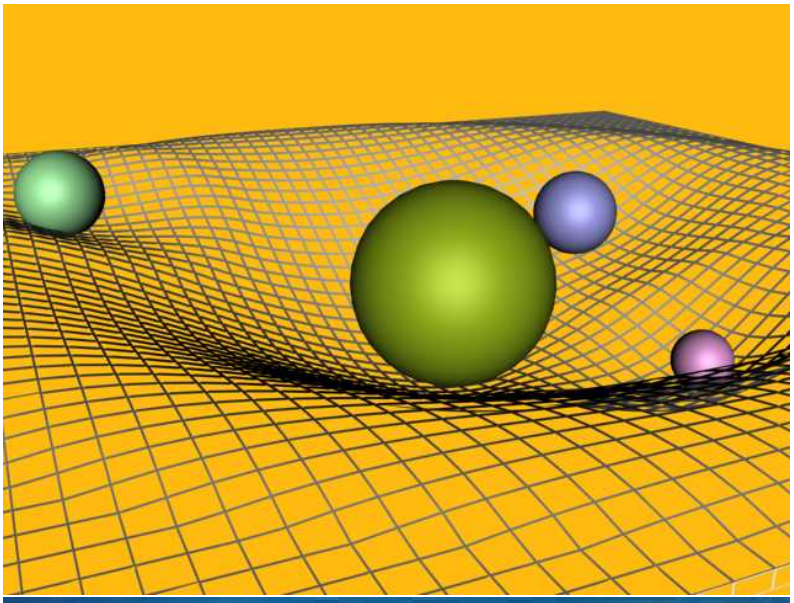




General Relativity and Cosmology

Steven Balbus

Version: January 16, 2022

These notes are intended for classroom use. There is little here that is truly original, but it is possible that 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 readers who have already done so: T. Acharya, V. Alimchandani, C. Arrowsmith, J. Bonifacio, C. Bowskill, T. Chan, Y. Chen, M. Elliot, L. Fraser-Taliente, G. Johnson, A. Mummery, A. Orr, J. Patton, V. Pratley, W. Potter, T. Surawy-Stepney, and H. Yeung.

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

Steven Balbus

Welcome to B5, General Relativity, Hilary Term 2022.

This set of notes is moderately revised from HT21, with a different treatment of the Schwarzschild orbits in Chapter 6. *Chapters 7 and 11 are off-syllabus and non-examinable.* They are included as a work in progress, toward the preparation of a textbook, for students who would like to be introduced to more advanced material. Your comments and questions are welcome.

The examinable material for HT22 is Chapters 1-6, 8.1–8.5, 9, 10.1–10.6. There are also a few individual sections that are marked as nonexaminable.

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 widely-used, clear, well-blended book, admirably covering the mathematics, physics, and astrophysics of GR. Good presentation of black holes (though there are some errors in the treatment of Kerr holes) and gravitational radiation. The explanation of the geodesic equation and the affine connection is particularly clear and enlightening. A little thin on physical cosmology, though a nice introduction to the physics of inflation and the early universe. Overall, a favourite student 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 is a classic reference, but in many ways it has become very dated. A paucity of material on (and a disparaging view of) 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. In this view, gravity is a field theory with a mere geometrical "analogy." 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 many 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 what is now classical physical cosmology are the text's main strength, and these are very fine indeed. Weinberg also has a more recent graduate-level text on cosmology *per se,* (*Cosmology* 2007, Oxford: Oxford University Press). While comprehensive in coverage, the level is advanced and not particularly student-friendly.

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 seemingly endless meanderings. Having said that, it is truly a monumental achievement. In the end, there is a lot wonderful material here, much that is difficult to find anywhere else, and it has all aged very well. Here, we find geometry is front and centre from start to finish, with lots of very detailed black hole and gravitational radiation physics, half-a-century on more timely than ever. I heartily recommend its deeply 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. A new hardbook edition was published in October 2017 by Princeton University Press. The core text is unchanged, but there is interesting new prefatory material, which reviews developments since the 1970s.

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, though written by a major contributor. 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 Temaine, S. 2008, *Galactic Dynamics,* (Princeton: Princeton University Press) Masterful text on galaxies with an excellent introduction to cosmology in the appendix. Very readable, given the high level of mathematics.

Carroll, S. M. 2019, *Spacetime and Geometry. An Introduction to General Relativity,* (Cambrige: Cambridge University Press) An up-to-date text with good coverage of many modern advanced topics such as differential geometry, singularity theorems, and field theory in curved spacetime. The treatment is very mathematical compared with HEL06, and there is also a tendency to state results without laying sufficient groundwork, which can be frustrating when encountering material for the first time. Still, the motivated student will find much of interest here, after mastering the material in this course.

Landau, L., and Lifschitz, E. M. 1962, *Classical Theory of Fields,* (Oxford: Pergamon) Classic advanced text; original and interesting treatment of gravitational radiation. Some material here can't be found elsewhere. Again, best after you've completed this course; otherwise dedicated students only!

Longair, M., 2006, *The Cosmic Century,* (Cambridge: Cambridge University Press) Excellent blend of observations, theory and the history of cosmology. This is a good general reference for anyone interested in astrophysics, a very lucid and informative read.

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,* (New York: Wiley) A text I grew up with. Very clear, with a nice summary of applications of GR to compact objects, lots of useful exercises, and very elucidating 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, θ, ϕ. Time is always the 0-component. Beware of extraneous factors of c in 0-index quantities, present in e.g. $T^{00} = \rho c^2$, $dx^0 = cdt$, 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 $\kappa, \lambda, \mu, \nu, \rho$ etc. are used to represent arbitrary spacetime components in all general relativity calculations.

- The Greek indices α, β, 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 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 dare even think of setting 2π to unity.

- It is "Lorentz invariance," but "Lorenz gauge." Not a typo, two different blokes.

Really Useful Numbers

$c = 2.99792458 \times 10^8$ m s^{-1} (Speed of light, exact.)

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

$h = 6.62607015 \times 10^{-34}$ J s (Planck's constant, exact.)

$a_\gamma = 7.56573325... \times 10^{-16}$ J m^{-3} K^{-4} (Blackbody radiation constant, exact.)

