metric in equation (487).  $\dot{R}$  is given by equation (463) quite generally, but we need to relate  $E$  and  $a$ . *That* requires general relativity. In §8.4, we show that  $2E = -c^2/a^2$ .

Let's look at a pure vacuum  $E = 0$  universe, but one with finite  $\rho_V$  and  $H_V$  so that we have something to work with! Then  $\dot{R} = H_V R$  and

$$r = c \int_R^1 \frac{dR}{\dot{R}R} = \frac{c}{H_V} \int_R^1 \frac{dR}{R^2} = \frac{c}{H_V} \left( \frac{1}{R} - 1 \right) = \frac{cz}{H_V} \quad (498)$$

How interesting. There is no horizon in such a universe. As  $z \rightarrow \infty$ ,  $r \rightarrow \infty$ . An early phase of exponential growth ("inflation") has in fact been proposed as a means to avoid the horizon problem. We've learnt something very important already.

Next, the volume of photons in a maximally symmetric space within redshift  $z_m$  is formally given by

$$V = 4\pi \int_0^r R^3 \frac{r'^2 dr'}{\sqrt{1 - r'^2/a^2}}, \quad (499)$$

a function of  $r$  and thus ultimately of redshift. But what do we use for  $R$  inside this integral? That depends on the question. If we are interested in the *current* net volume of these sources, then  $R = R_0$ , a constant, and life is simple. If we are interested in the net volume of all the sources occupied *at the time of their emission*, then we would need first to find  $R$  as a function of  $r$  to do the integral, then we need to find  $r(z_m)$  to get back to redshift  $z$ . That can be complicated! As an easier example, however, the current net volume  $V$  for sources out to a maximum redshift  $z_m$  with  $a^2 > 0$  is ( $R = R_0 = 1$ )

$$V = 4\pi \int_0^r \frac{r'^2 dr'}{\sqrt{1 - r'^2/a^2}} = 2\pi \left( a^3 \sin^{-1} \frac{r}{a} - a^2 r \sqrt{1 - \frac{r^2}{a^2}} \right). \quad (500)$$

You should be able to show that this reduces to the more familiar  $4\pi r^3/3$  when  $a$  is large. In that case, for an E-dS universe the volume to redshift  $z$  is (show!):

$$V(z) = \frac{32\pi}{3} \left( \frac{c}{H_0} \right)^3 \left( 1 - \frac{1}{\sqrt{1+z}} \right)^3. \quad (501)$$

## 8.4 Large scale dynamics

*Summary.* The day of reckoning can be put off no longer, as the Field Equations of general relativity are confronted. We show that our Newtonian dynamical equation for  $\dot{R}$  is valid



under all circumstances, and that  $2E$  from this equation is equal to  $-c^2/a^2$  where  $a$  is from the FRW metric. We show that the Field Equations can be modified by the addition of what is known as the cosmological constant, a term whose effect is greatest at the largest scales. Remarkably, this is mathematically identical to the effect of a vacuum energy density, and we therefore keep the original form of the Field Equations and allow for the possibility that the energy density  $\rho c^2$  may contain a contribution from the vacuum.

The student should read and study subsections 8.4.1 and 8.4.3, but may regard the rather heavy 8.4.2 as optional (off syllabus). Remember, however, the identification of  $2E$  with  $-c^2/a^2$ .

#### 8.4.1 The effect of a cosmological constant

Begin first with the Field Equations *including the cosmological constant*, equation (250):

$$R_{\mu\nu} - \frac{g_{\mu\nu}}{2}R = -\frac{8\pi G}{c^4}T_{\mu\nu} + \Lambda g_{\mu\nu} \quad (502)$$

Recall the stress energy tensor for a perfect fluid:

$$T_{\mu\nu} = P g_{\mu\nu} + (\rho + P/c^2)U_\mu U_\nu. \quad (503)$$

We may arrange the right side source term of (502) as follows:

$$-\frac{8\pi G}{c^4}T_{\mu\nu} + \Lambda g_{\mu\nu} = -\frac{8\pi G}{c^4} \left[ \tilde{P} g_{\mu\nu} + \left( \tilde{\rho} + \frac{\tilde{P}}{c^2} \right) U_\mu U_\nu \right], \quad (504)$$

where

$$\tilde{P} = P - \frac{c^4 \Lambda}{8\pi G}, \quad \tilde{\rho} = \rho + \frac{c^2 \Lambda}{8\pi G} \quad (505)$$

In other words, the effect of a cosmological constant is to leave the left side of the Field Equations untouched and to leave the right side of the Field Equations in the form of a stress tensor for a perfect fluid, but with the density acquiring a constant additive term  $c^2 \Lambda / 8\pi G$  and the pressure acquiring a constant term of the opposite sign,  $-c^4 \Lambda / 8\pi G$ !

This is simple, almost trivial, mathematics, but profound physics. The effect of a cosmological constant is as if the vacuum itself had an energy density  $\mathcal{E}_V = \rho_V c^2 = c^4 \Lambda / 8\pi G$  and a pressure  $P_V = -c^4 \Lambda / 8\pi G$ . Does it make sense that the vacuum has a negative pressure, equal to its energy density but opposite in sign? Yes! If the vacuum volume expands by  $dV$ , the change in energy per unit volume of expansion is just  $dE/dV \equiv \rho_V c^2$ . By the first law of thermodynamics, this must be  $-P_V$ . Yet more fundamentally, if we recall the form of the stress energy tensor of the vacuum, *but without assuming*  $P_V = -\rho_V c^2$ , then

$$T_{\mu\nu(V)} = P_V g_{\mu\nu} + (\rho_V + P_V/c^2)U_\mu U_\nu, \quad (506)$$

and the last group of terms would change the form of the vacuum stress energy going from one constant velocity observer to another. In other words, you could tell if you were moving relative to the vacuum. That is absolutely forbidden! The vacuum stress must always be proportional to  $g_{\mu\nu}$ , and to  $g_{\mu\nu}$  alone. The only way this can occur is if  $P_V = -\rho_V c^2$ .

An early general relativity advocate, Sir Arthur Eddington was particularly partial to a cosmological constant, and was fond of commenting that setting  $\Lambda = 0$  would be to

“knock the bottom out of space.” At the time this was probably viewed as Eddington in his customary curmudgeon mode; today the insight seems downright prescient. Nowadays, physicists like to think less in terms of a cosmological constant and place more conceptual emphasis on the notion of a vacuum energy density. What is the reason for its existence? Why does it have the value that it does? If  $\rho_V$  is not strictly constant, general relativity would be wrong. Are the actual observational data supportive of a truly constant value for  $\rho_V$ ? The value of  $\rho_V$  probably emerged from the same type of “renormalisation process” (for those of you familiar with this concept) that has produced finite values for the masses for the fundamental particles. How do we calculate this? These are some of the most difficult questions in all of physics.

For present purposes, we put these fascinating issues to the side, and continue our development of large scale models of the Universe without the formal appearance of a cosmological constant, but with the understanding that we may add the appropriate contributions to the density  $\rho$  and pressure  $P$  to account precisely for the effects of  $\Lambda$ .

#### 8.4.2 Formal analysis

*Regard this subsection as OPTIONAL—off-syllabus.*

We shall use the Field Equations in terms of the source function  $S_{\mu\nu} = T_{\mu\nu} - g_{\mu\nu}T/2$ ,

$$R_{\mu\nu} = -\frac{8\pi G}{c^4}S_{\mu\nu}, \quad (507)$$

and the FRW metric in the form of equation (497), which is more convenient for this analysis. In our comoving coordinates, the only component of the covariant 4-velocity  $U_\mu$  that does not vanish is  $U_0 = -c$  (from the relations  $g^{00}(U^0)^2 = -c^2$  and  $U_0 = g_{00}U^0$ ). The nonvanishing components of  $T_{\mu\nu}$  are then

$$T_{00} = \rho c^2, \quad T_{ij} = P g_{ij}. \quad (508)$$

Which means

$$T_0^0 = -\rho c^2, \quad T_j^i = \delta_j^i P, \quad T = T_\mu^\mu = -\rho c^2 + 3P \quad (509)$$

We shall need

$$S_{00} = T_{00} - \frac{g_{00}}{2}T = \frac{1}{2}(\rho c^2 + 3P), \quad S_{ij} = \frac{1}{2}g_{ij}(\rho c^2 - P) \quad (510)$$

To calculate  $R_{00}$ , begin with our expression for the Ricci tensor, (258):

$$R_{\mu\kappa} = \frac{1}{2} \frac{\partial^2 \ln |g|}{\partial x^\kappa \partial x^\mu} - \frac{\partial \Gamma_{\mu\kappa}^\lambda}{\partial x^\lambda} + \Gamma_{\mu\lambda}^\eta \Gamma_{\kappa\eta}^\lambda - \frac{\Gamma_{\mu\kappa}^\eta}{2} \frac{\partial \ln |g|}{\partial x^\eta}$$

where  $g$  is the determinant of the metric tensor  $g_{\mu\nu}$  given by (497). Defining

$$f(r) = 1/(1 - kr^2),$$

we have

$$|g| = R^6(t) r^4 f(r) \sin^2 \theta \quad (511)$$

For diagonal metrics,  $\Gamma_{ab}^a = \partial_b(\ln g_{aa})/2$  and  $\Gamma_{aa}^b = -(\partial_b g_{aa})/2g_{bb}$  (no sum on  $a$ ). Therefore  $\Gamma_{00}^\lambda = 0$ , and our expression for  $R_{00}$  simplifies to

$$R_{00} = \frac{1}{2} \frac{\partial^2 \ln |g|}{\partial x^0 \partial x^0} + \Gamma_{0\lambda}^\eta \Gamma_{0\eta}^\lambda = \frac{1}{2} \frac{\partial^2 \ln |g|}{\partial x^0 \partial x^0} + \Gamma_{0r}^r \Gamma_{0r}^r + \Gamma_{0\theta}^\theta \Gamma_{0\theta}^\theta + \Gamma_{0\phi}^\phi \Gamma_{0\phi}^\phi \quad (512)$$

With

$$\frac{1}{2} \frac{\partial^2 \ln |g|}{\partial x^0 \partial x^0} = \frac{3\ddot{R}}{Rc^2} - \frac{3\dot{R}^2}{R^2 c^2}, \quad \Gamma_{0r}^r = \Gamma_{0\theta}^\theta = \Gamma_{0\phi}^\phi = \frac{\dot{R}}{Rc}, \quad (513)$$

the 00 component of (507) becomes

$$\ddot{R} = -\frac{4\pi GR}{3} \left( \rho + \frac{3P}{c^2} \right) \quad (514)$$

which differs from our Newtonian equation (462) only by an additional, apparently very small, term of  $3P/c^2$  as an effective source of gravitation. However, during the time when the Universe was dominated by radiation, this term was important, and even now it turns out to be not only important, but *negative* as well! During the so-called inflationary phase,  $3P/c^2$  was hugely important. We will have much more to say about all of this later.

The  $rr$  component of the Field Equations is a bit more involved. Ready? With  $f' \equiv df/dr$ , we prepare a working table in advance of all the results we will need:

$$R_{rr} = \frac{1}{2} \frac{\partial^2 \ln |g|}{\partial r^2} - \frac{\partial \Gamma_{rr}^\lambda}{\partial x^\lambda} + \Gamma_{r\lambda}^\eta \Gamma_{r\eta}^\lambda - \frac{1}{2} \Gamma_{rr}^\eta \frac{\partial \ln |g|}{\partial x^\eta} \quad (515)$$

$$\Gamma_{rr}^r = \frac{1}{2g_{rr}} \frac{\partial g_{rr}}{\partial r} = \frac{f'}{2f}, \quad \Gamma_{rr}^0 = -\frac{1}{2cg_{00}} \frac{\partial g_{rr}}{\partial t} = \frac{fR\dot{R}}{c}, \quad \Gamma_{0r}^r = \frac{1}{2cg_{rr}} \frac{\partial g_{rr}}{\partial t} = \frac{\dot{R}}{Rc} \quad (516)$$

$$\Gamma_{\phi r}^\phi = \frac{1}{2g_{\phi\phi}} \frac{\partial g_{\phi\phi}}{\partial r} = \Gamma_{\theta r}^\theta = \frac{1}{2g_{\theta\theta}} \frac{\partial g_{\theta\theta}}{\partial r} = \frac{1}{r} \quad (517)$$

$$\frac{\partial \ln |g|}{\partial r} = \frac{f'}{f} + \frac{4}{r}, \quad \frac{1}{2} \frac{\partial^2 \ln |g|}{\partial r^2} = \frac{1}{2} \frac{f''}{f} - \frac{1}{2} \frac{(f')^2}{f^2} - \frac{2}{r^2}, \quad \frac{\partial \Gamma_{rr}^0}{\partial x^0} = \frac{1}{c^2} (f\dot{R}^2 + fR\ddot{R}) \quad (518)$$

$$\frac{\partial \ln |g|}{\partial t} = \frac{6\dot{R}}{R}, \quad \frac{\partial \Gamma_{rr}^r}{\partial r} = \frac{f''}{2f} - \frac{(f')^2}{2f^2} \quad (519)$$

Putting it all together:

$$\begin{aligned} R_{rr} &= \frac{1}{2} \frac{\partial^2 \ln |g|}{\partial r^2} - \frac{\partial \Gamma_{rr}^r}{\partial r} - \frac{\partial \Gamma_{rr}^0}{\partial x^0} + (\Gamma_{rr}^r)^2 + (\Gamma_{\theta r}^\theta)^2 + (\Gamma_{\phi r}^\phi)^2 + 2\Gamma_{rr}^0 \Gamma_{r0}^r - \frac{\Gamma_{rr}^r}{2} \frac{\partial \ln |g|}{\partial r} - \frac{\Gamma_{rr}^0}{2c} \frac{\partial \ln |g|}{\partial t} \\ &= \frac{f''}{2f} - \frac{(f')^2}{2f^2} - \frac{2}{r^2} - \left[ \frac{f''}{2f} - \frac{(f')^2}{2f^2} \right] - \frac{1}{c^2} (f\dot{R}^2 + fR\ddot{R}) + \frac{(f')^2}{4f^2} + \frac{2}{r^2} + \frac{2f\dot{R}^2}{c^2} - \frac{f'}{4f} \left( \frac{f'}{f} + \frac{4}{r} \right) - \frac{3f\dot{R}^2}{c^2} \end{aligned}$$

Thus, with  $f'/f = 2kr$ ,

$$R_{rr} = -\frac{2f\dot{R}^2}{c^2} - \frac{fR\ddot{R}}{c^2} - 2kf = -\frac{8\pi G}{c^4} S_{rr} = -\frac{4\pi G}{c^4} g_{rr}(\rho c^2 - P) = -\frac{4\pi R^2 G f}{c^4}(\rho c^2 - P)$$

or

$$2\dot{R}^2 + R\ddot{R} + 2kc^2 = 4\pi GR^2(\rho - P/c^2) \quad (520)$$

Notice that  $r$  has disappeared, as it must! (Why must it?) Eliminating  $\ddot{R}$  from (520) via equation (514) and simplifying the result leads to

$$\dot{R}^2 - \frac{8\pi G \rho R^2}{3} = -kc^2 \quad (521)$$

This is exactly the Newtonian equation (463) with the constant  $2E$  “identified with”  $-kc^2$ . *But be careful.* The Newtonian version (463) was formulated with  $R$  dimensionless. In equation (521),  $R$  has been rescaled to have dimensions of length, and  $k$  is either 1, 0, or  $-1$ . To compare like-with-like we should repeat the calculation with the radial line element of the metric (487). Don’t panic: this just amounts to replacing  $k$  with  $1/a^2$  with dimensions of an inverse length squared. We then have  $2E = -c^2/a^2$ , carrying dimensions of  $\dot{R}^2$  or  $1/t^2$ . The important point is that equation (463) is valid in full general relativity! And, as we have just seen, the dynamical Newtonian energy constant may be identified with the geometrical curvature of the space.

One final item in our analysis. We have not yet made use of the equation for the conservation of energy-momentum, based on (177) and the Bianchi Identities:

$$T_{;\mu}^{\mu\nu} = \frac{1}{\sqrt{|g|}} \frac{\partial(\sqrt{|g|}T^{\mu\nu})}{\partial x^\mu} + \Gamma_{\mu\lambda}^\nu T^{\mu\lambda} = 0. \quad (522)$$

This constraint is of course already built into the Field Equations themselves, and so in this sense adds no new information to our problem. But we may ask whether use of this equation from the start might have saved us some labour in getting to (521): it was a long derivation after all. The answer is an interesting “yes” and “no.”

The  $\nu = 0$  component of (522) reads

$$T_{;\mu}^{\mu 0} = \frac{1}{\sqrt{|g|}} \frac{\partial(\sqrt{|g|}T^{00})}{\partial x^0} + \Gamma_{\mu\lambda}^0 T^{\mu\lambda} = 0. \quad (523)$$

Only affine connections of the form  $\Gamma_{ii}^0$  (spatial index  $i$ , no sum) are present, and with

$$\Gamma_{ii}^0 = -\frac{1}{2g_{00}} \frac{\partial g_{ii}}{\partial x^0} = \frac{1}{2} \frac{\partial g_{ii}}{\partial x^0}, \quad T^{00} = \rho c^2, \quad T^{ii} = P g^{ii} = \frac{P}{g_{ii}}, \quad (\text{NO } i \text{ SUMS})$$

equation (523) becomes:

$$\frac{1}{R^3} \frac{\partial(R^3 \rho c^2)}{\partial t} + 3P \frac{\dot{R}}{R} = 0 \rightarrow \dot{\rho} + 3 \left( \rho + \frac{P}{c^2} \right) \frac{\dot{R}}{R} = 0. \quad (524)$$

Notice how this embodies at once the law of conservation of mass (in the large  $c$  limit), and the first law of thermodynamics,  $dE = -PdV$ .

*Exercise.* Justify the last statement, and show that a pure vacuum energy density universe satisfies (524).