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

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

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

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

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

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

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

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

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

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

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

1 AU $= 1.495978707 \times 10^{11}$m (1 Astronomical Unit, exact.)

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

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

For diagonal g_{ab},

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

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

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

Ricci tensor:

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

Contents

1 An overview **13**

1.1 The legacy of Maxwell . . . 13
1.2 The legacy of Newton . . . 14
1.3 The need for a geometrical framework . . . 14

2 The toolbox of geometrical theory: Minkowski spacetime **18**

2.1 4-vectors . . . 18
2.2 Transformation of gradients . . . 21
2.3 Transformation matrix . . . 22
2.4 Tensors . . . 24
2.4.1 The energy-momentum stress tensor . . . 24
2.5 Electrodynamics . . . 25
2.5.1 4-currents and 4-potentials . . . 25
2.5.2 Electromagnetic field tensor . . . 27
2.6 Conservation of $T^{\alpha\beta}$ . . . 27

3 The effects of gravity **31**

3.1 The Principle of Equivalence . . . 31
3.2 The geodesic equation . . . 33
3.3 The metric tensor . . . 35
3.4 The relationship between the metric tensor and affine connection . . . 36
3.5 Variational calculation of the geodesic equation . . . 37
3.6 The Newtonian limit . . . 38

4 Tensor Analysis **43**

4.1 Transformation laws . . . 43
4.2 The covariant derivative . . . 45
4.2.1 Contravariant vectors . . . 45
4.2.2 Transformation of the affine connection . . . 46
4.2.3 Covariant vectors . . . 48
4.2.4 Tensors of arbitrary rank . . . 48
4.2.5 Going from special to general relativity . . . 49
4.3 The affine connection and basis vectors . . . 50
4.4 Volume element . . . 52
4.5 Covariant div, grad, curl, and all that . . . 52

4.6 Hydrostatic equilibrium . . . 54
4.7 Covariant differentiation and parallel transport . . . 56

5 The curvature tensor 58
5.1 Commutation rule for covariant derivatives . . . 58
5.2 Parallel transport . . . 59
5.3 Algebraic identities of $R^{\sigma}{}_{\nu\lambda\rho}$ . . . 62
5.3.1 Remembering the curvature tensor formula. . . . 62
5.3.2 $R_{\lambda\mu\nu\kappa}$: fully covariant form . . . 62
5.3.3 How many independent components does the curvature have? . . . 64
5.4 The Ricci tensor . . . 65
5.5 Curvature and Newtonian gravity . . . 66
5.6 The Bianchi Identities . . . 66
5.7 The Einstein field equations . . . 68
5.8 An alternative formulation of the field equations . . . 71

6 Schwarzschild black holes 74
6.1 Schwarzschild metric . . . 74
6.2 Schwarzschild metric in Cartesian coordinates . . . 79
6.3 The Schwarzschild Radius . . . 80
6.4 Why does the determinant of $g_{\mu\nu}$ not change with M? . . . 82
6.5 Working with Schwarzschild spacetime. . . . 84
6.5.1 Radial photon geodesic . . . 84
6.5.2 Orbital equations . . . 85
6.6 The deflection of light by an intervening body. . . . 87
6.7 The advance of the perihelion of Mercury . . . 90
6.7.1 Newtonian orbits . . . 91
6.7.2 The relativistic orbit of Mercury . . . 93
6.8 General solution for Schwarzschild orbits . . . 95
6.8.1 Formulation . . . 95
6.8.2 Solution . . . 97
6.9 Gravitational Collapse . . . 98
6.10 Shapiro delay: the fourth protocol . . . 99

7 Kerr Black Holes 102
7.1 Rotating spacetime: Kerr-Schild metric . . . 102
7.2 Boyer-Lindquist coordinates . . . 104

7.3 Kerr geometry . . . 106
7.3.1 Singularities . . . 106
7.3.2 Stationary limit surface: $g_{00} = 0$ . . . 107
7.3.3 Event horizon: $g^{rr} = 0$ . . . 107
7.3.4 The angular velocity of the outer event horizon . . . 110
7.4 The Penrose Process . . . 110
7.4.1 Energy and angular momentum transferral . . . 110
7.4.2 Relation to the Hawking Area Theorem . . . 111
7.5 Orbital equations . . . 112
7.6 Circular and quasi-circular orbits . . . 114

8 Self-Gravitating Relativistic Hydrostatic Equilibrium 119
8.1 Historical Introduction . . . 119
8.2 Fundamentals . . . 121
8.3 Constant density stars . . . 123
8.4 White Dwarfs . . . 125
8.5 Neutron Stars . . . 128
8.6 Pumping iron: the physics of neutron star matter . . . 133

9 Gravitational Radiation 137
9.1 The linearised gravitational wave equation . . . 139
9.1.1 Come to think of it... . . . 144
9.2 Plane waves . . . 145
9.2.1 The transverse-traceless (TT) gauge . . . 145
9.3 The quadrupole formula . . . 147
9.4 Radiated Energy . . . 149
9.4.1 A useful toy problem . . . 149
9.4.2 A conserved energy flux for linearised gravity . . . 150
9.5 The energy loss formula for gravitational waves . . . 153
9.6 Gravitational radiation from binary stars . . . 156
9.7 Detection of gravitational radiation . . . 159
9.7.1 Preliminary comments . . . 159
9.7.2 Indirect methods: orbital energy loss in binary pulsars . . . 161
9.7.3 Direct methods: LIGO . . . 164
9.7.4 Direct methods: Pulsar timing array . . . 167

10 Cosmology 170

10.1 Introduction . . . 170
10.1.1 Newtonian cosmology . . . 170
10.1.2 The dynamical equation of motion . . . 172
10.1.3 Cosmological redshift . . . 173
10.2 Cosmology models for the impatient . . . 174
10.2.1 The large-scale spacetime metric . . . 174
10.2.2 The Einstein-de Sitter universe: a classic toy model . . . 175
10.3 The Friedmann-Robertson-Walker Metric . . . 178
10.3.1 Maximally symmetric 3-spaces . . . 179
10.4 Large scale dynamics . . . 182
10.4.1 The effect of a cosmological constant . . . 182
10.4.2 Formal analysis . . . 183
10.4.3 The Ω_0 parameter . . . 186
10.5 The classic, matter-dominated universes . . . 187
10.6 Our Universe . . . 189
10.6.1 Prologue . . . 189
10.6.2 A Universe of ordinary matter and vacuum energy . . . 190
10.6.3 Radiation-dominated era . . . 192
10.7 Observational foundations of cosmology . . . 195
10.7.1 The first detection of cosmological redshifts . . . 195
10.7.2 The cosmic distance ladder . . . 196
10.7.3 The parameter q_0 . . . 197
10.7.4 The redshift–magnitude relation . . . 198
10.8 The Cosmic Microwave Background Radiation (CMB) . . . 201
10.8.1 Overview . . . 201
10.8.2 An observable cosmic radiation background: the Gamow argument . . . 203
10.8.3 The cosmic microwave background (CMB): subsequent developments 207
10.9 Thermal history of the Universe . . . 207
10.9.1 Prologue . . . 207
10.9.2 Classical cosmology: helium nucleosynthesis . . . 210
10.9.3 Neutrino and photon temperatures . . . 212
10.9.4 Ionisation of hydrogen . . . 213

11 Supplemental Topics **216**
11.1 The Oppenheimer-Snyder problem . . . 216
11.2 The energy and angular momentum of black hole spacetimes . . . 220

11.3 Particle production and event horizons . . . 222
11.3.1 Minkowski and Rindler metrics . . . 223
11.3.2 Scalar field equation . . . 224
11.3.3 Review of quantum field theory . . . 225
11.3.4 Bogoliubov formalism . . . 226
11.3.5 Invariant inner product . . . 227
11.3.6 Minkowski and Rindler vacuums . . . 229
11.3.7 The Hawking effect . . . 231
11.3.8 Hawking luminosity and Bekenstein-Hawking entropy . . . 232
11.4 The Raychaudhuri equation . . . 234
11.5 Effects of electric charge . . . 237
11.5.1 The Reissner-Nordström metric . . . 237
11.5.2 Singular metric structure: worm holes and white holes . . . 239
11.5.3 Naked singularities . . . 240
11.6 The growth of density perturbations in an expanding universe . . . 240
11.7 Inflationary Models . . . 243
11.7.1 "Clouds on the horizon" . . . 243
11.7.2 The stress energy tensor of a scalar field . . . 244
11.7.3 Effective Potentials in Quantum Field Theory . . . 248
11.7.4 The slow roll inflationary scenario . . . 250
11.7.5 Inflated expectations . . . 252
11.7.6 A final word . . . 254

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 unforseen puzzles emerge that could not 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.

The 1861 synthesis provided by 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 propagated, and showed that the propagation velocity is $1/\sqrt{\epsilon_0\mu_0}$ (ϵ_0 is the permitivity, and μ_0 the permeability, of free space). This combination is precisely equal to the speed of light. Light is a form of electromagnetic radiation! The existence of electromagnetic raditation was later verified by brilliant experiments carried out by Heinrich Hertz in 1887, in which the radiation was both generated and detected.

But when one probed, Maxwell's theory, for all its success, had disquieting features. 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 aways $1/\sqrt{\epsilon_0\mu_0}$. A related point is that simple Galilean invariance is not obeyed, i.e. absolute velocities appear to affect the physics, something that had not been seen before. Lorentz and Larmor in the late nineteenth century discovered that Maxwell's equations in fact had a simple mathematical velocity transformation that did leave 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 very compelling experiments, most notably the famous Michelson-Morley investigation); and (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. In the process, electromagnetic radiation in vacuo took on a reality all its own: it no longer need be viewed as some sort of aetherial displacement. The increasingly problematic

aether host medium could be abandoned: the waves were "hosted" by the vacuum itself. Most stunningly, *all* equations of physics, not just electromagnetic phenomena, had to be invariant in form under the kinematic Lorentz transformations, even with the peculiar relative time variable. The transformations themselves really had nothing to do with electrodynamics. They were far more general. It just so happened that the Maxwell equations was the first such set to adhere to correct transformation laws. The others had to be fixed up!

These ideas, and the consequences that ensued collectively from them (e.g., the equivalence of energy and inertia) became known as *relativity theory*, in reference to the invariance of form with respect to relative velocities. The relativity theory that arose from the disinternment of the Lorrentz transformations from within the Maxwell 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? By the early twentieth century it had been verified with great precision, even predicting the existence of a new distant planet (Neptune) in the solar system. But Newtonian gravity is manifestly *not* compatible with relativity, because the governing Poisson equation

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

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

In Maxwell's theory, the *electrostatic* potential also 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 do not propagate instantaneously throughout all space. They propagate at the speed of light c. At least in retrospect, this is rather simple. Might it not be the same for gravity?

Alas, 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 in motion, but the mass does: mass, as a source of gravity, depends upon velocity. This deceptively simple detail complicates everything. It leads immediately to the realisation that in a relativistic theory, energy, like matter, must also be a source of a gravitational field, and this includes the distributed energy of the gravitational field itself. In other words, any 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 presented a remarkable insight. He argued that one should view the Lorentz transformations not merely as a set of rules for how space

and time coordinates change from one constant-velocity kinematical reference frame to another. Minkowski's idea was that these coordinates should be regarded as living in their own sort of pseudo-Euclidian geometry—a *spacetime,* if you will: Minkowski spacetime. There was even a ready-made mathematical structure at hand to be used: hyperbolic geometry.

To understand the motivation for this, start simply. We know that in ordinary Euclidian space we are free to choose any set of Cartesian coordinates we like, and it can make no difference to the description of the space itself, e.g., 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

$$dx^2 + dy^2 = dx'^2 + dy'^2 \tag{3}$$

i.e., the distance between two closely spaced points is the same number, regardless of the coordinates used. The quantity $dx^2 + dy^2$ is said to be an "invariant" under rotational transformations.

Now, an abstraction. From a mathematical viewpoint, there is nothing special 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 formal point-of-view, we needn't worry about the plus/minus sign. An invariant is an invariant. Indeed, the hyperbolic geometry arising from this form of invariant had been studied by Bolyai and Lobachevsky (and was known to Gauss) long before. With $dy^2 - dx^2$ as our invariant, we are in fact describing a hyperbolic Minkowski space, with $dy = cdt$ and dx a space interval as before. Under the simplest form of the Lorentz transformations, $c^2dt^2 - dx^2$ emerges as an invariant quantity, which is precisely what is needed to guarantee that the speed of light is always constant...an invariant. With this formulation, the interval $c^2dt^2 - dx^2$ always vanishes for light propagation along the x direction, in any coordinates (i.e., "observers") related by a Lorentz transformation. The obvious generalisation

$$c^2dt^2 - dx^2 - dy^2 - dz^2 = 0 \tag{4}$$

will guarantee the same is true in any spatial direction. Our Universe is one of four dimensions, but its hyperbolic geometry picks out one of these dimensions to appear with a different sign in formulating the invariant interval. This dimensional separation therefore has a different character to the other three, and is experienced by our consciousness in a very different way. We feel the need to separate this one particular form of dimensional separation from the other three, and give it its own name: time. And it is not just our subjective consciousness that treats this dimension spearately. Informally speaking, when we solve a partial differential equation (PDE) like the Laplace equation, with all coordinates on an equal footing, we specify the value of our function (or its gradient) on closed surfaces everywhere, and allow the equation to fill in the rest of the space. But when we solve a PDE like the wave equation, we fix the special time coordinate, and specify the function (and its time derivative) on all the non-time surfaces, and solve an initial value problem going forward. The Laplace equation is an example of what is known as an *elliptical PDE;* the wave equation is a *hyperbolic PDE.* Our conscious reality, flowing from one moment to the next in an external dimension of separation, "feels" like the solution to a hyperbolic PDE.