If we now use (524) to substitute for  $P$  in (514), after a little rearranging we easily arrive at the result

$$\frac{d}{dt} (\dot{R}^2) = \frac{8\pi G}{3} \frac{d}{dt} (\rho R^2) \quad (525)$$

which in turn immediately integrates to (521)! This is surely a much more efficient route to (521), except...except that we cannot relate the integration constant that emerges from (525) to the spatial curvature constant  $k$ . True, we have a faster route to our final equation, but with only dynamical information. Equation (521) is after all just a statement of energy conservation, as seen clearly from our Newtonian derivation. Without explicitly considering the Ricci  $R_{rr}$  component, we lose the geometrical connection between Newtonian  $E$  (just an integration constant) and  $-kc^2$ . We therefore have *consistency* between the energy conservation and Ricci approaches, but not true *equivalence*. A subtle and interesting distinction.

### 8.4.3 The $\Omega_0$ parameter

We work with the dynamical cosmological equation in its Newtonian form, with no loss of generality,

$$\dot{R}^2 - \frac{8\pi G \rho R^2}{3} = 2E = -\frac{c^2}{a^2}, \quad (526)$$

where  $E$  is an energy integration constant. Using observable quantities,  $E$  may be set by the convention that  $R = 1$  at the current time  $t_0$ . We write  $2E$  in terms of  $H_0$ , the Hubble constant  $\dot{R}_0$  and  $\rho_0 c^2$ , the current average value of the total energy density of the Universe:

$$H_0^2 \left( 1 - \frac{8\pi G \rho_0}{3H_0^2} \right) \equiv H_0^2 (1 - \Omega_0) = 2E = -\frac{c^2}{a^2}, \quad (527)$$

where the parameter  $\Omega_0$  is

$$\Omega_0 = \frac{8\pi G \rho_0}{3H_0^2}. \quad (528)$$

The Universe is therefore positively curved (closed) or negatively curved (open) according to whether the measured value of  $\Omega_0$  is larger or smaller than unity. Defining the critical mass density  $\rho_c$  by

$$\rho_c = \frac{3H_0^2}{8\pi G}, \quad (529)$$

and the critical *energy* density in the Universe is then  $\rho_c c^2$ . We have not yet assumed anything about the sources of  $\rho$ ; they can and do involve a vacuum energy density. The currently best measured value of  $H_0$  (in standard astronomical units) is  $67.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,<sup>12</sup> or  $H_0 \simeq 2.2 \times 10^{-18} \text{ s}^{-1}$ . This number implies a critical density of  $8.6 \times 10^{-27} \text{ kg}$  (about 5 hydrogen atoms) per cubic meter<sup>13</sup>. The best so-called “concordance models” all point to an  $\Omega_0 = 1$ ,  $E = 0$ , universe, but only if  $\sim 70\%$  of  $\rho_0 c^2$  comes from the vacuum! About 25% comes from dark matter, which is matter that is not luminous but whose presence is inferred from its gravitational effects, and just 5% comes from ordinary baryonic matter in the form of gas and stars. We will denote the current vacuum contribution to  $\Omega_0$  as  $\Omega_{V0}$  and the matter contribution as  $\Omega_{M0}$ . We have in addition a contribution from radiation,  $\Omega_{\gamma 0}$ , and while it is quite negligible now, in the early universe ( $z \geq 1500$ ) it was completely dominant, even over the vacuum component. The Universe went first through a radiation-dominated phase, followed by a matter-dominated phase, and then at about a redshift of 2, it started to switch to a vacuum-dominated phase. This latter transition is still ongoing.

We do live in interesting times.

## 8.5 The classic, matter-dominated universes

*Summary. Matter-dominated models of the Universe dominated cosmological thought throughout the 20th century. If the matter density is high, these universes all have a closed geometry*

<sup>12</sup>One “megaparsec,” or Mpc is one million parsecs. One parsec (“pc”) is the distance from the solar system at which the Earth–Sun semi-major axis would subtend an angle of one arcsecond:  $1 \text{ pc} = 3.0856 \times 10^{13} \text{ km}$ ,  $1 \text{ Mpc} = 3.0856 \times 10^{19} \text{ km}$ . Stars are typically separated by 1 pc in galaxies, galaxies from one another by about 1Mpc.

<sup>13</sup> The *pressure* in the vacuum is some  $5.4 \times 10^{-10} \text{ Pa}$  (for “Pascals”, the MKS pressure unit). This would be an ultra-vacuum in a terrestrial lab, but rather dense by astronomical standards: it is roughly the gas pressure in the core of a molecular cloud in the interstellar medium.

and recollapse after a finite time. By contrast, low density, open geometry models expand forever, the scale factor eventually increasing linearly with time. A unique critical density divides these two cases. In this unique case the geometry is flat (Euclidian) and the expansion rate is ever diminishing: the scale factor is proportional to  $t^{2/3}$  at all times  $t$ . This is the Einstein-de Sitter solution. All three types of solution may be studied analytically.

Let us return to the good old days, when the idea that the Universe was driven by “the energy density of the vacuum” was the stuff of *Star Trek* conventions, not something that serious-minded physicists dwelt upon. The expansion dynamics of the Universe was thought to be dominated by matter, pure and simple, after a brief early phase when it was radiation-dominated. Matter obeys the constraint that  $\rho(t)R^3(t)$  remains constant with time. With  $R_0 = 1$ , this constant must be the current mass density  $\rho_{M0}$ . The dynamical equation of the Universe may then be written

$$\dot{R}^2 = 2E + \frac{8\pi G\rho R^2}{3} = H_0^2 \left(1 - \Omega_{M0} + \frac{\Omega_{M0}}{R}\right) \quad (530)$$

or

$$\int_0^R \frac{dR'}{\sqrt{1 - \Omega_{M0} + (\Omega_{M0}/R')}} = \int_0^R \frac{R'^{1/2} dR'}{\sqrt{(1 - \Omega_{M0})R' + \Omega_{M0}}} = H_0 t \quad (531)$$

The nature of the integral depends upon whether  $\Omega_{M0}$  is less than, equal to, or greater than, 1. The case  $\Omega_{M0} = 1$  is trivial and leads immediately to

$$R = \left(\frac{3H_0 t}{2}\right)^{2/3} \quad (532)$$

our Einstein-de Sitter solution (472). According to (526) with  $E = 0$ , the average density is then given explicitly by

$$\rho = \frac{\dot{R}^2}{R^2} \frac{3}{8\pi G} = \frac{1}{6\pi G t^2} \quad (533)$$

Consider next the case  $\Omega_{M0} > 1$ , that of a Closed Universe. Then (531) may be written:

$$\int_0^R \frac{R'^{1/2} dR'}{\sqrt{1 - (1 - \Omega_{M0}^{-1})R'}} = \Omega_{M0}^{1/2} H_0 t \quad (534)$$

Set

$$(1 - \Omega_{M0}^{-1})R' = \sin^2 \theta' \quad (535)$$

Then (534) becomes

$$\frac{2}{(1 - \Omega_{M0}^{-1})^{3/2}} \int_0^\theta \sin^2 \theta' d\theta' = \frac{1}{(1 - \Omega_{M0}^{-1})^{3/2}} \int_0^\theta (1 - \cos 2\theta') d\theta' = \Omega_{M0}^{1/2} H_0 t \quad (536)$$

or

$$\frac{2\theta - \sin 2\theta}{2(1 - \Omega_{M0}^{-1})^{3/2}} = \Omega_{M0}^{1/2} H_0 t \quad (537)$$

With  $\eta = 2\theta$ , our final solution is in the form of a parameterisation:

$$R = \frac{1 - \cos \eta}{2(1 - \Omega_{M0}^{-1})}, \quad H_0 t = \frac{\eta - \sin \eta}{2\sqrt{\Omega_{M0}(1 - \Omega_{M0}^{-1})^3}} \quad (538)$$

Some readers may recognise these equations for  $R$  and  $H_0 t$  as the parameterisation of a cycloid, which is the path taken by a fixed point on the circumference of a wheel as the wheel rolls forward. (See the curve labelled “Closed” in figure [10].) Precisely analogous expressions emerge for the case  $\Omega_{M0} < 1$ , which are easily verified from (531):

$$R = \frac{\cosh \eta - 1}{2(\Omega_{M0}^{-1} - 1)}, \quad H_0 t = \frac{\sinh \eta - \eta}{2\sqrt{\Omega_{M0}(\Omega_{M0}^{-1} - 1)^3}} \quad (539)$$

This is the solution for negative curvature, an Open Universe.

*Exercise.* Show that for either equation (538) or (539),

$$\frac{d\eta}{dt} = \frac{c}{aR}$$

where  $a$  is interpreted as  $\sqrt{|\pm a^2|}$ . We will use this in the Problem Set!

*Exercise.* Show that the equations analogous to (532), (538) and (539) for the FRW metric in the form (496) are respectively

$$\begin{aligned} R &= \omega^{1/3} t^{2/3} \quad (\text{Einstein} - \text{deSitter}) \\ R &= \frac{\omega}{c^2} \left( \frac{1 - \cos \eta}{2} \right), \quad ct = \frac{\omega}{c^2} \left( \frac{\eta - \sin \eta}{2} \right) \quad (\text{closed}) \\ R &= \frac{\omega}{c^2} \left( \frac{\cosh \eta - 1}{2} \right), \quad ct = \frac{\omega}{c^2} \left( \frac{\sinh \eta - \eta}{2} \right) \quad (\text{open}) \end{aligned}$$

where  $\omega = 8\pi G \rho_{M0} R_0^3 / 3$ . Show that  $cdt = R d\eta$  for the last two cases.

*Exercise.* Show that as  $\eta \rightarrow 0$ , both (538) and (539) reduce to  $R = \Omega_0^{1/3} [3H_0 t / 2]^{2/3}$  (the same of course as [532] for  $\Omega_0 = 1$ ), and at late times (539) becomes the “coasting” solution,  $R = \sqrt{1 - \Omega_0} H_0 t$ . This means that a plot of all possible solutions of  $R(t)$  versus  $\Omega_0^{1/2} H_0 t$  would converge to exactly same solution at early times, regardless of  $\Omega_0$ . Figure (10) shows this behaviour quite clearly.

What if  $\Omega_0 = 0$ ? Show then that  $R = H_0 t$  for all times. Wait. Could a universe *really* be expanding if there is nothing in it? Expanding with respect to exactly what, please? See Problem Set.

## 8.6 Our Universe

### 8.6.1 Prologue

Throughout most the 20th century, the goal of cosmology was to figure out which of the three standard model scenarios actually holds: do we live in an open, closed, or critical universe? Solutions with a cosmological constant were relegated to the realm of disreputable speculation, perhaps the last small chapter of a textbook, under the rubric of “Alternative Cosmologies.” If you decided to sneak a look at this, you would be careful to lock your office door. All of that changed in 1998-9, when the results of two cosmological surveys of supernovae (Perlmutter et al. 1998 ApJ, **517**, 565; Riess et al. 1999 Astron J., 116, 1009) produced compelling evidence that the rate of expansion of the Universe is increasing with



time, and that the spatial geometry of the Universe was flat ( $k = 0$ ), even with what had seemed to be an under density of matter.

Though something of a shock at the time, for years there had been mounting evidence that something was seriously amiss with standard models. Without something to increase the rate of the Universe's expansion, the measured rather large value of  $H_0$  consistently gave an embarrassingly short lifetime for the Universe, less than the inferred ages of the oldest stars! (The current stellar record holder is HE 1523-0901, coming in at a spry 13.2 billion years.) People were aware that a cosmological constant could fix this, but to get an observationally reasonable balance between the energy density of ordinary matter and a vacuum energy density seemed like a desperate appeal to “it-just-so-happens” fine-tuning. But the new millenium brought with it unambiguous evidence that this is the way things are: our Universe is about 30% non-relativistic matter, 70% vacuum energy, and boasts a Euclidian spatial geometry. So right now *it just so happens* that there is a bit more than twice as much energy in the vacuum as there is in ordinary matter. Nobody has the foggiest idea why.

### 8.6.2 A Universe of ordinary matter and vacuum energy

It is perhaps some consolation that we can give a simple mathematical function for the scale factor  $R(t)$  of our Universe. With  $E = 0$  for a Euclidian space, (526) is

$$\dot{R} = \left( \frac{8\pi G \rho}{3} \right)^{1/2} R \quad (540)$$

The energy density  $\rho c^2$  is a combination of nonrelativistic matter  $\rho_M$ , for which  $\rho R^3$  is a constant, and a vacuum energy density  $\rho_V$  which remains constant. With  $\rho_{M0}$  the current value of  $\rho_M$  and  $R_0 = 1$ ,  $\rho_M = \rho_{M0}/R^3$  and therefore

$$\rho R^2 = \rho_V R^2 + \frac{\rho_{M0}}{R} \quad (541)$$

Substituting (541) into (540) leads to

$$\frac{R^{1/2} dR}{\sqrt{\rho_{M0} + \rho_V R^3}} = \sqrt{\frac{8\pi G}{3}} dt. \quad (542)$$

Integration then yields

$$\int \frac{R^{1/2} dR}{\sqrt{\rho_{M0} + \rho_V R^3}} = \frac{2}{3} \int \frac{d(R)^{3/2}}{\sqrt{\rho_{M0} + \rho_V R^3}} = \frac{2}{3} \frac{1}{\sqrt{\rho_V}} \sinh^{-1} \left( \sqrt{\frac{\rho_V}{\rho_{M0}}} R^{3/2} \right) = \sqrt{\frac{8\pi G}{3}} t \quad (543)$$

Equation (543) then tells us:

$$R^{3/2} = \sqrt{\frac{\rho_{M0}}{\rho_V}} \sinh \left( \frac{3}{2} \sqrt{\frac{8\pi G \rho_V}{3}} t \right). \quad (544)$$

In terms of the  $\Omega$  parameters, we have

$$\Omega_M = \frac{8\pi G \rho_{M0}}{3H_0^2}, \quad \Omega_V = \frac{8\pi G \rho_V}{3H_0^2}, \quad (545)$$



henceforth dropping the 0 subscript on the  $\Omega$ 's with the understanding that these parameters refer to current time. The dynamical equation of motion (540) at the present epoch tells us directly that

$$\Omega_M + \Omega_V = 1. \quad (546)$$

In terms of the observationally accessible quantity  $\Omega_M$ , the scale factor  $R$  becomes

$$R = \frac{1}{(\Omega_M^{-1} - 1)^{1/3}} \sinh^{2/3} \left( \frac{3}{2} \sqrt{1 - \Omega_M} H_0 t \right), \quad (547)$$

or, with  $\Omega_M = 0.31$ ,

$$R = 0.7659 \sinh^{2/3}(1.246 H_0 t). \quad (548)$$

This gives a current age of the Universe  $t_0$  of

$$H_0 t_0 = \frac{2 \sinh^{-1}(\sqrt{\Omega_M^{-1} - 1})}{3\sqrt{1 - \Omega_M}} = \frac{\sinh^{-1}[(0.7659)^{-3/2}]}{1.246} = 0.9553, \quad (549)$$

i.e.,  $t_0 = 1$  is rather close to  $1/H_0$ :  $t_0 \simeq 13.8$  billion years. (This calculation ignores the brief period of the Universe's history when it was radiation dominated.)

*Exercise.* Show that the redshift  $z$  is related to the time  $t$  since the big bang by

$$z = 1.306 \sinh^{-2/3} \left( \frac{1.19t}{t_0} \right) - 1$$

If a civilisation develops 5 billion years after the big bang and we detect their signals(!), at what redshift would they be coming from?

## 8.7 Observational foundations of cosmology

### 8.7.1 The first detection of cosmological redshifts

To me, the name of Vesto Melvin Slipher has always conjured up images of some 1930's J. Edgar Hoover FBI G-man in a fedora who went after the bad guys. But Vesto was a mild-mannered, careful astronomer. If we allow that Edwin Hubble was the father of modern observational cosmology, then Slipher deserves the title of grandfather.

In 1912, using a 24-inch reflecting telescope, Slipher was the first person to measure the redshifts of external galaxies. He didn't know that that was what he had done, because the notion of galaxies external to our own was not one that was well-formed at the time. Nebular spectroscopy was hard, tedious work, spreading out the light from the already very faint, low surface brightness smudges of spiral nebulae through highly dispersive prisms. Slipher worked at Lowell Observatory, a small, isolated, private outpost in Flagstaff, Arizona. Percival Lowell, Slipher's boss, was at the time obsessed with mapping what he thought were the Martian canals!

Slipher toiled away, far removed from the great centres of astronomical activity. By 1922, he had accumulated 41 spectra of spiral nebulae, of which 36 showed a shift toward the red end of the spectrum. But he had no way of organising these data to bring out the linear scaling of the redshift with distance for a very simple reason: he hadn't any idea what the distances to the nebulae were. Other observations were at this time showing an

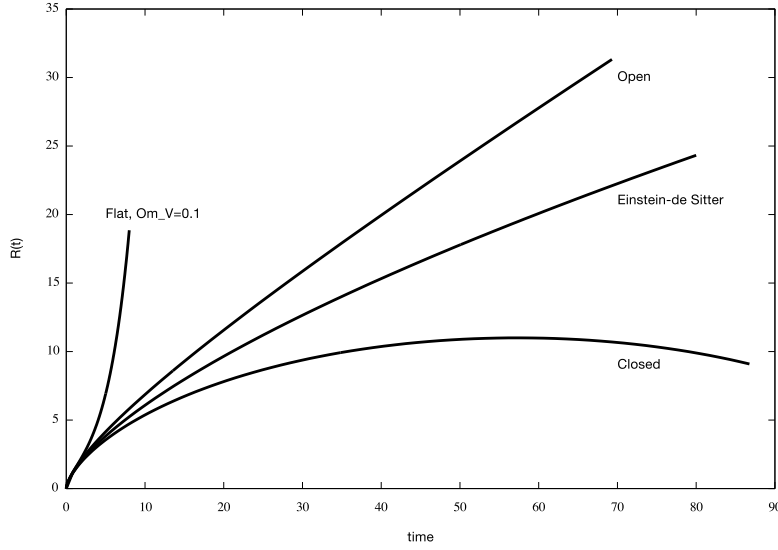


Figure 10:  $R(t)$  versus time, units of  $\Omega_M^{-1/2} H_0^{-1}$ , for four model universes. ‘Closed’ is eq. (538); ‘Einstein- de Sitter’ eq. (532); ‘Open’ eq. (539). ‘Closed’ has  $\Omega_M = 1.1$ , for ‘Einstein-de Sitter’  $\Omega_M = 1$ , ‘Open’ has  $\Omega_M = 0.9$ . The curve labelled ‘Flat,  $\text{Om}_V = 0.1$ ’, eq. (547), includes a cosmological constant with a vacuum contribution of  $\Omega_V = 0.1$ , so that  $\Omega_M + \Omega_V = 1$  and the spatial geometry is Euclidian.

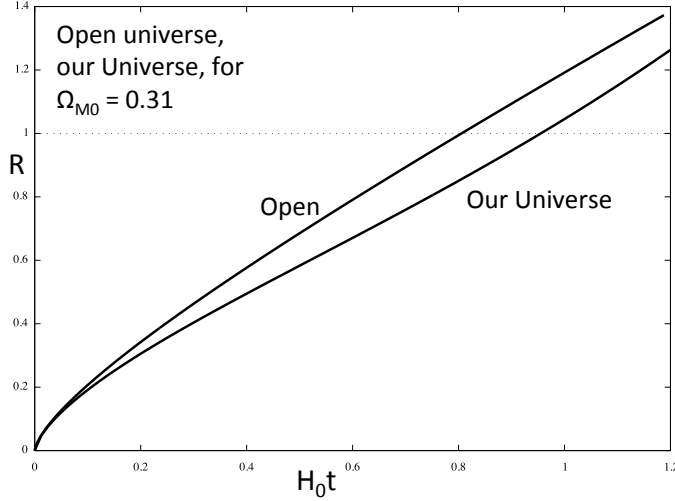


Figure 11: Comparison of  $R(t)$  versus  $\Omega_M^{1/2} H_0 t$  for our Universe eq. (547) and an open universe eq. (539), both with  $\Omega_M = 0.31$ . Crossing of the dotted line  $R = 1$  (at the current age) occurs earlier for the positive energy open universe, which is about 12% younger than our Universe. This would be too young for the oldest stars.

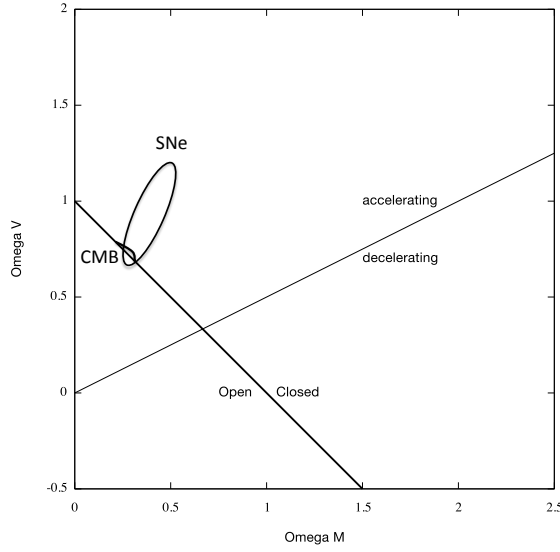


Figure 12: Parameter plane of  $\Omega_V$  versus  $\Omega_M$  assuming no radiation contribution. Regions of open/closed geometry and currently accelerating/decelerating dynamics are shown. Also shown are approximate zones of one standard deviation uncertainties for the distant supernova data (SNe) and —you have to squint—for fluctuations in the cosmic microwave background radiation (CMB), which came a decade later. Note the powerful constraint imposed by the latter: we no longer depend on the SNe data. That the Universe is accelerating is beyond reasonable doubt.

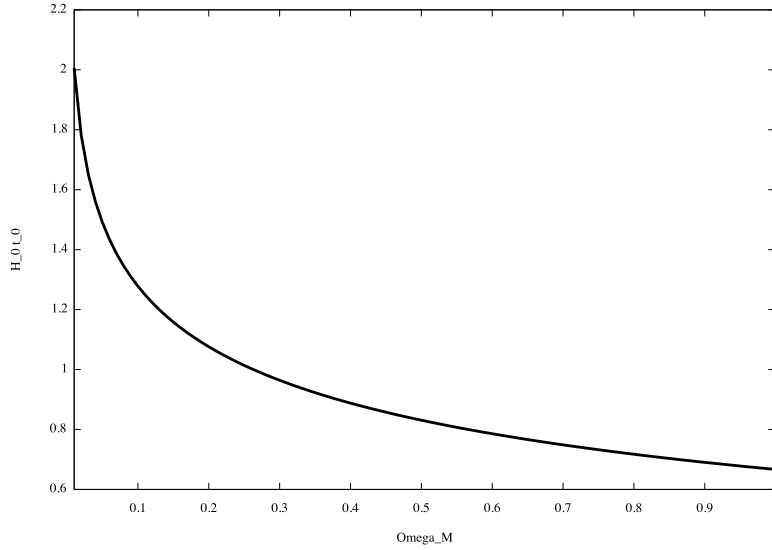


Figure 13: The current age of the Universe  $H_0 t_0$  as a function of  $\Omega_M$  for flat, matter plus vacuum energy models with negligible radiation. As  $\Omega_M$  approaches unity, the model recovers the Einstein–de Sitter value  $H_0 t_0 = 2/3$ ; as  $\Omega_M \rightarrow 0$ ,  $H_0 t_0$  becomes proportional to  $-\ln \Omega_M$  (Show!), and we recover a “logarithmic eternity,” first highlighted by Sir Arthur Eddington.

apparent trend of greater redshift with fainter nebulae, but the decisive step was taken by Edwin Hubble. With the aid of Milton Humason, Hubble found a *linear* relationship between galactic distance and redshift (E. Hubble 1929, PNAS, **15**, 168.) How did Hubble manage to determine the distances to the nebulae? Using the new 100-inch telescope on Mt Wilson, Hubble had earlier resolved individual Cepheid variable stars in the outer arms of the Andromeda spiral *galaxy*, as we shall henceforth refer to it. For in obtaining the distances to the “nebulae,” Hubble also showed that they must be galaxies in their own right.

Cepheid variable stars were at the time already well-studied in our own galaxy. Cepheids oscillate in brightness with periods ranging from days to months. What makes them an important candle for observational astronomers is that they have a well-defined relationship between their period time period and their absolute mean luminosity. To measure a distance, one proceeds as follows. Find a Cepheid variable. Measure its oscillation period. Determine thereby its true luminosity. Measure the star’s flux (energy per unit area per unit time crossing your detector—this is all that can actually be measured). The flux is the absolute luminosity  $L$  divided by  $4\pi r^2$ , where  $r$  is the distance to the star. Therefore, by measuring the flux and inferring  $L$  from the oscillation period, one may deduce  $r$ , the distance to the star. Simple—if you just happen to have a superb quality, 100-inch telescope handy. To such an instrument, only Hubble and a small handful of other astronomers had access<sup>14</sup>.

### 8.7.2 The cosmic distance ladder

As observations improved through the 1930’s the linear relation between velocity and distance, which became known as the *Hubble Law*,  $v = H_0 r$ , became more firmly established. There are two major problems with collecting data in support of the Hubble expansion.

First, galaxies need not be moving with the Hubble expansion (or “Hubble flow”): their motions are affected by neighbouring masses. The best known example of this is the Andromeda galaxy, whose redshift is in fact a blueshift! It is approaching our own Milky Way Galaxy at about 300 km per second. The redshift measurement problem is greatest for a nearby galaxy whose “peculiar velocity” (deviation from Hubble flow) is a large fraction of its Hubble recession velocity. Second, it is very difficult to establish distances to cosmological objects. We can establish distances to relatively nearby objects relatively easily, but these galaxies are precisely the ones affected by large peculiar velocities. Those galaxies unaffected by large peculiar velocities are just the ones whose distances are difficult to establish!

But observational astronomers are resourceful, and they have come up with a number of ingenious techniques which have served them well. The idea is to create a cosmic “distance ladder” (perhaps better described as a linked chain). You start with direct measurements on certain objects, and then use those measurements to calibrate the distances to other, somewhat more distant, objects. Then repeat. Here is how it actually works.

Start with our solar system. These days, we can bounce radar signals off planets and measure the time of flight (even testing general relativity in the process, as we have seen) to get extremely accurate distances. Next, we make use of our precise knowledge of the astronomical unit (denoted AU) thus obtained to use the classic technique of trigonometric parallax. The earth’s motion around the sun creates a baseline of about 2 AU, which allows us to have a different perspective on nearby stars as we orbit. We see nearby stars shift in angular position on the sky relative to their much more distant counterparts. The

---

<sup>14</sup>Much later, in 1952, Hubble’s quantitative results were found to be inaccurate. It turns out that Cepheids come in two quite separate populations, with very different Period-Luminosity relations! This was discovered by W. Baade, who was then motivated to introduce the concept of distinct stellar populations (differing by ages) into astronomy. This also revised the whole extragalactic distance scale, though it kept intact the linearity of the redshift-distance relation.

observed angular shift is inversely proportional to the distance of the nearer star from the solar system. We define the unit of distance known as a parsec (pc) as the distance at which a separation of 1 AU subtends 1 second of arc. We can turn this definition around and say that the parallax angle, perversely denoted as  $\pi$  in the astronomical literature and measured in arc seconds, corresponds to a distance to a measured star of  $1/\pi$  parsecs. (One parsec is  $\simeq 3.085678... \times 10^{16}$  m. Because 1 AU is now a defined exact quantity, so is 1 pc.)

The next rung up the ladder is known as spectroscopic parallax. There is no actual parallax angle in this case, it is just an analogous name. The idea is that you first use the method of direct trigonometric parallax to obtain distances to stars of a *given spectral type*. Knowing the distance to a star of a given spectral type and measuring its flux, you also know the intrinsic luminosity  $L$ , since  $F = L/4\pi r^2$  where  $F$  is the measured flux and  $r$  the measured distance to the star. Now you see a star with precisely the same spectrum, but much fainter and therefore more distant. You cannot measure its parallax, it is too far away. But the detailed agreement between spectra means this is the same type of star you've already measured, with the same mass and same luminosity. You then measure the new star's flux  $F$ , and deduce *its* distance!

Keep going. Distances to Cepheid variables can be calibrated by spectroscopic parallax, and these bright stars have, as we have explained, a well-defined period-luminosity relationship. They are bright enough that they can be seen individually in external galaxies. Measure, thereby, the distances to these external galaxies.

Keep going. In studying the properties of external galaxies, Tully and Fisher showed that the rotation velocities of spiral galaxies (measured by the Doppler shifts in the combined spectra of moving stars) was tightly correlated with the intrinsic luminosity of the galaxy. This is not terribly surprising in itself: the larger the stellar mass of a galaxy the larger the luminosity, and the larger the binding mass the larger the rotational velocity. Tully and Fisher crafted this notion into a widely used tool for establishing the distances to very distant galaxies. Elliptical galaxies, which are supported by the dispersion of stellar velocities rather than their systematic rotation were also turned into useful distance indicators. Here the correlation between luminosity and velocity dispersion is known as the Faber-Jackson relation.

The final step in the cosmic distance ladder involves Type Ia supernovae. Type Ia supernovae are thought to occur when a white dwarf in a tight binary system accretes just enough matter from its companion to tip itself over the "Chandrasekhar mass." (This mass is the maximum possible mass a white dwarf can sustain by electron degeneracy pressure, about 1.4 times the mass of the sun.) When this mass is exceeded, the white dwarf implodes, overwhelmed by its now unsupportable self-gravitational attraction. In the process, carbon and oxygen nuclei are converted to  $^{56}\text{Ni}$ , triggering a thermonuclear explosion that can be seen quite literally across the Universe. What is nice about type Ia supernovae, from an astronomical perspective, is that they always occur in a white dwarf of the same mass. Therefore there is little variation in the absolute intrinsic luminosity of the supernova explosion. To the extent that there is some variation, it is reflected not just in the luminosity, but in the rise and decay times of the emission, the "light curve." The slower the decline, the larger the luminosity. So you can correct for this. This relation has been well calibrated in many galaxies with well-determined distances.

The Type Ia supernova data were the first to provide compelling evidence that the Universe was expanding. To understand how this was obtained we need to return to our formula for the flux,  $F = L/4\pi r^2$ , and understand how this changes in an expanding, possibly curved, spatial geometry.

### 8.7.3 The parameter $q_0$

Observational astronomers often work with distances where the time it has taken light to reach us, measured from the present, is small compared with the age of the Universe. The goal is to measure small changes from conditions at the current epoch,  $t = t_0$ . Under these circumstances, it is mathematically convenient to expand our basic cosmological functions  $R(t)$  and  $H(t)$  to leading order in a Taylor series in  $t - t_0$ . For example, the Hubble parameter  $\dot{R}/R$  may be expanded in time:

$$H(t) = \frac{\dot{R}}{R} = H_0 + (t - t_0) \left( \frac{\ddot{R}_0}{R_0} - H_0^2 \right) + \dots \quad (550)$$

since  $\dot{R}^2/R^2$  at the present time  $t_0$  is  $H_0^2$ . (Remember  $t < t_0$  for observed sources.) This may also be written

$$H(t) = H_0 + H_0^2(t_0 - t)(1 + q_0) \quad (551)$$

where we have defined the *deceleration parameter*  $q_0$  by

$$q_0 \equiv -\frac{\ddot{R}_0 R_0}{\dot{R}_0^2} \rightarrow -\frac{\ddot{R}_0}{H_0^2} \quad \text{for } R_0 = 1. \quad (552)$$

This is important observational quantity because it embodies the observed evolution of the Hubble parameter. Next, use  $(1 + z)R(t) \equiv 1$  and expand  $R(t)$  in a similar Taylor series around  $t_0$ . Then, to leading order in  $z$ , we obtain  $z = H_0(t_0 - t)$ , and then (551) becomes

$$H(z) = H_0[1 + z(1 + q_0)]. \quad (553)$$

For the classical matter-dominated cosmological models,  $q_0$  is defined in such a way that should be a positive quantity. If  $q_0 < 0$ , the Universe would be accelerating, a possibility not taken seriously until after 1998. Consider the integral for the comoving coordinate  $r = \int c dt / R(t)$ . With  $R_0 = 1$  and  $t - t_0 \equiv \delta t$ , carry out the Taylor expansion in  $\delta t$  of

$$1 + z = \frac{1}{R(t)} = \frac{1}{R(t_0 + \delta t)} = 1 - \delta t \dot{R}_0 + (\delta t)^2 (\dot{R}_0^2 - \frac{\ddot{R}_0}{2}) + \dots \quad (554)$$

To linear order this gives  $\delta t = -z/\dot{R}_0 = -z/H_0$ . Using this result in the small quadratic correction term in  $(\delta t)^2$  in (554), the next order refinement is

$$-\delta t = t_0 - t = \frac{1}{H_0} \left[ z - \left( 1 + \frac{q_0}{2} \right) z^2 \right]. \quad (555)$$

Using our expansion for  $1/R$  we find,

$$l(z) \equiv c \int_t^{t_0} \frac{dt'}{R} = c(t_0 - t) + \frac{cH_0}{2}(t - t_0)^2 + \dots \quad (556)$$

and substituting with the help (555) through order  $z^2$ :

$$l(z) = \frac{c}{H_0} \left[ z - \frac{z^2}{2} (1 + q_0) + \dots \right] \quad (557)$$

Equation (557) is general for any FRW model.  $q_0$  embodies the leading order deviations from a simple Euclidian model in which  $l(z) \propto z$ .

#### 8.7.4 The redshift–magnitude relation

The flux that is measured from a source at cosmological distances differs from its simple  $L/4\pi r^2$  form for several important reasons. It is best to write down the answer, and then explain the appearance of each modification. For generality, we keep factors of  $R_0$  present, so that general FRW metrics are permitted. The flux from an object at redshift  $z$  is given by

$$\mathcal{F}(z) = \frac{LR^2(t)}{4\pi R_0^4 l^2(z)} = \frac{L}{4\pi(1+z)^2 R_0^2 l^2(z)} \quad (558)$$

where  $l(z)$  is computed, as always, by  $c \int_{t_0}^t dt'/R(t')$ , with this definite integral written as a function of  $z$  for the cosmological model at hand. (Recall  $R(t)/R_0 = 1/(1+z)$ , relating  $t$  to  $z$ .) The luminosity distance is then defined by  $\mathcal{F}(z) \equiv L/4\pi d_L^2$  or

$$d_L(z) = (1+z)R_0 l(z), \quad (559)$$

which reduces to the conventional Euclidian distance at small  $z$ .

Explanation: The two factors of  $1+z$  in the denominator of (558) arise from the change in the luminosity  $L$ . First, the photons are emitted with Doppler-shifted energies. But even if you were measuring only the rate at which the photons were being emitted (like bullets), you would require an additional second  $1+z$  factor, quite separate from the first, due to the emission interval time dilation. The photons have less energy and they are emitted less often. The proper radius of the sphere over which the photons from the distant source at  $z$  are now distributed is  $R_0 l(z)$ , where  $R_0$  is as usual the *current* value of  $R(t)$ . For the Einstein-de Sitter universe with  $R_0 = 1$ , the explicit formula is:

$$\mathcal{F}(z) = \frac{LH_0^2}{16\pi c^2(1+z)(\sqrt{1+z}-1)^2} \quad (\text{Einstein – de Sitter}) \quad (560)$$

The simple Euclidian value of  $L/4\pi r^2$  is recovered at small  $z$  by remembering  $cz = v = H_0 r$ , whereas at high  $z$ , all the photons come from the current horizon distance  $2c/H_0$ , and the  $1/z^2$  behaviour in  $\mathcal{F}$  is due entirely to  $1+z$  Doppler shifts.

A “magnitude” is an astronomical conventional unit used for, well, rather arcane reasons. It is a logarithmic measure of the flux. Explicitly:

$$\mathcal{F} = \mathcal{F}_0 10^{-0.4m} \quad (561)$$

where  $\mathcal{F}$  is the measured flux,  $\mathcal{F}_0$  is a constant that changes depending upon what wavelength range you’re measuring.  $m$  is then defined as the apparent magnitude. (Note: a larger magnitude is *fainter*. Potentially confusing.) The “bolometric magnitude” covers a wide wavelength range and is a measure of the total flux; in that case  $\mathcal{F}_0 = 2.52 \times 10^{-8} \text{ J m}^{-2}$ . Astronomers plot  $m$  versus  $z$  for many objects that ideally have the same intrinsic luminosity, like type Ia supernovae. Then they see whether the curve is well fit by a formula like (558) for an FRW model. It was just this kind of exercise that led to the discovery in 1998-9 by Perlmutter, Riess and Schmidt that our Universe must have a large value of  $\Omega_V$ : we live in an accelerating Universe!