Minkowski took what seemed to be a purely kinematical requirement—that the speed of light needs to be a universal constant—and gave it a geometrical interpretation in terms of an invariant quantity (a "quadratic form" as it is sometimes called) in Minkowski space. Rather, Minkowski *spacetime.*

Pause. And so what? Call it whatever you like. Who needs obfuscating mathematical pretence? Eschew obfuscation! One could argue that lots of things add together quadratically, and it is not particularly helpful to regard them as geometrical objects. The Lorentz transform stands on its own! I like my way! That was very much Einstein's initial take on Minkowski's pesky mathematical meddling with his theory.

However, as the reader has no doubt sensed, it is the geometrical viewpoint that is truly the more fundamental, a perspective that time (our fourth dimension!) has vindicated. In Minkowski's 1908 tour-de-force paper, we find more than just a simple analogy or mathematical pretence. We understand *why* the kinematics is changing from Galilean-Newtonian: it stems from the underlying geometry of the spacetime itself. In Minkowski's work, we find the first presentation of relativistic 4-vectors and tensors, of the Maxwell equations in manifestly covariant form, and the remarkable realisation that the magnetic vector and electrostatic potentials combine seamlessly to form their own 4-vector. Gone are the comforting kinematical tools, the clocks, the rods and the quaint Swiss trams of Einstein's 1905 relativity paper. This is no mere "überflüssige Gelehrsamkeit" (superfluous erudition), Einstein's dismissive term for the whole business.

It took until 1912 for Einstein to finally change his viewpoint. Five years earlier, he had had the seed of a very important idea, but had been unable to build upon it. It wasn't until 1912 that Einstein discovered what it was that he had discovered. His great revelation was that *the effect of the presence of matter (or its equivalent energy) throughout the Universe is to distort Minkowski's spacetime, and that this embedded geometrical distortion manifests itself as the gravitational field.* Minkowski spacetime is not some mere mathematical formalism, it is physical stuff, you can mess with it, and it takes real, tangible energy to do that. Of course, the spacetime distortions of which we speak must become, when sufficiently weak, familiar Newtonian gravitational theory. You thought gravity was a dynamical force? No. Gravity is a purely geometrical phenomenon.

Now that is one big idea. It is an idea that physicists are still taking in. How did Einstein make this leap? What was the path that led him to change his mind? Where did the notion of a gravity-geometry connection come from?

It came from that "important idea" he had in 1907, a year before Minkowski published his geometrical interpretation of relativity. In a freely falling elevator, or more safely in an expertly piloted aircraft executing a ballistic parabolic arc, one feels weightless. That is, the effect of gravity can be made to disappear locally in the appropriate reference frame—i.e., the right coordinates. Let that sit. The physics of gravity must be intimately associated with coordinate transformations! This is because gravity has exactly the same effect on all types of mass, regardless of bulk or elemental composition. But how does one go from "intimately associated with coordinate transformations" to a rigorous theory? In 1912 Einstein, with Grossman's help, grasped that this is precisely what would naturally emerge from a theory in which objects were responding to background geometrical distortions instead of to an applied force. In a state of free-fall, gravity is in effect absent, and we locally return to the environment of an 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. It is easy to be fooled into thinking that the Earth is flat, if your view is too local. "Tangent plane coordinates" on small-scale road maps locally eliminate spherical geometry complications, but if we are flying from Oxford 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, but that it is always possible to find coordinates in which local distortions may be eliminated *to leading local 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. Space*time.* Once again, bringing in time, you see, is everything. Who would have thought of doing that 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 mathematical language of gravitational theory, a point impressed upon Einstein by his friend and colleague the mathematician Marcel Grossman. Indeed, half a century before Einstein, Riemann himself had the notion that gravity might arise from a non-Euclidian curvature

in a space of three or more dimensions! He got nowhere, because time was not part of his geometry. Riemann was thinking only of space. It was the genius of Minkowski, who by showing 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 dimension we associate with *time*. And that is why Riemann failed.

In brief: Riemann created the mathematics of non-Euclidian geometry. Minkoswki 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 the force we call gravity arises from the distortions of Minkowski's flat spacetime which are created by the presence 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 draw 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 spacetime

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 geometrical structure in which it is always possible to find coordinates in which the space looks locally smooth is known as a *Riemannian manifold.* Crudely speaking, manifolds are smooth surfaces in n dimensions. A circle is a manifold, a spherical surface is a manifold, an egg shell is a manifold, but a tetrahedron is not. Mathematicians would say that an n-dimensional manifold is *homeomorphic* to n-dimensional Euclidian space. Actually, since our local invariant interval $c^2dt^2 - dx^2$ is not a simple sum of squares, but contains that notorious 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 with perfect Greek lettering) but are actually imposters. Their credentials are fake.

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 4-vectors

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

$$ct' = \gamma(ct - vx/c) = \gamma(ct - \beta x) \tag{5}$$

$$x' = \gamma(x - vt) = \gamma(x - \beta ct) \tag{6}$$

$$y' = y \tag{7}$$

$$z' = z \tag{8}$$

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

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \equiv \frac{1}{\sqrt{1 - \beta^2}} \tag{9}$$

is the *Lorentz factor*. The primed frame can be thought of as the frame moving with an object we are studying, 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

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

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

where $c\boldsymbol{\beta} = \boldsymbol{v}$ is the vector velocity and boldface $\boldsymbol{x}$'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 individual 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:

$$x' = x\cos\theta + y\sin\theta \tag{12}$$

$$y' = -x\sin\theta + y\cos\theta \tag{13}$$

$$z' = z \tag{14}$$

Show that if we define

$$\beta \equiv \tanh\zeta, \tag{15}$$

then

$$\gamma = \cosh\zeta, \quad \gamma\beta = \sinh\zeta, \tag{16}$$

and

$$ct' = ct\cosh\zeta - x\sinh\zeta, \tag{17}$$

$$x' = -ct\sinh\zeta + x\cosh\zeta. \tag{18}$$

What happens if we apply this transformation twice, once with "angle" ζ from (x, t) to (x', t'), then with angle ξ 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