*Exercise. Show that*

$$\mathcal{F}(z) = \frac{LH_0^2}{4\pi c^2 z^2} [1 + (q_0 - 1)z + \dots]$$

for any FRW model. Thus, with knowledge of  $L$ , observers can read off the value of  $H_0$  from the dominant  $1/z^2$  leading order behaviour of the redshift-magnitude data, but that knowledge of  $q_0$  comes only once the leading order behaviour is subtracted off.

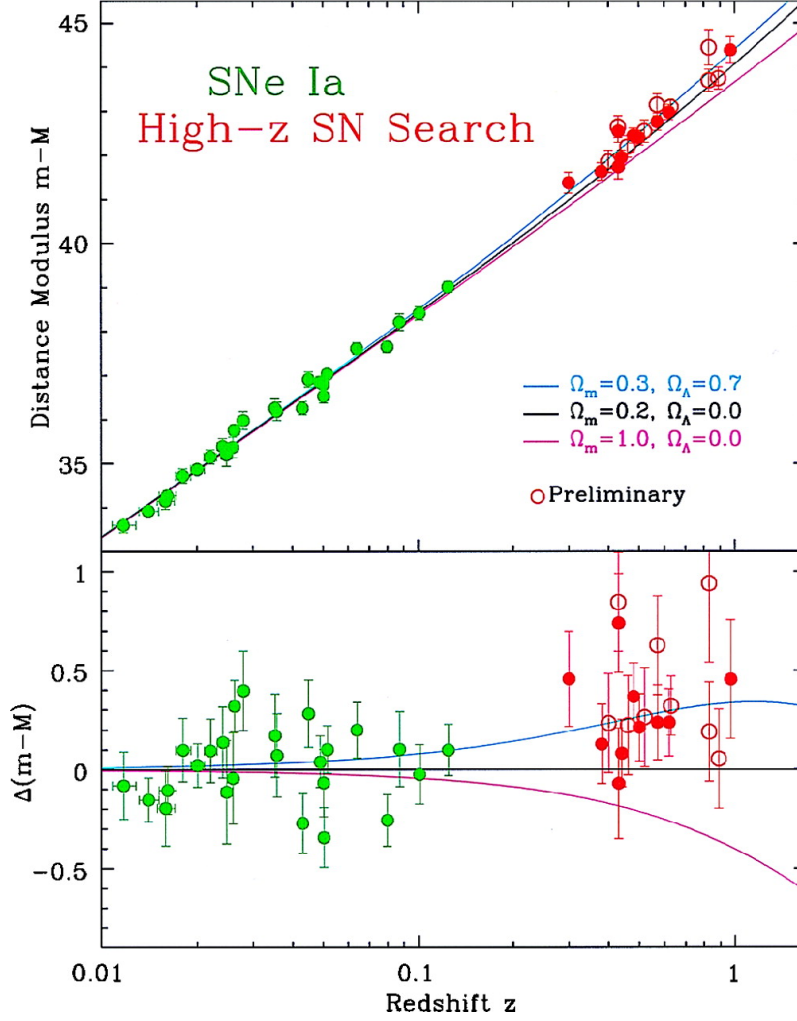


Figure 14: Evidence for an accelerating Universe from type Ia supernovae. The top figure shows a redshift-magnitude plot for three different FRW models,  $\Omega_{M0} = 1$ ,  $\Omega_V = 0$ , an Einstein-de Sitter model in mangenta;  $\Omega_{M0} = 0.2$ ,  $\Omega_V = 0$ , an open model, in black, and  $\Omega_{M0} = 0.3$ ,  $\Omega_V = 0.7$  an accelerating model, in blue. The bottom panel shows the same with the inverse square slope removed. The data are much better fit by the accelerating model.



With the determination of  $\Omega_V$ , the classical problem of the large scale structure of the Universe has been solved. There were 6 quantities to be determined:

- The Hubble constant,  $H_0 = \dot{R}/R$  at the present epoch.
- The age of the Universe,  $t_0$ .
- The curvature of the Universe, in essence the integration constant  $2E = -c^2/a$ .
- The ratio  $\Omega_0$  of the current mass density  $\rho_0$  to the critical mass density  $3H_0^2/8\pi G$ .
- The value of the cosmological constant or equivalently, the ratio  $\Omega_V$  of the vacuum energy density  $\rho_V c^2$  to the critical energy density.
- The value of the  $q_0$  parameter.

Within the context of FRW models, these parameters are not completely independent, but are related by the dynamical equations for  $\ddot{R}$  (514) and  $\dot{R}^2$  (526). A quick summary:

- $H_0 \simeq 67.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .
- $t_0 \simeq 13.8 \text{ billion years}$
- $E \simeq 0$
- $\Omega_{M0} \simeq 0.31$
- $\Omega_V \simeq 0.69$ .
- $q_0 = \Omega_0/2 - \Omega_V \simeq -0.54$

*Exercise. Derive the last result on this list.*

This brief list of values hardly does justice to the century-long effort to describe our Universe with precision. Because astronomers were forced to use galaxies as “standard candles” (the colloquial term for calibrated luminosity sources), their measuring tools were fraught with uncertainties that never could be fully compensated for. It was only the combination of establishing a truly standard candle via the type Ia supernovae, together with the technical capabilities of high receiver sensitivity and automated search techniques that allowed the programme (“The Supervova Cosmology Project”) to succeed.

Since the 1998/9 breakthrough, cosmologists have not been idle. The development of extremely sensitive receivers and sophisticated modelling techniques have turned the remnant microwave radiation “noise” from the big bang itself into a vast treasure trove of information. In particular, the nature of the tiny fluctuations that are present in the radiation intensity—more specifically the radiation temperature—allow one to set very tight constraints on the the large scale parameters of our Universe. Not only are these measurements completely consistent with the supernova data, the results of the missions known as WMAP and Planck render this same data all but obsolete! Figure (12) speaks for itself.

## 8.8 Radiation-dominated universe

The early stages of the Universe were dominated by radiation, which includes very relativistic particles along with photons. We denote the energy density of radiation as  $\rho_\gamma c^2$ . In such a universe,

$$\rho_\gamma \propto 1/R^4.$$

This follows from the fact that the equilibrium energy density of radiation is proportional to  $T_\gamma^4$ , where  $T_\gamma$  is the radiation temperature. The temperature  $T_\gamma$  has the same  $R$ -scaling as a photon with energy  $h\nu$ : both the frequency  $\nu$  as well as  $T_\gamma$  (which measures average energy per particle) decrease as  $1/R$ . In these early times, the dynamical equation of motion (526) is dominated by the two terms on the left side of the equation, both of which are very large (becoming infinite as  $t \rightarrow 0$ ) compared with the constant on the right side. Then, using  $\rho_\gamma R^4 = \text{constant}$ , the dynamical equation of motion at early times may then be written

$$R^2 \dot{R}^2 = \frac{8\pi G \rho_\gamma R^4}{3} = (\text{constant}). \quad (562)$$

This leads to a simple power law time dependence for the scale factor:

$$R(t) \propto t^{1/2} \quad (\text{radiation dominated universe}). \quad (563)$$

Knowing only that  $R \sim t^{1/2}$ , if we now return to the dynamical equation of motion, we may solve explicitly for  $\rho_\gamma c^2$ :

$$\rho_\gamma c^2 = \frac{3c^2}{8\pi G} \left( \frac{\dot{R}^2}{R^2} \right) = \frac{3c^2}{32\pi G t^2}. \quad (564)$$

We then have an *exact* expression for what the total energy density in *all* relativistic particles must be, even if we don't know how many relativistic species there are. (In practise, we have a good idea of this number from other physics!)

This is a rather neat result. At a time of one second, we know that the Universe had an energy density of  $4 \times 10^{25} \text{ J m}^{-3}$ , and that is that. Moreover, since the total energy density is fixed, the greater the number of relativistic particle populations that were present in the early Universe (e.g., the various neutrino species), the *smaller* the temperature at a given time. During the epoch when hydrogen was being fused into helium and a few other low atomic number nuclei, the production rate of these latter isotopes was sensitive to temperature. (By contrast, helium production is not so temperature sensitive, which makes a prediction of its abundance rather robust.) This temperature sensitivity of the light isotopes, combined with their observationally-determined abundances, has been used to limit the number of different types of neutrinos that could have been present during the era of nucleosynthesis.

## 8.9 The Cosmic Microwave Background Radiation (CMB)

### 8.9.1 Overview

The Universe is expanding, and expanding systems cool adiabatically. That means if we follow history backwards, the compressing Universe becomes hot and dense. We have seen that the energy density of a radiation scales as  $1/R^4$ , whereas the energy density of matter (dominated by its rest mass) scales as  $1/R^3$ . Since  $1/R^4$  gets bigger than  $1/R^3$  as  $R$  gets smaller, the Universe must at some point in the past have become radiation dominated. This, we shall see, occurred at a redshift  $z$  of about 1000.

You should now be thinking of the radiation as true photons; electrons and positrons (and probably neutrinos as well) are relativistic only at the very earliest times. At redshifts larger than 1000, not only was radiation dominant, the photons were also “well-coupled” to ordinary matter. That means that photons would scatter many times off of electrons during the time it took the Universe to  $\sim$  double in size. These collisions kept the radiation in a state of complete thermodynamic equilibrium, the radiation temperature did not differ from the matter temperature. In thermodynamic equilibrium, the number density of photons is *fixed*, given by the Planck function associated with a blackbody spectrum:

$$n(\nu, T) d\nu = \frac{(8\pi\nu^2/c^3)d\nu}{\exp(h\nu/kT_\gamma) - 1} \quad (565)$$

You will recall that if radiation is in thermal equilibrium with matter, you don’t get to choose how many photons there are at a given frequency. Mother Nature does. It is worth dwelling on this point. Unlike a gas bag of atoms in thermal equilibrium, whose density can be whatever we like, a gas bag of photons has its number density completely determined by its temperature! The spectrum may be directly proportional a Planck function, but if the overall number density is not given *exactly* by (565), the photons are simply not in thermal equilibrium with their surroundings. So does the inevitable decline in number density with universal expansion mean that we lose a thermal equilibrium distribution for the radiation?

At about the time that the radiation became subdominant relative to matter, the photons also lost good coupling with the matter. Thermal equilibrium could no longer be maintained via photon-electron collisions or induced exchanges with atoms. Imagine now a situation in which the photons “move along” with the expansion of the Universe, but are otherwise unaffected by interactions. From what we’ve said, you might guess that while the overall *shape* of the frequency spectrum would be maintained, the photon number density would be diluted below that of a true blackbody. (Rather like sunlight reaching the earth.) Not so. In fact, both the number density and effective temperature of the photons vary in just the right way to maintain an exact blackbody distribution! The ever decreasing number density is always exactly the correct one for a true thermal Planck spectrum for the evolving and ever cooling photon temperature.

To see this, start with the spectrum (565) at the time  $t$  of last scattering, by which we mean the last moment thermal equilibrium was maintained by collisions. We follow a little group of photons with frequency  $\nu$  and little dispersion  $d\nu$  as these photons evolve with the Universe. Now it is some much later time  $t'$ . The  $t'$  number density  $n(\nu', t', T')$  is diluted by an overall (volume) factor of  $R^3(t)/R^3(t')$  relative to the original blackbody number density at the time of last scattering  $t$ . In addition, each photon at  $t'$  that we now observe at frequency  $\nu'$  must have come from a frequency way back at time  $t$  given by  $\nu = R'\nu'/R$  (where  $R' \equiv R(t')$ ), in the original blackbody distribution. Therefore, to deduce the later time  $t'$  spectrum, follow this rule: take the original time- $t$  blackbody distribution, dilute it by  $R^3/R'^3$ , and then replace  $\nu$  everywhere in the Planck formula by  $R'\nu'/R$ . Carrying this through,

$$n(\nu', t')d\nu' = \frac{R^3}{R'^3}n(\nu, t)d\nu = \frac{R^3}{R'^3} \frac{(8\pi/c^3)(R'\nu'/R)^2 d(\nu'R'/R)}{\exp(hR'\nu'/RkT_\gamma) - 1} = \frac{8\pi\nu'^2 d\nu'/c^3}{\exp(h\nu'/kT'_\gamma) - 1} \quad (566)$$

where  $T'_\gamma = T_\gamma R/R'$ , a *cooler* temperature. The remarkable point is that the distribution (566) is *still* a blackbody at time  $t'$ , but with a new temperature  $T'_\gamma$  that has cooled in proportion to  $1/R$  with the expansion. The reason this is remarkable is that there is nothing to maintain this thermal equilibrium! In fact, you can see that this scaling result follows mathematically for any number spectrum of the form  $\nu^2 F(\nu/T) d\nu$ , where  $F$  is an *arbitrary* function.

Might we just be playing a mathematical game? What if, in the real Universe, other interactions scatter photons and change their frequencies in the process? For many years, there were arguments as to whether the actual cosmic radiation spectrum was truly a blackbody. Adding to the confusion, there seemed to be outlandish observations. These were largely based on strikingly un-blackbody spectra obtained from rocket-borne instruments. Now, to significantly change the radiation spectrum of the *entire* Universe would take a vast amount of energy. From where would that have come? Theorists tied themselves into knots trying to come up with explanations! All doubts were finally erased in 1992 when the COBE (COsmic Background Explorer) satellite was launched. This satellite had (submillimeter) detectors of unprecedented accuracy, and they showed that the background radiation was almost a perfect blackbody at  $T_\gamma = 2.728K$ . Figure (15) summarises the data. The error bars are shown at 400 times their actual value, just in order to be visible. (The earlier misleading rocket observations were by this time retracted.)

You may have noticed that I wrote “almost a perfect blackbody.” The photon distribution is in fact not *exactly* a blackbody for two very interesting reasons. The first is that the earth is not at rest relative to the CMB frame. In fact, the local group of galaxies seems to be moving at about  $630 \text{ km s}^{-1}$  relative to the CMB, a surprisingly large value. This translates to a measured radiation temperature that is about  $2 \times 10^{-3}K$  warmer in one direction, and the same amount cooler in the opposite direction. This induces no change in the overall energy of the radiation.

The second reason is much deeper in its physics. The objects that comprise our Universe collapsed out of an expanding gaseous background, forming great clusters of galaxies and ultimately individual galaxies and stars. The seeds to form these structures could not have appeared recently, after the last scattering event. There simply hasn’t been enough time for them to have grown from seeds, collapse and form the objects we see throughout the Universe today. The seeds must already have been present during the era when the Universe was radiation-dominated, the same era when matter and radiation were still interacting strongly with one another. This interaction would have left its imprint, at the time of last scattering, in the form of temperature *fluctuations* in the microwave background radiation. Scrunch the matter, you also compress the radiation, and you heat it up a bit. Some places are little hotter, others a bit cooler (where there has been dilation). The simplest calculations suggested that the relative temperature fluctuation  $\Delta T/T$  should have been at least  $\sim 10^{-3}$ , in order for these initially tiny compressions to be able to form nonlinear structures by the current epoch.

COBE found the temperature fluctuations. But when COBE (and later satellites in great detail) did find them, they were an order of magnitude smaller. So how does structure form in the Universe? The consensus answer is that there is more to matter in the Universe than the usual atomic nuclei and electrons that we know about. The Universe is also pervaded by what is known as “dark matter,” which turns out to be most of the so-called ordinary matter, and apparently leaves no imprint on the microwave background. Dark Matter feels the force of gravity and evidently little else, and had been postulated long before COBE as something that was needed to hold galaxies together. COBE offered another compelling reason to believe in Dark Matter.

The directors of the COBE project, John Mather and George Smoot, won the 2006 Nobel Prize in Physics for their work.

## 8.9.2 An observable cosmic radiation background: the Gamow argument

The idea that the Universe had a residual radiation field left over from its formation *and that this radiation is potentially observable*, seems to have originated with George Gamow in

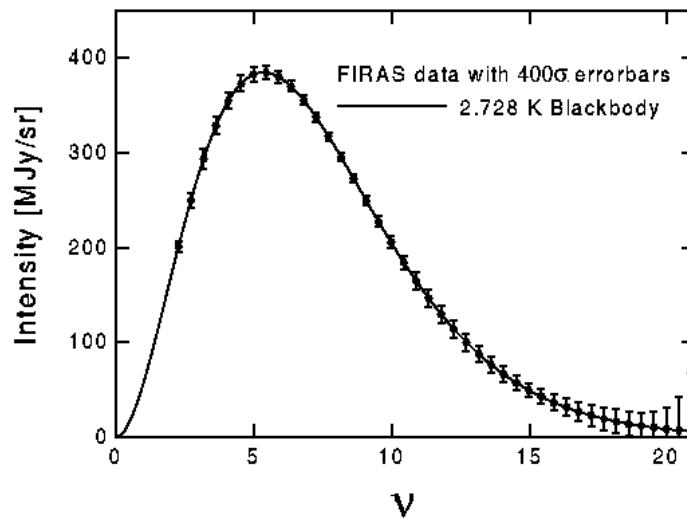


Figure 15: COBE satellite data showing a perfect fit to a blackbody at 2.728 K. To be visible, the error bars are shown at 400 times their actual value! When shown at an American Astronomical Society Meeting, this plot triggered a spontaneous standing ovation. Units of  $\nu$  are  $\text{cm}^{-1}$ , i.e., the wavelength in cm is the reciprocal of the number on the  $x$  axis.

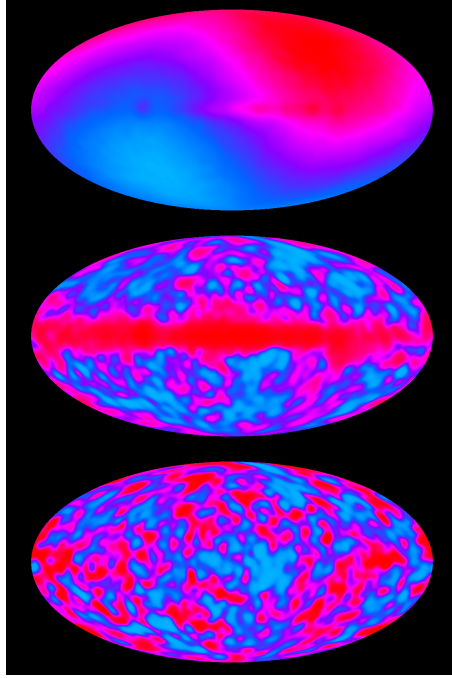
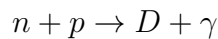


Figure 16: COBE temperature fluctuations. *Top*: Temperature dipole variation dominated by motion of the galaxy. *Middle*: Residual with kinematic dipole subtracted off. *Bottom*: Further residual with galactic foreground subtracted off. These represent primordial fluctuations in the CMB.

the 1940's. Gamow was a theoretical physicist with a brilliant common sense instinct that allowed him to make contributions to a wide variety of important problems, from the theory of radioactivity to how DNA coding works. He spent many years developing, and making respectable, what has become known as the Big Bang theory<sup>15</sup>.

Gamow drew attention to the fact that Helium, with about 25% of the mass of the Universe, is neither an overwhelming nor a tiny constituent. Gamow was convinced, we now know correctly, that Helium was made during a brief interval of nucleosynthesis in the early Universe. (Gamow thought all the other elements were made this way as well, but that turns out to be false. Heavier elements are made in stars.) The 25% number implies that at the time of nucleosynthesis, the expansion rate of the Universe and the nuclear reaction rate could not have been very different. Too rapid an expansion, there is no Helium made. Too slow an expansion rate, *all* the protons get fused into Helium, and there is no Hydrogen left over. The delicate balance turns out to be a remarkably tight constraint, leading to a prediction that there should today be an observable residual radiation field of about 10 K. Let's see how it works.

At some point early in its history, the Universe's temperature passes downward through the range of  $\sim 10^9$  K. When that happens, neutrons  $n$  and protons  $p$  can begin to combine to form Deuterium (a proton  $p$  plus a neutron  $n$ ) nuclei. The reaction is pretty simple:



where  $\gamma$  is a gamma ray. Helium synthesis follows rapidly thereafter. Gamow assumed equal numbers of protons, neutrons and electrons. As we have emphasised, in order to get

---

<sup>15</sup>The name was coined by Fred Hoyle, a rival cosmologist who meant the description to be a disparaging moniker. But proponents of the theory loved the name and adopted it as their own.

a 25% yield of Helium, the reaction rate and the age of the Universe (which is of order the expansion time to double in size) should be comparable. Denote the cross section for Deuterium formation to be  $\sigma$  (units of area). Then, the reaction rate per proton due to an incoming flux  $nv$  of neutrons is  $nv\sigma$  (units of pure number per time). At the nucleosynthesis age  $t$ , on average very roughly one reaction per proton should have occurred, because of the order unity fraction of Helium that is observed—that is the heart of the Gamow argument:

$$nv\sigma t \sim 1, \quad (567)$$

where  $t$  is also the age of the Universe! The product  $\sigma v$  turns out to be just a laboratory number, nearly independent of  $v$  (because the cross section  $\sigma$  depends inversely on  $v$ ) and, as Gamow knew from nuclear physics experiments, is equal to about  $4.6 \times 10^{-26} \text{ m}^3 \text{ s}^{-1}$ . As for the time  $t$ , we are interested in the epoch when the temperature  $T_\gamma \sim 10^9 \text{ K}$ . Using equation (564) for the density in relativistic particles, and following Gamow by assuming (not quite correctly, but let it go) that this was all radiation, the time  $t$  is

$$t = \frac{c}{T_\gamma^2} \left( \frac{3}{32\pi G a} \right)^{1/2} \quad (568)$$

which amounts to 230 s for  $T_\gamma = 10^9 \text{ K}$ . This gives  $n$  very close to  $10^{23} \text{ m}^{-3}$ . Next, Gamow estimated a present day average particle density of about 0.1 per cubic meter based on astronomical estimates and observations, so that the Universe had expanded by a factor of  $10^8$  in the scale factor  $R$ . (The factor  $10^8$  reduces  $n$  by  $10^{24}$ .) But an expansion of a factor of  $10^8$  in  $R$  means that the current  $T_\gamma$  should be  $10^9/10^8 = 10 \text{ K}$ ! The millimeter detectors that would have been required for this wavelength of observation were just at the leading edge of technology in the late 1940s, a by-product of the development of radar during the Second World War (which was instrumental to winning the Battle of Britain). But Gamow did not pursue this result aggressively, and the prediction was gradually forgotten. This CMB could have been detected in the late 1940's, but it wasn't.

The argument would be put slightly differently now, but it is in essence correct, and a brilliant piece of intuitive reasoning.

*Exercise.* Repeat the Gamow argument with modern cosmological numbers. Keep  $T_\gamma = 10^9 \text{ K}$  and the  $\sigma v$  value, but note that relativistic species present at the time include not only photons but three types of neutrinos and three types of antineutrinos, each neutrino-anti neutrino combination contributing  $(7/8)aT_\gamma^4$  to the background energy density. We can neglect  $e^+$  and  $e^-$  pairs in our relativistic fermion population. Why? The 7/8 factor arises because the neutrinos obey fermi statistics. Recall from your statistical physics that the occupation number for a state with momentum  $p$  is

$$n_\pm = \frac{1}{\exp(pc/kT) \pm 1}$$

where the  $+$  sign is for bosons (e.g. photons) and the  $-$  sign for fermions (e.g. neutrinos). The density of states (per unit volume) is the same for both types of particles:  $4\pi p^2 dp/h^3$ , where  $h$  is the Planck's constant. The energy is the same for each type of particle,  $pc$ . Therefore, the difference in the total energy densities of a population of bosons or fermions comes down to the difference in two simple integrals:

$$\int_0^\infty \frac{x^3 dx}{e^x + 1} = \frac{7}{8} \int_0^\infty \frac{x^3 dx}{e^x - 1}$$



Can you prove this mathematically *without* doing either integral explicitly? (Hint: consider the difference

$$\frac{1}{e^x - 1} - \frac{1}{e^x + 1},$$

and integrate  $x^3$  times the residual.) Note that for each neutrino specie the factor is actually  $7/16$ , because only *one* spin helicity is present, as opposed to *two* spin states for photons. But then the neutrinos get back up to  $7/8$  with their antiparticle cousins...which the photons lack! You may take the current density in ordinary baryons to be 5% of the critical density  $3H_0^2/8\pi G$ .

Almost 20 years later, in 1965, the problem of determining the current  $T_\gamma$  attracted the attention of an impressive team of physicists at Princeton University. They rediscovered for themselves the Gamow argument. The senior investigator, Robert Dicke, a hugely talented physicist (both theoretical and experimental), realised that there was likely to be a measurable background radiation field that survived to the present day. Moreover, it could be easily detected by an instrument, the Dicke radiometer, that he himself had invented twenty years before! (A Dicke radiometer is a device that switches 100 times per second, looking between the sky source and a carefully calibrated thermal heat bath of liquid helium. This imprints a 100Hz fourier component on the desired signal, and thereby reduces nuisance variability occurring on longer time scales.)

The A-Team assembled: Dicke, the scientific leader; J. Peebles, the brilliant young theorist who would become the world's leading cosmologist in the decades ahead, and P. Roll and D. Wilkinson, superb craftsmen who designed and built the contemporary Dicke radiometer. They were all set up to do the observation from the roof of their Princeton office building(!), when a phone call came from nearby Bell Telephone Laboratories. Two radio engineers named Arno Penzias and Robert Wilson had found a weird extraneous signal in their detector. Their device was designed to receive longrange signals relayed by some of the first commercial television satellites. (The *Telstar* series.) Naturally, they were trying to chase down all possible sources of background confusion. The "effective radiation temperature" of the nuisance diffuse signal was about 3 K. Penzias and Wilson had no idea what to make of it, but were advised by a colleague to give the Dicke team at Princeton a call. Those guys are very clever you know, they might just be able to help.

The 1978 Nobel Prize in Physics went to Penzias and Wilson for the discovery of the cosmic microwave background radiation.

### 8.9.3 The cosmic microwave background (CMB): subsequent developments

The initial observations of the CMB were at one wavelength: 7 cm. Needless to say, a single point does not establish an entire spectrum! The task of establishing the broader spectrum was fraught with difficulties, however, with many disputed and ultimately withdrawn claims of large deviations from a blackbody. In 1992, we have seen that matters were laid to rest when the COBE satellite returned its dramatically undramatic finding that the CMB is, very nearly, but not *exactly*, a blackbody. There are small fluctuations in the temperature, the largest of which amount to a few parts in  $10^4$ . This value is smaller than first expected, but ultimately of the order needed to account for the nonlinear structure in the Universe that we see today, provided that there is a healthy component of *dark matter* that does not react with the radiation. This unseen, and perhaps unseeable, dark matter component had been invoked, long before the COBE results, for a very different reason: to account for the large stellar velocities measured in galaxies and within clusters of galaxies. These velocities are unsustainable unless most of the gravitating mass of the galaxy or cluster is in the form of dark, non-light-emitting matter. The COBE results are yet further evidence for the presence of dark matter in the Universe.



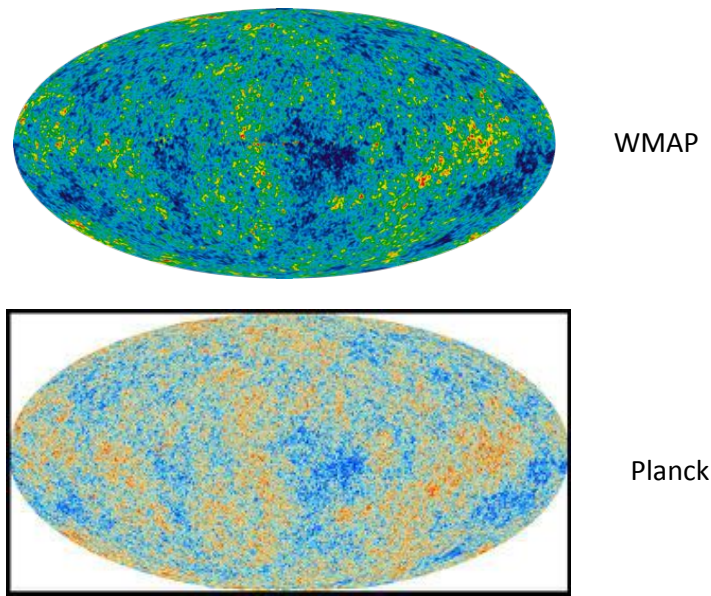


Figure 17: WMAP (top) and Planck (bottom) temperature fluctuations, with different colour scales. Think of these as the Universe’s baby pictures. Note the much higher angular resolution compared with COBE; Planck is another factor of 3 better than WMAP. The angular information encoded on these smaller scales is valuable as a means of tightly constraining cosmological parameters.

Following COBE, the next important CMB probe was the Wilkinson Microwave Anisotropy Probe, or WMAP. This is named for David Wilkinson, the same Princeton researcher who had been instrumental in the earliest CMB studies. With the launch of WMAP in 2003, cosmology truly became a precision science. WMAP revealed the structure of the temperature fluctuations in such exquisitely fine angular resolution on the sky, it became possible to determine all the key physical parameters of the Universe:  $H_0$ ,  $\Omega_{M0}$ ,  $\Omega_V$ ,  $t_0$  and many others, to an excellent level of accuracy. WMAP was followed by the Planck satellite, launched in 2009, which provided a further shrinking of the error bars, higher angular resolution coverage of the CMB on the sky, broader frequency coverage (very important for subtracting off the effects of the Milky Way Galaxy), and better constraints on null results (absence of CMB polarisation, for example). But somewhat disappointingly, there were no qualitatively new physical findings. The current story remains essentially the one revealed by WMAP.

Think of the evolution of the maps and globes of the Earth, from ancient to modern times. Gross inaccuracies in basic geometry gradually evolved to Googlemap standard over a period of thousands of years. Contrast this with the observation that, well within the professional careers of currently active researchers, serious models of the Universe went from fundamental misconceptions to three-significant-figure accuracy in structural parameters. By any measure, this is one of the great scientific achievements of our time. As with any great scientific advance, its true meaning and its full implications will take years to elucidate.

## 8.10 Thermal history of the Universe

### 8.10.1 Prologue

The recorded history of civilisation began when Sumerian merchants inscribed their grain inventories on clay tablets. (BTW, the concept of a “grain inventory” would not have been possible without the invention of large scale agriculture, which would not have been possible without the invention of astronomy. Please, stick with this for a moment: despite what you’re thinking, I am on-message...) Etched in the matrix of clay, these markings were frozen in time because the clay, once dried, remained unaltered and intact. It turns out that the Universe has its own gigantic clay tablet, which allowed “inscriptions” first to be imprinted, then to be preserved, pretty much unaltered for 13.8 billion years. We know the inscriptions as temperature fluctuations, and the clay tablet as the cosmic microwave background.

The temperature fluctuations are a tracer of the initial density fluctuations, which coupled to the radiation by electron-photon scattering, also known as *Thompson scattering*. While this scattering remained vigorous, the CMB dutifully recorded and then re-recorded the ever evolving changes in the density. Then, matter and radiation suddenly became decoupled, and the pattern imprinted on the CMB abruptly stopped being recorded. The CMB instead retained only the very last pattern that happened to be imprinted upon it, at the time of the “last scattering.” It is *this* pattern that we receive, redshifted by the enormous subsequent expansion of the Universe, in our detectors today. Think of this either as an ancient inscription passed on through the aeons, or more congenially as the Universe’s baby picture.

*Exercise.* Let  $\sigma_T$  be the Thompson cross section for a photon to collide with an electron, a constant number equal to  $6.7 \times 10^{-29} \text{ m}^2$ . When an electron has moved relative to the photon gas a distance  $l$  such that the swept-out volume  $l\sigma_T$  captures a single photon,  $l$  is said to be one *mean free path (mfp)*: the average distance between scatterings.

Justify this definition, and show that the scattering rate per electron is  $n_\gamma c \sigma_T = c/l$ . (What is  $n_\gamma$  here?) In a radiation-dominated Universe, show that the ratio of the photon scattering mfp to the horizon size grows like  $\sqrt{t}$ , and that the mfp relative to the scale factor  $R$  grows like  $t$ . ( $t$  as usual is time.)

Exactly how to decipher the ancient CMB inscriptions and turn them into a model of the Universe is a *very* complicated business. While it is a local Oxford Astrophysics speciality, it is one that we will be able to treat only very superficially in this course. We shall go about this in two steps. In the first, we will describe what I will call the “classical theory” of the thermal history of the Universe: how matter and radiation behaved in each other’s presence from temperatures of  $10^{12} \text{ K}$  through  $3000\text{K}$ , the time when hydrogen ions recombine with electrons to form neutral H atoms. In the second step, which the reader may regard as optional “off-syllabus” material, we will discuss a more modern theory of the very early Universe. This puts a premium on the notion of *inflation*, a period in the history of the Universe in which it seems to have undergone a very rapid (exponential!) growth phase. First put forth in 1980, this idea has been the single most important theoretical advance in modern cosmology in recent decades. There are very good theoretical reasons for invoking the process of inflation, even if the mechanism is not well-understood (not an unusual state of affairs in science), and there is by now compelling observational support for it. To my mind, the best evidence there is for inflation is that we have in fact entered another, albeit far more mild, inflationary stage of the Universe’s history.

In what follows, the reader should bear in mind the characteristic physical parameters on what is known as the “Planck scale.” These are the scales set by the three fundamental dimensional constants of physics: Newton’s  $G$ , Planck’s  $h$ , and the speed of light  $c$ . There

are unique dimensional combinations to form a mass  $m_P$ , a length  $l_P$  and a time  $t_P$  from these constants:

$$m_P = \left(\frac{hc}{G}\right)^{1/2} = 5.456 \times 10^{-8} \text{ kg} \quad (569)$$

$$l_P = \left(\frac{hG}{c^3}\right)^{1/2} = 4.051 \times 10^{-35} \text{ m} \quad (570)$$

$$t_P = l_P/c = \left(\frac{hG}{c^5}\right)^{1/2} = 1.351 \times 10^{-43} \text{ s} \quad (571)$$

With these three quantities, we can form any other dimensional quantity we like. These are, in some sense, the limits of our knowledge, for it is on these scales that we may expect quantum gravity effects to be important, a theory of which we remain quite ignorant. We cannot expect to have anything like a classical picture of the early Universe for time  $t < t_P$ , or for horizon scales  $c/H < l_P$ . The Planck mass may not, at first glance, appear remarkable, but remember the natural comparison is with elementary particle masses. An electron, for example, is close to  $5 \times 10^{-4}$  GeV in rest energy; the Planck mass energy is  $m_P c^2 = 3.06 \times 10^{19}$  GeV. By these Planckian standards, we will always be working with very low mass particles, and very long length and time scales.

### 8.10.2 Classical cosmology: Helium nucleosynthesis

Let us begin the story when the temperature of the Universe is just under  $10^{12}$  K. This is very early on, only about  $10^{-4}$  s after the big bang. At this stage, the Universe consists of a relativistic cocktail of photons, neutrinos and their antiparticles, muons and their antiparticles and electrons and their antiparticles (positrons). This cocktail is well-mixed (shaken, not stirred). Even the neutrinos are in complete thermal equilibrium, freely created and destroyed.<sup>16</sup> There is also a population of protons and neutrons, which are energetically unimportant at these temperatures. But keep track of them! They are going to make the Universe we know as home.

Once the temperature slips below  $10^{12}$  K, the muons and antimuons can annihilate, but there isn't enough energy for muon production from the other relativistic populations present. The energy from the photons and  $e^+e^-$  pairs produced in the annihilation is shared amongst all the other populations. Everyone is heated and takes part, including, at these ultra high densities, the normally elusive neutrinos.