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

Then show that (12) and (13) become

$$x' = \Gamma(x + \alpha y) \tag{20}$$

$$y' = \Gamma(y - \alpha x) \tag{21}$$

Thus, while a 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 (ordinary trigonometric) symmetry amongst themselves. Timelike extensions exhibit Lorentzian (hyperbolic trigonometric) symmetry with a spacelike extension, and probably nothing with another possible timelike extension. Maybe the need for internal symmetry between all of the dimensional extensions 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 $\boldsymbol{V}$, 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^α, with α taking on the values 0 through 3. Symbolically,

$$V^\alpha = (V^0, \boldsymbol{V}) \tag{22}$$

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

$$p^\alpha = (E/c, \boldsymbol{p}) \tag{23}$$

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

$$\boldsymbol{p} = \gamma m \boldsymbol{v}, \quad E = \gamma m c^2 \tag{24}$$

where m is a particle's rest mass. Just as

$$(ct)^2 - x^2 \tag{25}$$

is an invariant quantity under Lorentz transformations, so too is

$$E^2 - (pc)^2 = m^2 c^4 \tag{26}$$

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

$$U^\alpha = \gamma(c, \boldsymbol{v}) \tag{27}$$

Note that in the rest frame of a particle, $U^0 = c$ (a constant) and the ordinary 3-velocity components $\boldsymbol{U} = 0$. To get to any other frame, just use ("boost with") the Lorentz transformation along the $\boldsymbol{v}$ 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 ρ_q in a fluid at rest. We may create a 4-vector which, in this 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^α, only with a different normalisation constant. Now boost! The resulting 4-vector is denoted

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

The time component gives the charge density in any frame, and the 3-vector components are the corresponding standard current density $\boldsymbol{J}$. 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 Φ, and $\boldsymbol{J}$ is the source of the magnetic field vector potential $\boldsymbol{A}$. Moreover, as we will later see,

$$A^\alpha = (\Phi, \boldsymbol{A}) \tag{29}$$

is itself a 4-vector[1]. 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'^α 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

$$ct = \gamma(ct' + \beta x') \tag{30}$$

$$x = \gamma(x' + \beta ct') \tag{31}$$

we find

$$\frac{\partial}{\partial t'} = \frac{\partial t}{\partial t'}\frac{\partial}{\partial t} + \frac{\partial x}{\partial t'}\frac{\partial}{\partial x} = \gamma\frac{\partial}{\partial t} + \gamma\beta c\frac{\partial}{\partial x} \tag{32}$$

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

In other words,

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

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

and for completeness,

$$\frac{\partial}{\partial y'} = \frac{\partial}{\partial y} \tag{36}$$

$$\frac{\partial}{\partial z'} = \frac{\partial}{\partial z}. \tag{37}$$

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

$$cdt' = \gamma(cdt - \beta dx), \tag{38}$$

$$dx' = \gamma(dx - \beta cdt), \tag{39}$$

$$dy' = dy, \tag{40}$$

$$dz' = dz. \tag{41}$$

[1]We are using relativistic esu units, in which $\boldsymbol{A}$ and Φ have the same dimensions, which is natural.

This has a very important consequence:

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

or simplifying,

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

Adding y and z into the mixture changes nothing. Thus, a scalar product exists between dx^α and $\partial/\partial x^\alpha$ that yields a *Lorentz scalar,* much as $\boldsymbol{dx}\cdot\boldsymbol{\nabla}$, 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 use of the summation sign Σ_α. For example:

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

I will often further shorten this to $dx^\alpha\partial_\alpha$. This brings us to another important notational convention. I was careful to write ∂_α, not ∂^α, which differs from the former by the sign of its time (0) component. Superscripts will be reserved for vectors, like dx^α, which transform according to (5) through (8) from one frame to another, with the x' axis moving at a velocity v relative to the x axis. Subscripts will be used to indicate vectors that transform as the gradient components in equations (34)–(37). Superscipt vectors like dx^α are referred to as *contravariant* vectors; their subscripted counterparts are *covariant.* (These names will acquire significance later.) The co- contra- difference is a very important distinction in general relativity, and we begin by respecting it in special relativity.

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

$$[-cdt'] = \gamma([-cdt] + \beta dx) \tag{45}$$

$$dx' = \gamma(dx + \beta[-cdt]) \tag{46}$$

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

$$dx^\alpha = (cdt, dx, dy, dz) \tag{47}$$

$$dx_\alpha = (-cdt, dx, dy, dz) \tag{48}$$

It is simple to go between covariant and contravariant forms in special relativity: flip the sign of the time component. We are motivated to formalise this by introducing a matrix $\eta_{\alpha\beta}$, which is defined as

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

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

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

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 its 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

$$\Lambda^\alpha{}_\beta \, dx^\beta = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} dx^0 \\ dx^1 \\ dx^2 \\ dx^3 \end{pmatrix} \tag{51}$$

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

$$\Lambda^\alpha{}_\beta \Lambda^\epsilon{}_\gamma \eta_{\alpha\epsilon} = \eta_{\beta\gamma} \tag{52}$$

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

$$V'^\alpha W'_\alpha = V'^\alpha W'^\beta \eta_{\beta\alpha} = \Lambda^\alpha{}_\gamma V^\gamma \Lambda^\beta_\epsilon W^\epsilon \eta_{\beta\alpha} = V^\gamma W^\epsilon \eta_{\gamma\epsilon} = V^\gamma W_\gamma \tag{53}$$

For covariant vectors, for example ∂_α, 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 β reversed:

$$\tilde{\Lambda}^\alpha{}_\beta = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{54}$$

Note that

$$\tilde{\Lambda}^\alpha{}_\beta \Lambda^\beta{}_\gamma = \delta^\alpha_\gamma, \tag{55}$$

where δ^α_γ 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_α 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 $\boldsymbol{v}$. The equations of constant time in this frame are given by the requirement that dx^α and U_α are orthogonal.