The Universe continues to expand, the temperature falls. To understand the next important phase, recall that neutrons and protons do not have quite the same mass. More precisely, with  $m_p$  the proton mass and  $m_n$  the neutron mass,

$$\Delta mc^2 \equiv (m_n - m_p)c^2 = 1.293 \text{ MeV} = 2.072 \times 10^{-13} \text{ J}. \quad (572)$$

This is very small compared with either  $m_n c^2$  or  $m_p c^2$ . Except when we are specifically concerned with the mass difference or doing very accurate calculations, there is normally no need to distinguish  $m_n$  for  $m_p$ . Textbooks on statistical mechanics tells us that the ratio of the probabilities for finding a system in some state  $i$  or  $j$  depends only on the energies  $E_i$  and  $E_j$  of these states, and nothing else. In particular, the ratio of the probability of finding the system in  $i$  to the probability of finding it in  $j$  is given by the *Boltzmann equation*:

$$\frac{P_i}{P_j} = \frac{g_i}{g_j} \exp \left[ - \left( \frac{E_i - E_j}{kT} \right) \right] \quad (573)$$

---

<sup>16</sup>By using the word “freely,” we ignore the particle rest masses.

where the  $g$ 's are known as the *statistical weights*, a fancy name for how many distinct states there are with energy  $E_i$  or  $E_j$ . (If you've forgotten this and don't have a statistical mechanics textbook handy, and you are curious, see the **Exercise** below for a quick justification.) Below  $10^{11}$ K, the so-called Boltzmann factor  $\exp(-\Delta mc^2/kT)$  starts to differ noticeably from unity, and as we approach  $10^{10}$  it becomes rather small. Since neutrons are more massive than protons, the former become more scarce at cooler temperatures. The critical temperature corresponding to  $kT = \Delta mc^2$  is  $1.5 \times 10^{10}$ K. (By way of comparison, for electrons and positrons,  $m_e c^2/k = 5.9 \times 10^9$ K.)

*Exercise.* Let the probability of finding a system in state  $i$  be  $f(E_i)$ , where  $E_i$  is the energy of the state. The ratio of the probability of finding the system in state  $i$  versus  $j$  is then  $f(E_i)/f(E_j)$ . But this ratio must be a function only of the difference of the energies, because a constant additive constant in energy can't affect the physics! The potential energy is always defined only up to such a constant. Hence, with  $F$  an arbitrary function,

$$\frac{f(E_i)}{f(E_j)} = F(E_i - E_j).$$

If  $E_i = E_j + \delta$ , show that if this equation is to hold even to first order in  $\delta$ , that

$$f(E) \propto \exp(-\beta E)$$

where  $\beta$  is an as yet undetermined constant. We can determine  $\beta$  by physics, e.g. by demanding that the ideal gas equation of state be satisfied, pressure  $P$  equals  $nkT$ , where  $n$  is the number density and  $k$  is the Boltzmann constant and  $T$  the temperature. Then it follows (show!)  $\beta = 1/kT$ .

The reactions that determine the neutron  $n$  proton  $p$  balance are ( $\bar{\nu}$  means “antineutrino”):

$$n + \nu \leftrightarrow p + e$$

$$n + e^+ \leftrightarrow p + \bar{\nu}$$

$$n \leftrightarrow p + e + \bar{\nu}$$

The reaction rates for the two body processes are about 0.1 per second per nucleon at  $T = 10^{10}$ K. However, this rate drops rapidly as the thermometer falls. Below  $10^{10}$ K, the reactions can't keep up with the expansion rate of the Universe, they don't take place fast enough. Once this occurs, whatever the ratio of  $n/p$  happens to be, it remains “frozen” with time, since the Universe is too cold to keep the reactions cooking! At  $T = 10^{10}$ K, the Universe is about 1 s old. The  $n/p$  ratio at this temperature is  $\exp(-\Delta mc^2/kT) \simeq 0.2$ , and it remains stuck at this value, i.e. it is frozen in.

*Exercise.* Show that the early Universe “temperature clock” is given conveniently by

$$t \simeq 1/T_{10}^2,$$

where  $t$  is the time in seconds and  $T_{10}$  the temperature in  $10^{10}$ K. Assume a Universe of photons, electrons and positrons, and three neutrinos and anti-neutrinos.

This figure of 20% is interesting, because it is neither close to unity nor tiny (a coincidence related to the 25% Helium mass fraction!). Without further production of neutrons, they will decay in minutes by the third reaction channel. If all the neutrons had decayed into protons and electrons and antineutrinos, there would have been no cosmological nucleosynthesis.

But just as the neutrons start to decay as we approach  $10^9\text{K}$ , the remaining neutrons are salvaged by being safely packed into stable  $^4\text{He}$  nuclei. Helium nuclei are environmentally sound, perfectly safe, radioactive containment vehicles. Here neutron decay is absolutely forbidden, in essence because the statistical phase space within the realm of the nuclear potential is degenerate. “Degenerate” means that when the neutrons are confined to the nuclear potential, there are no available states for them to decay into. They are already occupied by protons. It is rather like trying to use Oxford public transportation on a very rainy day: SORRY, BUS FULL.

Back to our story. The neutrons are scarfed up into He nuclei. This offers a firm prediction for the observed mass fraction of the Universe, since later stellar nucleosynthesis does not change the Helium number significantly. (By sharp contrast, almost all of the heavier element abundances, collectively not much more than one percent of the total, are indeed dominated by stellar nucleosynthesis.) From what we have just seen, we may estimate the mass fraction in Helium to be

$$\frac{4m_p \times (n/2)}{m_p(n+p)} = \frac{2}{1+p/n} \simeq 0.33. \quad (574)$$

Much more detailed, time-dependent calculations (pioneered by Peebles in 1965) give a number close to 0.27 (a slight refinement of the 25% figure mentioned above), but the essential physics is captured by our simple estimate. There is not very much wiggle room here. We cannot use the precise value of this abundance to determine retrospectively what sort of Universe (open, closed, critical) we live in. The curse and strength of the calculation is that the results are essentially the same for *any* FRW model. The Big Bang Theory predicts something close to 27% of the mass of the Universe is in the form of Helium, and that is that. Happily, this turns out to be very close to what observations reveal.

### 8.10.3 Neutrino and photon temperatures

*This subsection is non-examinable.*

As the temperature slides from  $5 \times 10^9 \text{ K}$  to below  $10^9$ , the electron positron pairs annihilate into photons,

$$e^+e^- \rightarrow 2\gamma$$

The falling temperature means that the photons cannot maintain an equilibrium population of these  $e^+e^-$  pairs. In other words, there is a conversion of electrons and positrons into photons. However, it is not possible to just add photons and keep the temperature the same. In thermal equilibrium, the number of photons at a given temperature is *fixed*. If you add more via  $e^+e^-$  annihilation, the extra photons force a new thermal equilibrium, one at a higher temperature that is compatible with the increased photon number. The photons are, in effect, heated. By contrast, the relativistic neutrinos that are also present just march blithely along, enjoying the expansion of the Universe without a care, oblivious to everything. (When it was the muons that annihilated, the Universe was so dense that even the neutrinos took part in the thermal collisions, but not this time!) This means that present day background neutrinos are cooler than the CMB photons. The question is, by how much?

This turns out to be a relatively simple problem, because if we fix our attention on a comoving volume of the Universe, the entropy in this volume is conserved by the conversion of electron-positron pairs into photons. For photons or relativistic electrons/positrons, the entropy per unit volume is a function only of the temperature. It is most easily calculated from a standard thermodynamic identity,

$$E + PV - TS + \mu\mathcal{N} = 0,$$

for a gas with zero chemical potential  $\mu$ . The latter is true of a relativistic gas that freely creates and annihilates its own particles. So for the entropy per unit volume  $s$ ,

$$s \equiv \frac{S}{V} = \frac{E}{VT} + \frac{P}{T} \quad (575)$$

With  $\rho c^2$  denoting the energy per unit volume and  $P = \rho c^2/3$  for a relativistic gas,

$$s = \frac{4\rho c^2}{3T} \quad (576)$$

With a relativistic energy density always proportional to  $T^4$ , the entropy per unit volume is proportional to  $T^3$ , and the entropy in a comoving volume  $R^3$  is proportional to  $(TR)^3$ . It is conserved with the expansion, in essence because the entropy and particle number are proportional to one another, and particle number is conserved.

Before the  $e^+e^-$  annihilation, the entropy in volume  $R^3$  in photons, electrons and positrons is

$$sR^3 = \frac{4a(TR)^3}{3} \left(1 + 2 \times \frac{7}{8}\right) = \frac{11a(TR)_{before}^3}{3} \quad (577)$$

Afterwards, and after the reestablishment of thermal equilibrium, it is all photons:

$$sR^3 = \frac{4a(TR)_{after}^3}{3} \quad (578)$$

This entire process conserves entropy  $sR^3$ : it conserves the sum of the number of particles in photons plus the number of  $e^+e^-$  pairs. It is therefore reversible. Thus, one could recompress the adiabatic expansion backwards to produce anew the pairs. In other words,

$$\frac{4(TR)_{after}^3}{3} = \frac{11(TR)_{before}^3}{3} \quad (579)$$

Now for an interesting point of physics. Recall that the muon-antimuon annihilation heated *all* relativistic populations, because the reaction rates demand this at these earlier times, when the Universe was so very dense. But the  $e^+e^-$  population annihilation we are now considering ignores the neutrinos. The density has now fallen to a level at which the neutrinos pass through everything, so that their  $TR$  value does not increase during the  $e^+e^-$  annihilation. This process must therefore produce different cosmic background photon and neutrino temperatures. The ratio of neutrino temperature  $T_\nu$  to photon temperature  $T_\gamma$  is

$$\frac{T_\nu}{T_\gamma} = \left(\frac{4}{11}\right)^{1/3} = 0.7138 \quad (580)$$

This would correspond to a current value of  $T_\nu = 1.95$  K — if the neutrinos remained a  $\mu = 0$  relativistic population for all time. (We now know they do not—at least not quite.) At one time, people wondered whether there was any possible way to measure this temperature difference. But a temperature of 1.95K corresponds to an energy  $kT$  of  $1.7 \times 10^{-4}$  eV, whereas the average mass per neutrino species (the best we can measure at the current time) is about 0.1 eV, with a corresponding temperature of 1160 K. Neutrinos thus became “cold” at a redshift of  $z \sim 400$ . As we shall see, this is well after hydrogen recombined, but probably before galaxies formed. These neutrinos are a part, but only a very small part, of the dark matter in galaxies. Even if it *had* turned out that neutrinos had zero rest mass, it would have been impossible with present technology to measure a 2K neutrino background! Nevertheless, the physics of this problem is enlightening, and the formal difference in neutrino and photon temperatures surprising.

### 8.10.4 Ionisation of Hydrogen

Between the time of Helium synthesis, a few minutes into the baby Universe's life, and a few hundred thousand years, almost nothing happens. Radiation remains the dominant source of energy and pressure, and the Universe simply expands with  $R$  scaling like  $t^{1/2}$ . The matter and the radiation remain tightly coupled, so that density fluctuations in the matter correspond also to energy fluctuations in the radiation. But finally at some point, the Universe cools enough that hydrogen recombines and the matter ceases to be an ionised plasma, becoming a neutral gas. At what temperature does this occur? To answer this question, we need to use an interesting sort of variation of the Boltzmann equation, known as the Saha equation. It tells us the ionisation fraction of an atomic element as a function of  $T$ .

Recall the Boltzmann equation. The probability  $P_i$  for an atom to be in state  $i$  relative to the probability  $P_j$  to be in state  $j$  is

$$\frac{P_i}{P_j} = \frac{g_i}{g_j} \exp\left(-\frac{E_i - E_j}{kT}\right) \quad (581)$$

where  $E_i, E_j$  are the energies of the states, and  $g_i, g_j$  are the statistical weights (i.e. the number of states at each level). For simplicity, we will consider a gas of pure hydrogen, and interpret this equation as follows. State  $j$  is the ground state of hydrogen with one electron, the 1s state. Let's set  $j$  equal to 0 so we think of neutrality. State  $i$  is then the ionised state with a bare proton and a free electron with kinetic energy between  $E$  and  $E + dE$ . Set  $i$  equal to 1. Whether bound or free, the electron has two spin states available to it, so this factor of 2 appears in both numerator and denominator, cancelling in the process. Then  $g_0$  is unity. We think of  $g_1$  as the number of states available to the electron with energy in the range  $dE$ . Do you remember how to do this odd-sounding calculation? In terms of momentum  $p$  (a scalar magnitude here), the number of states per electron around a shell of thickness  $dp$  in momentum space is

$$\frac{4\pi p^2}{n_e h^3} dp$$

where the factor  $1/n_e$  represents the volume per electron and  $h$  is Planck's constant. This is derived in any standard text on statistical mechanics, but if this is not familiar, now is a good time to have a look at *Blundell and Blundell, Concepts of Thermal Physics*. The concept comes from the stratagem of putting the electrons in big cube of volume  $V$ , counting the eigenstates (now discrete because of the box walls) for each electron, and then noticing that the artificial box appears in the calculation only as a volume per electron, which is just  $1/n_e$ . Drawing these threads together, (581) becomes, upon adding up all possible free electron states,

$$\frac{n_e n_p}{n_0} = \exp\left(-\frac{\Phi}{kT}\right) \frac{4\pi}{h^3} \int_0^\infty p^2 e^{-p^2/2m_e kT} dp \quad (582)$$

where  $\Phi$  is the ionisation potential of hydrogen,  $m_e$  the electron mass,  $n_e$  the electron density,  $n_p$  the proton density and  $n_0$  the neutral hydrogen density. Note that we have replaced the ratio of relative probabilities of the ionised H state (or  $p$ ) to the neutral H state (subscript 0) by their number densities  $n_p$  and  $n_0$ . Now

$$\int_0^\infty p^2 e^{-p^2/2m_e kT} dp = (2m_e kT)^{3/2} \int_0^\infty x^2 e^{-x^2} dx = (2m_e kT)^{3/2} \times \frac{\pi^{1/2}}{4} \quad (583)$$

The Saha equation becomes

$$\frac{n_e n_p}{n_0} = \left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} \exp\left(-\frac{\Phi}{kT}\right) \quad (584)$$



The final step is to note that for pure hydrogen,  $n_e = n_p$  and thus  $n_e + n_0$  is the total hydrogen density  $n_H$ . This remains unchanged regardless of the ionisation state: if  $n_e$  goes up by 1,  $n_0$  has gone down by 1. With  $x = n_e/n_H$ , our equation therefore becomes

$$\frac{x^2}{1-x} = \frac{1}{n_H} \left( \frac{2\pi m_e kT}{h^2} \right)^{3/2} \exp\left(-\frac{\Phi}{kT}\right) = \left( \frac{2.415 \times 10^{21} T^{3/2}}{n_H} \right) \exp\left(-\frac{1.578 \times 10^5}{T}\right) \quad (585)$$

in MKS units. The value  $x = 0.5$  is attained at a redshift of about 1400, a temperature of 3800 K. This is remarkable, because it is much less than the formal Boltzmann ionisation temperature of  $1.58 \times 10^5$  K. At redshifts less than 1400, the Universe becomes transparent to photons, the energy density is already dominated by matter. Interestingly, the intergalactic medium seems to have been reionised shortly after galaxies were able to form at  $z \sim 10$ , presumably by the very radiation produced by the accretion process that gave rise to these galaxies. To pursue this active area of current astrophysical research would take us too far afield at this point. Now that we have a sense of the basics of Helium nucleosynthesis and hydrogen recombination, we return to the very early Universe. It is there that we will learn about what seems to have been a key process for creating the Universe as we know it: inflation.



*From a small seed a mighty trunk may  
grow.*

— *Aeschylus*

## 9 The Seeds of Structure

### 9.1 The growth of density perturbations in an expanding universe

*This section is optional and off-syllabus.*

The Universe is expanding and the density of nonrelativistic matter decreasing as  $1/R^3$ . In such an expanding background, as the reservoir of material to form condensed objects diminishes so quickly, does gravity even permit the sort of runaway collapse we think of when we envision a star or a galaxy forming?

Determining the fate of a small overdensity or underdensity of nonrelativistic matter in an FRW universe is a problem that can be approached via an analysis of how small disturbances behave in a uniformly expanding background. We require two fundamental equations. The first expresses the conservation of ordinary matter. The mass within a volume  $V$ ,  $\int_V \rho dV$  is changed only if matter flows into or out from the boundaries of  $V$ . The flux of mass (mass per square meter per second) is  $\rho \mathbf{v}$ . Hence

$$\frac{d}{dt} \int_V \rho dV = \int_V \frac{\partial \rho}{\partial t} dV = - \int_{\partial V} \rho \mathbf{v} \cdot d\mathbf{A} = - \int_V \nabla \cdot (\rho \mathbf{v}) dV$$

where  $\partial V$  is the volume's boundary, and we have used the divergence theorem. The volume  $V$  is arbitrary, so we must have

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (586)$$

which is the equation of mass conservation. Newton's equation of motion states that if a mass element of fluid  $\rho dV$  is accelerating, then it is acted on by a gravitational force given by  $-\rho dV \nabla \Phi$ , where  $\Phi$  is the associated potential function. In other words, the force equation reads after cancellation of  $\rho dV$ ,

$$\left[ \frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla) \right] \mathbf{v} = -\nabla \Phi \quad (587)$$

Note that the acceleration measured relative to a fixed space-time coordinate background means that the “total time derivative” must be used,  $\partial_t + v_i \partial_i$  in index notation.

The local time behaviour of the density is entirely Newtonian. The expansion of the universe is described by

$$\mathbf{v} = \frac{\dot{R}}{R} \mathbf{r} \quad (588)$$

a familiar Hubble law. The mass equation then becomes

$$\frac{\partial \rho}{\partial t} + \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^3 \rho \frac{\dot{R}}{R} \right) = 0 \quad (589)$$

With the background  $\rho$  independent of position,

$$\frac{d \ln \rho}{dt} + \frac{3\dot{R}}{R} = 0 \quad (590)$$

whence  $\rho R^3$  is a constant, as we know. As for the force equation, a straightforward calculation yields

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = \frac{\ddot{R}}{R} \mathbf{r} \quad (591)$$

and, as the discussion of §8.1.1 shows that  $\ddot{R}/R = -4\pi G\rho/3$ , our solution for the background expansion is correct. The Poisson equation

$$\nabla^2 \Phi = \frac{1}{r} \frac{\partial^2 (r\Phi)}{\partial r^2} = 4\pi G\rho \quad (592)$$

is likewise solved by our solution. (Try  $\Phi = 2\pi G\rho r^2/3$ .)

We are interested in the fate of small disturbances  $\delta\rho$ ,  $\delta\mathbf{v}$  and  $\delta\Phi$  expressed as small additions to this background:

$$\rho \rightarrow \rho + \delta\rho, \quad \mathbf{v} \rightarrow \mathbf{v} + \delta\mathbf{v}, \quad \Phi \rightarrow \Phi + \delta\Phi \quad (593)$$

Using our fundamental equations, we replace our dynamical variables as shown, and because the  $\delta$ -quantities are small, we retain them only through linear order. We ignore quadratic and higher order terms. The mass equation is then

$$\frac{\partial \delta\rho}{\partial t} + \nabla \cdot (\rho \delta\mathbf{v}) + \nabla \cdot (\mathbf{v} \delta\rho) = 0 \quad (594)$$

Note that the gradient of  $\rho$  vanishes and that the equilibrium  $\mathbf{v}$  satisfies  $\nabla \cdot \mathbf{v} = -\partial_t \ln \rho$ . It is then straightforward to show that the perturbed linearised mass conservation equation simplifies to

$$\left[ \frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla) \right] \frac{\delta\rho}{\rho} + \nabla \cdot \delta\mathbf{v} = 0 \quad (595)$$

The linearised equation of motion

$$\left[ \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \delta\mathbf{v} + (\delta\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla \delta\Phi \quad (596)$$

becomes

$$\left[ \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \delta\mathbf{v} + \frac{\dot{R}}{R} \delta\mathbf{v} = -\nabla \delta\Phi \quad (597)$$

Next, we change to comoving coordinates. This is straightforward. Let  $\mathbf{r} = R(t')\mathbf{r}'$  and  $t = t'$ . Then  $\mathbf{r}'$  (or  $x'_i$  in index notation) is a comoving spatial coordinate. The partial derivative transformation is (sum over repeated  $i$ ):

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t'} + \frac{\partial x'_i}{\partial t} \frac{\partial}{\partial x'_i}, \quad \nabla = \frac{1}{R} \nabla', \quad \frac{\partial x'_i}{\partial t} = -\frac{x_i}{R^2} \dot{R} = -\frac{v_i}{R}, \quad (598)$$

so that

$$\frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla) = \frac{\partial}{\partial t'}, \quad (599)$$

which is a time derivative following a fluid element of the unperturbed expansion. Then, our two equations for mass conservation and dynamics are

$$\frac{\partial}{\partial t'} \frac{\delta \rho}{\rho} + \frac{1}{R} \nabla' \cdot \delta \mathbf{v} = 0 \quad (600)$$

$$\frac{\partial \delta \mathbf{v}}{\partial t'} + \frac{\dot{R}}{R} \delta \mathbf{v} = -\frac{1}{R} \nabla' \delta \Phi \quad (601)$$

Taking  $\nabla' \cdot$  of (601) and using (600) leads to

$$\frac{\partial^2}{\partial t'^2} \frac{\delta \rho}{\rho} + \frac{2\dot{R}}{R} \frac{\partial}{\partial t'} \frac{\delta \rho}{\rho} = \frac{1}{R^2} \nabla'^2 \delta \Phi = 4\pi G \rho \left( \frac{\delta \rho}{\rho} \right) \quad (602)$$

where the last equality is the linearised Poisson equation  $\nabla^2 \delta \phi = 4\pi G \delta \rho$ . For an Einstein-de Sitter universe, recall that  $\rho = 1/6\pi G t^2$  and  $\dot{R}/R = 2/3t$ , where we have dropped the primes. Using the notation  $\delta = \delta \rho / \rho$  and a dot  $\dot{\delta}$  for a time derivative, our differential equation for the growth of small perturbations in an Einstein de-Sitter universe takes on a very elegant form:

$$\ddot{\delta} + \frac{4}{3t} \dot{\delta} - \frac{2}{3t^2} \delta = 0 \quad (603)$$

This differential equation has two very simple linearly independent solutions, one where  $\delta$  decays as  $1/t$ , the other where it grows as  $t^{2/3}$ . (Show!) The important point is that because of the background expansion, the perturbations display none of the rapid growth one usually finds against a static background. The growing solution of the small perturbation is  $t^{2/3}$ , growing no faster than the universe itself expands. This is a pretty torpid tempo. For the musically inclined, think *adagio*.

How our Universe grew both its large and small scale structures has long been a great mystery, one that remains far from well-understood. To make things grow in the barren soil of the Universe, one needs to start out with very healthy-sized seeds<sup>17</sup>. The questions of where those might come from are the topics of the next section.

## 9.2 Inflationary Models

### 9.2.1 “Clouds on the horizon”

We begin by posing two profound mysteries associated with classical FRW universes: the horizon problem and the flatness problem.

We have already encountered the first, the horizon problem, on page (116). The CMB is homogeneous to 1 part in  $10^4$  on all angular scales, yet the angular size of the horizon at the redshift of hydrogen recombination is of the order of the diameter of the full moon. How can we possibly understand this degree of homogeneity between regions that have never been in causal contact?

The second problem is known as the flatness problem. Consider the dynamical equation of motion in the form

$$1 - \frac{8\pi G \rho}{H^2} \equiv 1 - \Omega = \frac{2E}{\dot{R}^2} \quad (604)$$

---

<sup>17</sup>One also needs a great deal of Dark Matter, a topic we must alas leave untouched.

Now if  $E$  happens to be zero,  $\Omega$  is unity for all time. Fair enough. But the measured value of  $\Omega$ , at least in terms of ordinary matter, was thought to be a number of order 25%, including unseen dark matter, with only 5% for ordinary baryons. This is decidedly smaller than unity, but not infinitesimally small. Now normally either  $\Omega$  is very close to unity if  $E$  is small and  $\dot{R}$  large, or  $\Omega$  is very small indeed, proportional to  $1/R\dot{R}^2$  throughout most of the vast history of a matter dominated universe. But  $\Omega$  passes through “a number less than but not very different from unity” during a tiny, fleeting moment of a universe’s history. And this is the period we just so happen to be observing it? That certainly is a coincidence. We don’t like coincidences like that.

To see how the concept of inflation can resolve both of these problems, consider the integral that is done on page (116) to calculate the horizon distance. The problem is that the horizon length at cosmic time  $t$  is very finite, of order  $c/H(t)$ . Formally, this distance is proportional to

$$\int \frac{dt'}{R(t')} = \int \frac{dR}{R\dot{R}} \quad (605)$$

Imagine that at very small  $R$ , the dominant behaviour of  $\dot{R}$  is  $\dot{R} \sim R^p$ , where  $p$  is some number. If  $p \geq 0$ , then the integral diverges like  $\ln R$  or  $R^{-p}$  at small  $R$ , and in this case divergence is good: it is *what we want*. Then the horizon problem goes away because the horizon is unbounded! In essence a small patch of universe, small enough to communicate with itself completely, can rapidly grow to encompass an arbitrarily large segment of sky. With  $p \geq 0$ , then  $\ddot{R} \geq 0$  at small  $R$  so that the universe would be accelerating (or just not decelerating). The problem with standard models is that they are radiation-dominated,  $p = -1$ , and highly decelerating. A matter-dominated universe,  $p = -1/2$ , is no help.

What instead appears to have happened is that the Universe, early in its history, went through a phase of exponential expansion with  $\dot{R} \sim R$ ,  $p = +1$ . As we have seen, the Universe has begun such an inflationary period recently, at redshifts of order unity. Exponential expansion is the hallmark of a vacuum energy density  $\rho_V$ , with a corresponding pressure  $P_V = -\rho_V c^2$ . This rapid expansion makes an entire Universe from a once very tiny region that was in complete causal self-contact<sup>18</sup>. The rapid large expansion also has the effect of killing off the  $2E/\dot{R}$  term in equation (627). In other words it resolves the flatness problem by, well, flattening the Universe! Think of being on patch of sphere and then having the radius expand by an *enormous* factor. The new surface would look very flat indeed. Dynamically, (627) shows us that  $\Omega$  must then be equal to unity to great accuracy. This is just what observations show.

At this point the student may wish to move on to the closing remarks of §X. The next subsection is an overview of the modern ideas of inflation, but is entirely of syllabus. I hope that if it is impractical to read this now, you will return to it at your leisure at some future point.

## 9.2.2 The stress energy tensor of a field

*This subsection is optional and off-syllabus, containing advanced material.*

Can we make a case for early inflation? If we are to understand the physics of the vacuum quantitatively, we need to learn a little about quantum field theory, the domain of physics where the vacuum makes a starring appearance in a leading role.

---

<sup>18</sup>Note that the rest of the original universe is still hanging around! Inflationary models therefore lead naturally to the concept of a “multiverse,” which in itself would help us to understand many otherwise mysterious cooperations between physical scales.

Let us start easy, with spin 0 particles. Spin 0 particles, so-called scalar fields (“scalar” because they have only one component), satisfy the Klein-Gordon equation. In Minkowski spacetime, this is

$$\square\Phi - \mu^2\Phi = 0, \quad \mu^2 = (mc/\hbar)^2 \quad (606)$$

where as usual  $\square = \partial^\alpha\partial_\alpha$ ,  $m$  is the mass of the particle (the “quantum of the field”), and  $\hbar$  is Planck’s constant over  $2\pi$ . Now when we solve the Einstein Field Equations, we need to know what the stress energy tensor  $T^{\mu\nu}$  is. For ordinary stuff, this is not a problem:

$$T^{\mu\nu} = Pg^{\mu\nu} + (\rho + P/c^2)U^\mu U^\nu$$

The question is, what is the  $T^{\mu\nu}$  for a *field* obeying equation (606)? Is there something we can identify as the energy density? The pressure? Once we have a  $T^{\mu\nu}$  from a field equation, there is nothing to prevent us from treating it on an equal footing with the usual  $T^{\mu\nu}$  given above, and using it as source term in the Einstein Field Equations.

We are not entirely in the dark on how to proceed. Indeed, there is light. Literally. Electrodynamics is a field theory. There is both an energy density in the electromagnetic field and a corresponding pressure. In short, there is a perfectly good stress tensor available for whatever legitimate use we would like. We may certainly use this stress tensor on the right side of the Einstein Field Equations. With the help of our “4-potential”  $A_\alpha$  (space components equal to the usual vector potential  $\mathbf{A}$  and time component equal to minus the electrostatic potential), we first define the field tensor  $F_{\alpha\beta}$ :

$$F_{\alpha\beta} = \frac{\partial A_\beta}{\partial x^\alpha} - \frac{\partial A_\alpha}{\partial x^\beta} \quad (607)$$

and then the stress tensor is

$$T^{\alpha\beta} = F^\alpha_\gamma F^{\beta\gamma} - \frac{1}{4}\eta^{\alpha\beta} F_{\gamma\delta} F^{\gamma\delta}. \quad (608)$$

(The texts of Jackson [1998] or W72 are good references if needed.) There is in fact only one conserved tensorial combination that is quadratic in the derivatives of the 4-potential. You’re looking at it. The overall normalisation constant can be determined by looking at the interaction between the fields and the particles via the Lorentz equation of motion. (More precisely the work done by the fields upon the particles.) A tensor quantity that is quadratic in the derivatives of the potentials and is conserved: these are the essential features we seek.

With this thought in mind, it is a surprisingly simple matter to find the  $T_{\alpha\beta}$  for the Klein-Gordon field, and even to pick out its corresponding  $\rho$  and  $P$ . Multiply (606) by  $\partial_\beta\Phi$ . For ease of future generality, let’s call the  $\mu^2$  term  $dV(\Phi)/d\Phi$ , and refer to it as “the potential derivative”. For the K-G equation,  $V = \mu^2\Phi^2/2$ . We will shortly consider other forms. The equation becomes:

$$\frac{\partial\Phi}{\partial x^\beta} \frac{\partial^2\Phi}{\partial x^\alpha\partial x_\alpha} = \frac{dV(\Phi)}{d\Phi} \frac{\partial\Phi}{\partial x^\beta} = \frac{\partial V}{\partial x^\beta} = \frac{\partial(\eta_{\alpha\beta}V)}{\partial x_\alpha} \quad (609)$$

Integrate the left side by parts:

$$\frac{\partial\Phi}{\partial x^\beta} \frac{\partial^2\Phi}{\partial x^\alpha\partial x_\alpha} = \frac{\partial}{\partial x_\alpha} \left( \frac{\partial\Phi}{\partial x^\alpha} \frac{\partial\Phi}{\partial x^\beta} \right) - \frac{\partial\Phi}{\partial x^\alpha} \frac{\partial^2\Phi}{\partial x^\beta\partial x_\alpha} \quad (610)$$

But the final term of (610) is

$$- \frac{\partial\Phi}{\partial x^\alpha} \frac{\partial^2\Phi}{\partial x^\beta\partial x_\alpha} = -\frac{1}{2} \frac{\partial}{\partial x_\alpha} \left( \eta_{\alpha\beta} \frac{\partial\Phi}{\partial x^\gamma} \frac{\partial\Phi}{\partial x_\gamma} \right) \quad (611)$$

Putting the last three equations together:

$$\frac{\partial T_{\alpha\beta}}{\partial x_\alpha} = 0, \quad \text{where} \quad T_{\alpha\beta} = -\eta_{\alpha\beta} \left[ \frac{1}{2} \frac{\partial \Phi}{\partial x^\gamma} \frac{\partial \Phi}{\partial x_\gamma} + V(\Phi) \right] + \left( \frac{\partial \Phi}{\partial x^\alpha} \frac{\partial \Phi}{\partial x^\beta} \right), \quad (612)$$

with frame-independent trace  $T = -(\partial_\alpha \Phi)(\partial^\alpha \Phi) - 4V$ .  $T_{00}$  may be read off directly:

$$T_{00} = \rho c^2 = \frac{1}{2c^2} \dot{\Phi}^2 + \frac{1}{2} |\nabla \Phi|^2 + V, \quad (613)$$

which really does look like an energy density, but corresponds to the true rest frame  $\rho c^2$  *only* in the rest frame of the field! The first term is a kinetic energy density (the dot as usual means time derivative), the second is the effective potential energy density from the spring-like coupling that produces the simple harmonic motion of  $\Phi$ , and the final term is the *external* potential from an external driver. In textbooks on quantum fields, Chapter 1 often begins by examining the quantum mechanics of a collection of masses on springs, because that problem is not just similar to, but is practically identical with, the problem of the excitation of massless spin 0 particles. The Hamiltonian density found for this analogue mechanical problem is precisely our expression (613), where  $\Phi$  corresponds to the mass displacement.

To extract the energy density and pressure, begin by recalling the quantum mechanical association between the derivative operator  $\partial_\alpha$  and the 4-momentum  $p_\alpha$ . (Don't worry about any constant coefficients, they will ultimately cancel.) In equation (612), the final term combination  $\partial_\alpha \Phi \partial_\beta \Phi$  should then be identified with the  $U_\alpha U_\beta$  term in our classical stress tensor, which is  $(\rho + P/c^2) U_\alpha U_\beta$ . To turn the  $p_\alpha$  terms into 4-velocity  $U_\alpha$  terms, divide and multiply by the effective squared rest mass scalar " $E^2 - p^2 c^2$ ". This is not a matter of inserting a constant factor of  $m^2 c^4$  somewhere. We must express everything in terms of the field  $\Phi$ ! The mass expression we require is then  $-c^2 \partial_\gamma \Phi \partial^\gamma \Phi$  (Why, what does  $\partial_t$  correspond to?). To go from 4-momentum to 4-velocity is a simple sleight-of-hand:

$$\partial_\alpha \Phi \partial_\beta \Phi = \frac{c^2 \partial_\alpha \Phi \partial_\beta \Phi}{-c^2 \partial_\gamma \Phi \partial^\gamma \Phi} \times \frac{1}{c^2} (-\partial_\gamma \Phi \partial^\gamma \Phi). \quad (614)$$

Now read off the 4-vector(s), and then  $\rho + P/c^2$  must be the final multiplicative field scalar:

$$U_\alpha = c(\partial_\alpha \Phi) (-\partial_\gamma \Phi \partial^\gamma \Phi)^{-1/2}, \quad \rho + \frac{P}{c^2} = \frac{1}{c^2} (-\partial_\gamma \Phi \partial^\gamma \Phi). \quad (615)$$

The pressure  $P$  is just the coefficient of  $\eta_{\alpha\beta}$  in (612):

$$P = - \left( \frac{1}{2} \partial^\gamma \Phi \partial_\gamma \Phi + V(\Phi) \right) \quad (616)$$

Then, from the second of equations in (615):

$$\rho c^2 = - \left( \frac{1}{2} \partial^\gamma \Phi \partial_\gamma \Phi - V(\Phi) \right) \quad (617)$$

Remember, these are the rest frame pressure and the rest frame energy density, expressed in a form that is valid for any frame. Notice the difference between  $T_{00}$  in (613) and  $\rho c^2$  in (617)! When do the two expressions agree with one another?

The idea now is to upgrade from  $\eta_{\alpha\beta}$  to  $g_{\mu\nu}$  as per the usual GR prescription, and then use this quantum form of  $T_{\mu\nu}$  in the classical spacetime cosmological equations during the earliest phase of the Universe. But why go through the trouble of picking out  $\rho$  and  $P$ , why not just work directly with the stress tensor itself? The reason for explicitly identifying  $\rho$  and  $P$  is that it becomes a relatively easy matter to analyse the “quantum cosmological field equations” by analogy to the classical case.

The fundamental dynamical equation in the absence of curvature is

$$H^2 = \frac{8\pi G\rho}{3} \quad (618)$$

where  $H = \dot{R}/R$ . We need a second equation to know how  $\rho$  (and  $P$ ) depend upon  $R$ . This is the energy conservation equation, (524):

$$\dot{\rho} + 3H \left( \rho + \frac{P}{c^2} \right) = 0. \quad (619)$$

Remember that this equation comes from  $\partial^\mu T_{\mu\nu} = 0$ . But this amounts to solving the Klein-Gordon equation itself, since the way we formed our stress tensor (612) was by contracting the K-G equation with a 4-gradient of  $\Phi$ . So all the relevant equations are embodied in (618) and (619), with (617) and (616) for  $\Phi$ .