Exercise. Show that the general Lorentz transformation matrix is:

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

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

2.4 Tensors

2.4.1 The energy-momentum stress tensor

There is more to relativistic life than vectors and scalars. There are objects called tensors, which carry several indices. 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

$$T'^{\alpha\beta} = \Lambda^{\alpha}{}_{\gamma}\Lambda^{\beta}{}_{\epsilon}T^{\gamma\epsilon}, \tag{57}$$

while

$$T'_{\alpha\beta} = \tilde{\Lambda}^{\gamma}{}_{\alpha}\tilde{\Lambda}^{\epsilon}{}_{\beta}T_{\gamma\epsilon}, \tag{58}$$

and of course

$$T'^{\alpha}{}_{\beta} = \Lambda^{\alpha}{}_{\gamma}\tilde{\Lambda}^{\epsilon}{}_{\beta}T^{\gamma}{}_{\epsilon}, \tag{59}$$

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

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

$$\tilde{\Lambda}^{\alpha}{}_{\beta}\tilde{\Lambda}^{\epsilon}{}_{\gamma}\eta_{\alpha\epsilon} = \eta_{\beta\gamma} \tag{60}$$

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

$$\Lambda^{\epsilon}{}_{\gamma}\tilde{\Lambda}^{\alpha}{}_{\beta}\delta^{\gamma}_{\alpha} = \Lambda^{\epsilon}{}_{\gamma}\tilde{\Lambda}^{\gamma}{}_{\beta} = \delta^{\epsilon}_{\beta}. \tag{61}$$

Here is another tensor, slightly less trivial:

$$W^{\alpha\beta} = U^{\alpha}U^{\beta} \tag{62}$$

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

$$T^{\alpha\beta} = \rho_r\langle u^{\alpha}u^{\beta}\rangle \tag{63}$$

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 these 4-velocity products preserves its tensor properties: the linear average of all the little particle tensors must itself be a tensor.) ρ_r is a local rest density, a scalar number. (Here, r is not an index.)

The component T^{00} is just ρc^2, the energy density of the swarm, where ρ (without the r) includes both a rest mass energy and a thermal contribution. (The thermal part comes from averaging the "swarming" γ 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 γ factors when the u^i are relativistic?) Hence, in the frame in which the *swarm* has no net translational *bulk* motion,

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

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^α we could adopt a brute force method, boost away, and apply the Λ 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:

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

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

2.5 Electrodynamics

2.5.1 4-currents and 4-potentials

The vacuum Maxwell equations for the electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$ produced by a charge density ρ_q and vector current $\boldsymbol{J}$ may be written:

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = 4\pi\rho_q \tag{66}$$

$$\boldsymbol{\nabla}\times\boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} \tag{67}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0 \tag{68}$$

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \frac{4\pi}{c}\boldsymbol{J} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} \tag{69}$$

We have used gaussian esu units rather than SI. The former is the natural choice for relativistic field applications (see Jackson 1999 for an interesting discussion on the topic of units in electrodynamics).

To cast these equations into a Lorentz invariant form, we first introduce the standard vector potential $\boldsymbol{A}$, so that if

$$\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A} \tag{70}$$

equation (68) is automatically satisfied in full generality. Then, equation (67) becomes

$$\boldsymbol{\nabla}\times\left(\boldsymbol{E} + \frac{1}{c}\frac{\partial\boldsymbol{A}}{\partial t}\right) = 0 \tag{71}$$

This means that the bracketed combination is derivable from a gradient, leading to

$$\boldsymbol{E} = -\boldsymbol{\nabla}\Phi - \frac{1}{c}\frac{\partial\boldsymbol{A}}{\partial t} \tag{72}$$

where Φ is the standard electrostatic potential with its conventional minus sign.

The Lorentz invariant properties of Maxwellian electrodynamics become more apparent when we write the sourced equations (66) and (69) in terms of Φ and $\boldsymbol{A}$:

$$-\frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2} + \nabla^2\Phi + \frac{1}{c}\frac{\partial}{\partial t}\left(\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial\Phi}{\partial t}\right) = -4\pi\rho_q \tag{73}$$

$$-\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} + \nabla^2\boldsymbol{A} - \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial\Phi}{\partial t}\right) = -\frac{4\pi}{c}\boldsymbol{J} \tag{74}$$

We have alread noted in §(2.1), equation (28), that the combination

$$J^\alpha = \gamma(c\rho_0, \boldsymbol{v}\rho_0) = (c\rho_q, \boldsymbol{v}\rho_q) \tag{75}$$

is a relativistic 4-vector, provided that ρ_q is identified with the "boosted" rest frame charge density ρ_0. This leads immediately to a relativistic formulation of (73) and (74) potentials. We combine these two equations into a single, relativistically invariant, 4-vector equation:

$$\partial_\alpha\partial^\alpha A^\beta - \partial^\beta(\partial_\alpha A^\alpha) = -\frac{4\pi}{c}J^\beta \tag{76}$$

where

$$A^\alpha = (\Phi, \boldsymbol{A}) \tag{77}$$

must itself be a 4-vector, because that is the only kind of object that will transform properly from one frame to another, given the 4-vector J^β on the right side of the equation. It is straightforward to show that the physical $\boldsymbol{E}$ and $\boldsymbol{B}$ fields remain unchanged if we add a 4-gradient of a scalar function Λ to A^α:

$$A^\alpha \to A^\alpha + \partial^\alpha\Lambda, \tag{78}$$

an important property of the Maxwell equations known as gauge invariance. Among other consequences, it means that we are free to choose, via an appropriate choice of Λ, whatever value of $\partial_\alpha A^\alpha$ is convenient for the task at hand. In particular, if we work in the so-called *Lorenz gauage* (note spelling!)[2]:

$$\partial_\alpha A^\alpha = 0, \tag{79}$$

[2]Ludvig Lorenz (1829-1891) and Hendrick Lorentz (1853-1928), a sort of Thomson and Thompson duo of Victorian-era physics, both made seminal contributions to electrodynamic theory. It is perhaps inevitable that there is a *Lorentz-Lorenz equation,* discovered by both quite independently of one another, that relates the polarisability of a medium to its density and refractive index. It is Ludvig's gauge, however.

then equation (76) takes on the elegant form

$$\partial_\alpha \partial^\alpha A^\beta = -\frac{4\pi}{c} J^\beta, \tag{80}$$

a set of uncoupled classical wave equations with sources, and the starting point for most treatments of classical electrodynamical radiation theory.

Finally, we note that the 4-divergence of equation (76) gives zero identically on the left side of that equation, leading to an equation of charge conservation,

$$\partial_\alpha J^\alpha = 0, \tag{81}$$

a constraint that is very clear in the Lorenz gauge wave equation (80).