What is this  $V(\Phi)$ ? It helps to have a concrete mechanical model. If I have a one-dimensional collection of masses on springs (we can even make the masses out of concrete), and  $\phi_n$  is the lateral displacement of mass  $n$ , the equation of motion for spring constant  $k$  is:

$$\ddot{\phi}_n = -k(\phi_n - \phi_{n-1}) - k(\phi_n - \phi_{n+1}) \quad (620)$$

Now when  $n$  is very large and the masses are closely spaced with a small separation  $\Delta x$ , I can take the limit

$$\phi_n - \phi_{n-1} \simeq \Delta x \phi'_{n-1/2}$$

so that

$$-k(\phi_n - \phi_{n-1}) - k(\phi_n - \phi_{n+1}) \rightarrow -k\Delta x(\phi'_{n-1/2} - \phi'_{n+1/2}) \quad (621)$$

where the spatial  $x$  derivative  $\phi'$  is formally defined halfway between the integer  $n$ 's. A second use of this limit brings us to

$$-k\Delta x(\phi'_{n-1/2} - \phi'_{n+1/2}) = k(\Delta x)^2 \phi''_n \quad (622)$$

or

$$\frac{\partial^2 \phi}{\partial t^2} = c^2 \frac{\partial^2 \phi}{\partial x^2} \quad (623)$$

where  $c^2 = k(\Delta x)^2$  becomes the velocity of a propagating  $\phi$  disturbance as  $k$  gets large and  $\Delta x$  small. A three-dimensional extension of this argument would introduce nothing new, so this is a mechanical analogue of the standard wave equation. But this is not yet the Klein-Gordon equation: where is  $V$ ?

Do the same problem, but this time hang the masses from strings of length  $l$  in a vertical gravitational field  $g$ , while they slide back and forth on their connecting springs, as in figure [18]. Then, our final partial differential equation becomes

$$\frac{\partial^2 \phi}{\partial t^2} = c^2 \frac{\partial^2 \phi}{\partial x^2} - \frac{g}{l} \phi \quad (624)$$

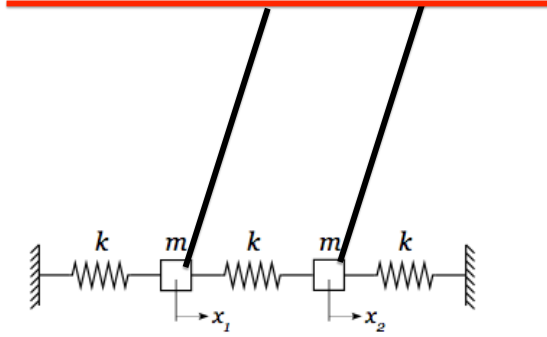


Figure 18: Mechanical analogue of the Klein-Gordon equation. Spring-like coupling between adjacent masses gives rise to the wave equation, while the pendula produce an acceleration directly proportional to the displacement, not the  $\partial_x$  derivative thereof. This  $g/l$  force is exactly analogous to the mass term in the KG equation, which evidently arises from similar “external couplings” of the scalar field  $\Phi$ . (Only two masses are shown instead of an infinite continuum.)

This is the Klein-Gordon equation with our potential  $V$  term. The final term is an external coupling between the displacement  $\phi$  (or the “field”  $\Phi$ ) and some external interaction. Note: an *interaction*. In the K-G equation, the coefficient embodying this interaction  $g/l$  is called “mass.” This sets up Richard Feynman’s famous quotation: “*All mass is interaction.*” This kind of loading is the price you have to pay to jiggle the masses on the springs, or to jiggle the field, when there is some kind of external coupling present. The mass of electron arises, at least in part, from the fact that when you try to accelerate it, you produce disturbances in the photon field and the charge radiates.

We could imagine putting our concrete masses in some kind of a yet more complicated external force. Maybe  $g\phi/l$  is only the first term in a Taylor series of, say,  $g \sin \phi/l$ . (Can you think of simple mechanical system that would, in fact, have this property?) The point here is that as long as the equilibrium  $\phi = 0$  null displacement state is one that is stable, any small deviation from  $\phi = 0$  will be generically linear in the interaction, and  $V(\phi)$  (or  $V(\Phi)$ ) *quadratic*. This is why we think of  $V$  as a sort of potential function. We speak casually of the field being a “ball sitting at the bottom of a potential well.” The rollback oscillation frequency then becomes the mass of the field particle!

So in our example, what is giving rise to this so-called interaction? Where is the external pendulum force coming from that is affecting the  $\Phi$  field? The answer is that comes from the vacuum itself.

### *Effective Potentials in Quantum Field Theory*

One thing you have to understand about the quantum vacuum: *it is a jungle out there.*

The vacuum, *even* the vacuum, is *full* of fluctuations in the varied collection of harmonic oscillators that we are pleased to call particles. Depending upon our precise circumstances,



a simple Klein-Gordon equation, which would represent completely free massive particles, doesn't capture the dynamics of the scalar field in question. The scalar field could interact with all these other fields, and since interaction is mass, in the process create a new mass coefficient. We could easily imagine a non-linear “pendulum,” coupling to other fields. Then, the mass constant  $\mu^2$  is not, in fact, a constant, but would depend on the field strength  $\Phi$ . In the simplest case,  $\mu^2$  would depend additively upon  $\Phi^2$ , so that only the magnitude of  $\Phi$ , not its sign, affects the distortion of  $\mu^2$ . Then, the resulting potential  $V(\Phi)$  would take the form

$$V(\Phi) = \mu^2 \Phi^2 / 2 + \beta \Phi^4 \quad (625)$$

In fact, for decades before the notion of the inflation became popular, precisely this model potential had been in wide use amongst particle physicists for completely different reasons. If for some reason  $\mu^2$  was, despite its mathematical form, a *negative* quantity, and  $\beta$  positive, then  $\Phi = 0$  is not a stable vacuum solution. The true, stable vacuum state would be the global minimum,  $\Phi^2 = |\mu^2 / 4\beta|$ . With respect to disturbances about this displaced minimum, the potential once again appears stable-quadratic, and we are back to (to leading order) to the Klein-Gordon Equation. But the mass coefficient has changed! For understanding the dynamics of the early Universe, it is not so much the location of the new equilibrium that is important, it is the journey we take to reach it: down the potential slide! This is from whence inflation may originate—assuming, of course, that we start off at or near the top of the slide.

Try not to be put off by the vagueness of all this. This is forefront physics and we are groping a bit. As Mark Twain once quipped about Wagnerian music, it is better than it sounds. The real world's  $V(\Phi)$  may well be a complicated function (lots of interactions), but we have long ago landed in a stable local parabolic minimum of  $V$ , so to us things now look deceptively, Klein-Gordonly, simple. It is OK to grope in the dark a bit to try to understand what new physics might in principle lie beyond the usual theories and to see what the most robust predictions might be. But within reason, of course. As long as certain ground rules like Lorentz invariance are respected, there is considerable freedom in choosing the form of the interaction. As noted, the type of quartic potential (625) was already well-known to particle physicists and in-play at the time it was appropriated by cosmologists. The particle physicists had borrowed it from condensed matter physicists, who had in turn used it (in the form of a *thermodynamic* potential) to describe ferromagnets. This so-called broken symmetry approach has proven to be enormously useful in particle physics. (The Higgs Boson, for example.) Just as a ferromagnet can spontaneously magnetise itself, so too can certain types of particle spontaneously acquire mass. These physically distinct processes turn out to share very similar mathematics.

In a ferromagnet, the equilibrium state minimises the Gibbs Free Energy  $G$ , as is usual for thermodynamical systems at fixed temperature and pressure.  $G$  is a function of the magnetisation  $M$ , and typically takes the form of (625):

$$G = \alpha M^2 + \beta M^4 \quad (626)$$

At high temperatures, above some critical  $T_c$ , both  $\alpha$  and  $\beta$  are positive. Below  $T_c$ , however,  $\alpha$  changes sign, and it becomes energetically more favourable —  $G$  attains a smaller value — when  $M^2$  is finite and equal to  $-\alpha/2\beta$ . The system, in other words, becomes spontaneously magnetised. A similar mathematical arrangement also produces the phenomenon of superconductivity. And in particle physics, this is the core of the argument for how a class of field particles acquires mass under conditions that are otherwise mysterious.

### *The slow roll inflationary scenario*

The essence of the so-called slow roll inflationary is easily grasped. Start by using (613)

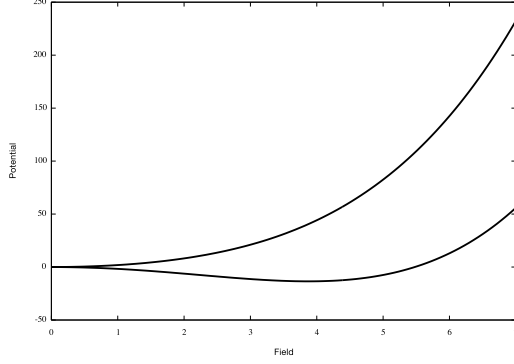


Figure 19:  $V(\Phi)$  potential functions of the form (625). The upper curve corresponds to  $\mu^2 > 0$ , the lower to  $\mu^2 < 0$ . If there is a scalar field described by the  $\mu^2 < 0$  potential in the early Universe, it could trigger an episode of inflation. Before cosmological applications, these potential functions were used to describe ferromagnets, and to explain how symmetry is broken in fundamental particle physics.

and (616) in (618) and (619). This leads to the equations

$$H^2 = \frac{8\pi G}{3c^2} \left( V + \frac{\dot{\Phi}^2}{2c^2} \right) \quad (627)$$

and

$$\ddot{\Phi} + 3H\dot{\Phi} + c^2 \frac{dV}{d\Phi} = 0 \quad (628)$$

The idea is that the gross form of the potential  $V(\Phi)$  itself changes while the Universe expands and cools, going from the top form of figure (19) early on to the bottom form as things cool, much as a ferromagnet's free energy does when the temperature changes from  $T > T_c$  to  $T < T_c$ . Moreover, if  $V$  is very flat, so that  $dV/d\Phi$  is small, then  $\dot{\Phi}$  is also small by equation (628) and there is an extended period when (627) is simply

$$H = \frac{\dot{R}}{R} \simeq \sqrt{\frac{8\pi G V_0}{3c^2}} \quad (629)$$

where  $V_0$  is the (approximately) constant of  $V(\Phi)$ .  $R$  grows exponentially,

$$R \propto \exp \left( \sqrt{\frac{8\pi G V_0}{3c^2}} t \right) \quad (630)$$

and the Universe enters its inflationary phase.  $\Phi$  meanwhile grows slowly, but it does grow, and eventually, after many e-foldings, the inflation stops when the minimum of  $V(\Phi)$  is reached. It was Alan Guth, a particle physicist, who put together this picture in 1980, and brought to the fore the concept of inflation as a phase of the history of the early Universe. In particular, he argued that the peculiar model potentials then in widespread use to understand ferromagnets and symmetry breaking in particle physics, might also be relevant to fundamental problems in cosmology.

*Exercise.* The slow-roll equations of inflationary cosmology, from (627) and (628), are

$$H^2 = \left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi GV}{3c^2}, \quad \dot{\Phi} = - \left(\frac{c^5}{24\pi GV}\right)^{1/2} \frac{dV}{d\Phi}.$$

What are the conditions for their validity? Next, try a potential of the form

$$V(\Phi) = V_0(1 - \epsilon^2\Phi^2)^2,$$

where  $\epsilon$  is small and  $V_0$  constant. Plot  $V$  as a function of  $\Phi$ . Show that this  $V$  satisfies the slow-roll constraints, and solve the above differential equations for  $R(t)$  and  $\Phi(t)$  with no further approximations.

The theoretical arguments for the mechanism of inflation are not based on fundamental theory (at least not yet). They are what physicists call “phenomenological.” That means they are motivated by the existence of an as yet unexplained phenomenon, and rely on the detailed mathematical exploration of a what-if theory to see how the ideas might lead to the behaviour in question. Perhaps the theory, if framed carefully, will explain something it wasn’t specifically designed to do. That would be encouraging! Inflation models have this property, which is why there are very attractive.

Here is an example. During the period of inflation, small fluctuations go through two types of behaviour in sequence. At first, they oscillate, like a sound wave. But as the Universe rapidly expands, at some point the peak of a wave and the trough of a wave find themselves outside of each other’s horizons! A wave can’t possibly oscillate coherently under those conditions, so the disturbance remains “frozen” with the expansion.

But the rapid inflation eventually slows while the Universe is still practically a newborn baby. Then, the expansion no longer is accelerating, but decelerating. It does not take very long before the ever expanding horizon scale can once again enclose a wavelength. In the parlance of cosmology, the wave “enters” the horizon, and the oscillation can then restart! Suppose that between this (very early) moment of restart when the oscillation starts off with zero velocity, and the time of hydrogen recombination (several hundred thousand years later), when we see the imprint of the fluctuation in the radiation, there is one full contraction of the oscillation. (Or, one full expansion, if the disturbance starts on its expanding phase.) This half-period of oscillation corresponds to a particular wavelength, and for this particular wavelength we would expect to see a peak when we plot the spectrum of fluctuations as a function of wavelength. And then another peak for the shorter wavelength that would allow a compression followed by a complete expansion. And so on. The so-called power spectrum<sup>19</sup> Figure (19) shows a sequence of peaks on certain angular scales on the sky. In order to exist as well-defined entities, instead of just a smear, these peaks need to have oscillations recommence at nearly the same time. This is possible only because of inflation. During inflation, the rapid expansion prevents the random oscillations that would otherwise be present. It then allows the dance to begin again from zero velocity. This happens very early on, once the rapid expansion has slowed and the oscillations enter the horizon. The coherent release at only slightly different (shorter wavelengths enter the horizon a bit sooner!) early times is what makes the peaks possible in the first place. Their very existence is powerful evidence for the concept of inflation.

Let us recap. The concept of an early inflationary period of the Universe explains both why there is such uniformity in the CMB temperature across the sky and why the Universe is flat. *It explains why there is an FRW metric at all!* But there is subtlety as well in the

---

<sup>19</sup>If  $T(\theta, \phi)$  is the CMB temperature as a function of angle (angular separation), and  $\mathcal{T}$  is the usual fourier transform in wavenumber space, the power spectrum is  $|\mathcal{T}|^2$ .

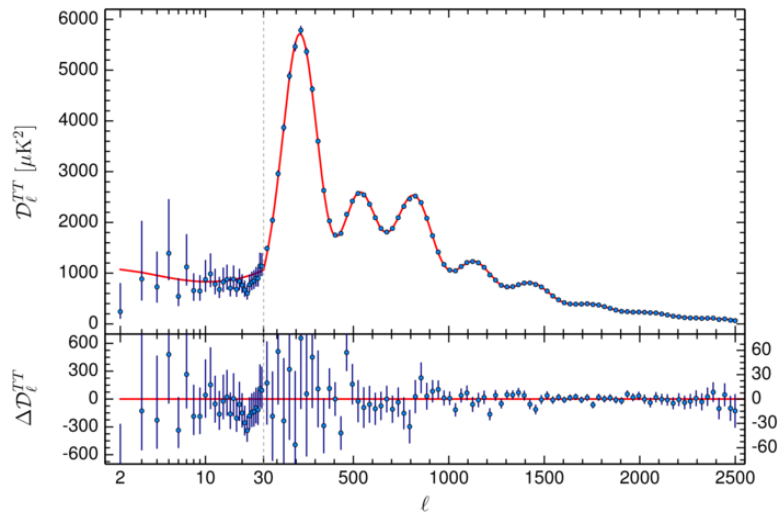


Figure 20: The temperature power spectrum of fluctuations (known as “TT spectrum” in the literature) as a function of angular scale from recent Planck data. (The bottom plot is a measure of the discrepancy between the data and a model, and does not concern us here.) The excellent model fit in the top plot is strong evidence for inflation.

predictions of inflation, including statistical predictions of where the power spectrum should have its peaks. There is so much more that, alas, we don't have time to go into in this course. There are, for example, firm predictions for how large scale structure in the Universe evolves—how galaxies cluster—from the initial seed fluctuations. (Quantum fluctuations!) These predictions rely on the notion of *Cold Dark Matter* taking part in the gravitational response to the seeds. Cold Dark Matter (CDM) is cold in the sense that it responds readily to gravitational perturbations, forming the bulk of the mass distribution in galaxies and clusters. Though we don't know very much about CDM, there are reasons to think that it may be some kind of weakly interacting massive particle, and there are searches underway to try to detect such particles via their (perhaps) more easily found decay products<sup>20</sup>. But it all must begin with some kind of inflationary process. There really is no other explanation for how vast stretches of the Universe could ever have been in causal contact. Inflation is a powerful, unifying concept without which we can not make sense of even the most basic cosmological observations. And don't forget: the Universe is inflating right now! We are living through mild inflation that will, with time, become much more dramatic.

But from whence inflation? Why? What is the underlying imperative to inflate, and then exit gracefully from inflation? While there are some promising ideas afoot, we still don't really know how and why inflation occurred. Maybe somebody reading these notes will settle the matter.<sup>21</sup>

### 9.3 A Final Word

Astrophysics can be a very a messy and speculative business. But every once and a while, something truly outstanding is accomplished. The development of stellar structure and evolution is one such triumph. This led to a new field of science: nuclear astrophysics, and ultimately a precision theory for the origin of all the chemical elements. We now understand where atoms come from, and even how to make them ourselves, a stunning achievement. Another milestone is the full blossoming of the theory of black holes, brilliantly confirmed in the last year by the LIGO detections of gravitational radiation from a number of merging black hole binaries. Surely the development of precision cosmology, the discovery and construction of a model of the Universe, must also rank as one of the great advances in science. I view this on a par with the Rutherford nuclear atom, Gödel's theorem, or the Crick-Watson DNA model: not just a technical advance, but a transformative understanding. We have taken the measure of the entire Universe, and the vacuum is a full player. What we thought of a generation ago as the stuff of the Universe is foam on a sea, less than 5% of the true stuff of the Universe. The Universe started out dominated by its vacuum, and it will spend most of its life dominated by its vacuum. Let that sink in. The time when matter and radiation have much of a say in regulating the Universe will be a fleeting instant in its history, one which we are only just now emerging from. We ensembles of ordinary matter are the condensed fumes from the Universe's exhaust vapour, waste products of the vacuum-driven engine. Even that doesn't begin to capture our insignificance. The concept of inflation suggests that an incomprehensibly vast multiverse is a viable description of the true reality. But we can take heart, all is not bleak. Important questions remain for us, tiny restless pieces of a Universe that has produced a recursive understanding of itself. Thank goodness for that. Along the way, we have deduced our true age, our dynamics, and a great deal of our history. I think it very unlikely that very many universes are fortunate enough to have internal bits that are able to justify such a remarkable and outlandish claim.

---

<sup>20</sup>Update: No luck so far with the searches! An alternative view that black holes may be the dark matter is now starting to be taken seriously. We'll see.

<sup>21</sup>I'm quite serious: given the historical track record, when this problem is solved, it would not be at all surprising to me if it is by someone who was an Oxford undergraduate.

End of notes April 3, 2018.