2.5.2 Electromagnetic field tensor

The natural extension of the equation $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$ leads us to define the electromagnetic field tensor:

$$F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha, \tag{82}$$

which is manifestly antisymmetric in its indices. We then have for the electric and magnetic field components, E_i, B_i:

$$F_{i0} = \partial_i A_0 - \partial_0 A_i = -\frac{\partial \Phi}{\partial x^i} - \frac{1}{c}\frac{\partial A_i}{\partial t} = E_i \tag{83}$$

and

$$F_{ij} = \partial_i A_j - \partial_j A_i \equiv \epsilon_{ijk} B^k \tag{84}$$

where $\epsilon^{ijk} = \epsilon_{ijk}$ is the Levi-Civita symbol, defined to be $+1$ when ijk is an even permutation of 123, equal to -1 for an odd permutation, and 0 when any index repeats. The indices of $F_{\alpha\beta}$ are raised and lowered in the usual way, by contraction with $\eta_{\alpha\beta}$.

Equation (76) becomes:

$$\partial_\alpha F^{\alpha\beta} = -\frac{4\pi}{c} J^\alpha. \tag{85}$$

It is now a straightforward exercise to show that the remaining two unsourced Mawell equations (67, 68) may be combined into a single equation expressing the vanishing of the cyclic derivative combination

$$\partial_\alpha F^{\beta\gamma} + \partial_\gamma F^{\alpha\beta} + \partial_\beta F^{\gamma\alpha} = 0. \tag{86}$$

Exercise. Verify equation (86) and show that it may also be written in terms of the four dimensional Levi-Civita symbol $\epsilon^{\alpha\beta\gamma\delta}$ as $\epsilon^{\alpha\beta\gamma\delta}\partial_\beta F^{\gamma\delta} = 0$.

2.6 Conservation of $T^{\alpha\beta}$

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

$$\frac{\partial T^{\alpha\beta}}{\partial x^\alpha} = 0 \tag{87}$$

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

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

$$\frac{\partial}{\partial t}\left[\gamma^2\left(\rho+\frac{Pv^2}{c^4}\right)\right]+\boldsymbol{\nabla}\cdot\left[\gamma^2\left(\rho+\frac{P}{c^2}\right)\boldsymbol{v}\right]=0 \tag{88}$$

which is the relativistic version of mass conservation,

$$\frac{\partial\rho}{\partial t}+\boldsymbol{\nabla}\cdot(\rho\boldsymbol{v})=0. \tag{89}$$

Elevated in special relativity, this becomes a statement of *energy conservation.* So one of the things we are conserving is energy. This is good.

The spatial part of the conservation equation reads

$$\frac{\partial}{\partial t}\left[\gamma^2\left(\rho+\frac{P}{c^2}\right)v_i\right]+\left(\frac{\partial}{\partial x^j}\right)\left[\gamma^2\left(\rho+\frac{P}{c^2}\right)v_iv_j\right]+\frac{\partial P}{\partial x^i}=0 \tag{90}$$

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

Now do the following *Exercise:*

Exercise. By using the explicit form (65) of $T^{\alpha\beta}$ and contracting equation (87) with U_β, prove the interesting identity:

$$0=\frac{\partial(\rho c^2U^\alpha)}{\partial x^\alpha}+P\frac{\partial U^\alpha}{\partial x^\alpha}.$$

(Remember that $U^\alpha U_\alpha=-c^2$, a constant.) What is the physical interpretation of this equation? If we use this result back in equation (87), show that takes a more familiar form:

$$\left(\eta^{\alpha\beta}+\frac{U^\alpha U^\beta}{c^2}\right)\frac{\partial P}{\partial x^\alpha}+\left(\rho+\frac{P}{c^2}\right)U^\alpha\frac{\partial U^\beta}{\partial x^\alpha}=0$$

This is recognisable as the standard equation of motion with relativistic corrections of order v^2/c^2.

What if there are other external forces? These forces are included in this formalism by expressing them in terms of the divergence of their own associated stress tensors. Then, it is the summed total of the $T^{\alpha\beta}$ that is conserved.

An important example which follows this description is the Maxwellian electromagnetic stress tensor. To construct this, begin with the following contracted form of the relativistic Maxwell equation (85):

$$\frac{1}{4\pi}F^\alpha{}_\gamma\partial_\beta F^{\beta\gamma}=-\frac{1}{c}J^\gamma F^\alpha{}_\gamma. \tag{91}$$

In particular, the 0 component of the right side of this equation reads:

$$-\frac{1}{c}J_iF^0{}_i=-\frac{1}{c}\boldsymbol{J}\cdot\boldsymbol{E} \tag{92}$$

This is immediately recognisable as minus the rate at which the electromagnetic fields do work on their sources, with a lead $1/c$ factor. (This factor is present for dimensional reasons,

and is in effect compensated by a $1/c$ factor in the $x^0 = ct$ time-coordinate partial derivative.) It is natural, then, to expect the left side of (91) to be expressible as the divergence of the $T^{\alpha 0}$ (energy) component of a full, symmetric electromagnetic stress tensor for the fields, $T^{\alpha\beta}_{\rm em}$. In other words, we seek an explicit form for the mathematical object $T^{\alpha\beta}_{\rm em}$ such that the equation

$$\frac{1}{4\pi} F^{\alpha}_{\ \gamma} \partial_\beta F^{\beta\gamma} = \partial_\beta T^{\alpha\beta}_{\rm em} \tag{93}$$

is satisfied.

Start with an integration by parts:

$$F^{\alpha}_{\ \gamma} \partial_\beta F^{\beta\gamma} = \partial_\beta (F^{\alpha}_{\ \gamma} F^{\beta\gamma}) - F^{\beta\gamma} \partial_\beta F^{\alpha}_{\ \gamma}. \tag{94}$$

The first term on the right is now an explicit divergence, and this is part of what we want. The second term may be written

$$F_{\beta\gamma} \partial^\beta F^{\alpha\gamma} = -F_{\beta\gamma} \partial^\beta F^{\gamma\alpha} = -F_{\beta\gamma} \partial^\gamma F^{\alpha\beta}, \tag{95}$$

where the second equality follows from switching the dummy indices γ and β and using antisymmetry twice, once on each F-tensor. Watch carefully now:

$$\begin{aligned} F_{\beta\gamma} \partial^\beta F^{\alpha\gamma} &= -\frac{1}{2} F_{\beta\gamma} (\partial^\beta F^{\gamma\alpha} + \partial^\gamma F^{\alpha\beta}) \\ &= -\frac{1}{2} F_{\beta\gamma} (\partial^\beta F^{\gamma\alpha} + \partial^\gamma F^{\alpha\beta} + \partial^\alpha F^{\beta\gamma}) + \frac{1}{2} F_{\beta\gamma} \partial^\alpha F^{\beta\gamma}. \end{aligned} \tag{96}$$

The first group of three terms has been constructed so that it explicitly vanishes by equation (86), the cyclic identity for the sourceless Maxwell equations. We conclude that

$$F_{\beta\gamma} \partial^\beta F^{\alpha\gamma} = \frac{1}{2} F_{\beta\gamma} \partial^\alpha F^{\beta\gamma} = \frac{1}{4} \eta^{\alpha\beta} \partial_\beta (F_{\delta\gamma} F^{\delta\gamma}). \tag{97}$$

Putting together equations (91), (93), (94) and (97), we arrive at the equation

$$\partial_\beta T^{\alpha\beta}_{\rm em} = -\frac{1}{c} J^\gamma F^{\alpha}_{\ \gamma}, \tag{98}$$

where

$$T^{\alpha\beta}_{\rm em} = \frac{1}{4\pi} \left(F^{\alpha}_{\ \gamma} F^{\beta\gamma} - \frac{1}{4} \eta^{\alpha\beta} F_{\delta\gamma} F^{\delta\gamma} \right) \tag{99}$$

This is our desired electromagnetic stress energy tensor. The right side of equation (98) is in fact expressible as minus the divergence of the stress energy tensor *for the particle sources*, for whom the electromagnetic field is an external disturbance (see W72). Thus, the total stress tensor for the sum of the particles and fields is conserved. We shall make explicit use of equation (99) in our discussion of electrically charged black holes in Chapter 10.

Exercise. Verify that the trace $T_{\rm em} \equiv \eta_{\alpha\beta} T^{\alpha\beta}_{\rm em} = 0$. This means that if we tried to make a scalar relativistic field theory for gravity of the form $\partial^\alpha \partial_\alpha \Phi = 4\pi G T^{\beta}_{\ \beta}$, it would fail immediately. Why?

Exercise. Show that:

$$T^{00}_{\rm em} = \frac{1}{8\pi} (E^2 + B^2),$$

$$T^{0i}_{\rm em} = \frac{1}{4\pi}(\boldsymbol{E}\times\boldsymbol{B})\cdot\boldsymbol{e_i},$$

where $\boldsymbol{e_i}$ is a unit vector in the i direction, and

$$T^{ij}_{\rm em} = -\frac{1}{4\pi}\left[E_iE_j + B_iB_j - \frac{1}{2}\delta_{ij}(E^2+B^2)\right].$$

Finally, show that $T_{\rm em} = 0$ implies $3P_{\rm rad} = \rho_{\rm rad}c^2$ where $P_{\rm rad}$ and $\rho_{\rm rad}$ represent radiation pressure and energy density respectively.

The reader perhaps may be gaining a sense of the difficulty in constructing a theory of gravity compatible with relativity. The mass-energy density ρ is but a single component of a stress tensor, and it is the entire stress tensor from all possible contributors, fields and particles alike, that would have to be the source of what we identify as the gravitational potential field. Recall that electromagnetic theory starts with electrostatics and Coulomb's law, but in the ultimate Maxwellian theory it is the entire 4-current J^α that is the source of electromagnetic fields. We should expect something similar here. Relativistic theories must work with scalars, vectors and tensors to preserve their invariance properties from one frame to another. This simple insight is already an achievement: we can, for example, expect pressure to play a role in generating gravitational fields. Would you have guessed that? Perhaps our relativistic gravity equation ought to look something like:

$$\nabla^2 G^{\mu\nu} - \frac{1}{c^2}\frac{\partial^2 G^{\mu\nu}}{\partial t^2} = T^{\mu\nu}, \tag{100}$$

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^α) 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 proportional to (J^α) on the right. So now we just need to find a $G^{\mu\nu}$ tensor to go with $T^{\mu\nu}$. Right?

Actually, this is not a bad 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 [3] 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 gravity-related "massy" property that might vary from one type of body to another with the same inertial bulk mass m_I. In this case, the acceleration a is $m_g g/m_I$, i.e., something that varies with the ratio of inertial to gravitational mass from one body to another. How well can we actually measure this ratio, or posing the question more accurately, how well do we know that this ratio 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

[3] With apologies to any readers who may actually have fallen off the roof of a house—safe space statement.

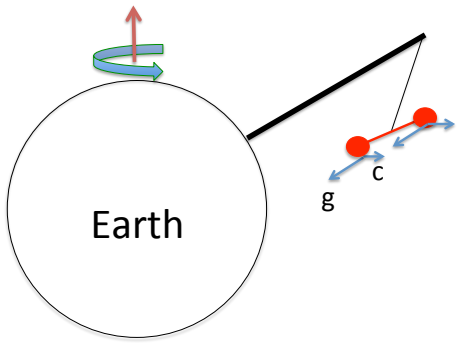


Figure 1: Schematic diagram of the Eötvös experiment. (Not to scale!) A barbell-shaped mass, the red object above, is hung from a pendulum on the surface of a rotating Earth. Its two end-masses are of different of material, say copper and lead. Each mass is affected by gravity (g), pulling it to the centre of the Earth with a force proportional to a *gravitational mass* m_g, and by 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. From the combined g and c forces, a difference between the inertial to gravitational mass ratio in the different composition end-masses will produce an unbalanced torque about the axis of the suspending fibre of the barbell.

Pisa. As with all classic precision gravity experiments (including those of Galileo!) we use a pendulum. The first direct measurement of the gravitational to inertial mass ratio actually predates relativity, the so-called Eötvös experiment, after Baron Loràn̈d 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 rotating masses would be sometimes hindered by any putative torque, and sometimes pushed forward. 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 from the Earth 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 responds identically to 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, whether 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

$$\frac{d^2x}{dt^2} = 0, \tag{101}$$

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

$$\frac{d^2x'}{dt^2} = -g, \tag{102}$$

which the reader will agree 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, then *locally the physics is exactly the same as if we were accelerating upwards at* g. But the effect of kinematic acceleration of an observer on the measured frequency of light is 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 $v = gh/c$ relative to the bottom of the elevator at the time of release. The photon is measured to be redshifted by an amount $v/c = gh/c^2$, or Φ/c^2 with Φ the gravitational potential per unit mass at h. This is the classical gravitational redshift, the simplest nontrivial prediction of general relativity. The gravitational redshift was first 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. Despite all those popular pictures of deformed rubber membranes (one of which graces the cover of these notes), 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 ξ^α our freely falling inertial coordinate frame in which the effects of gravity are locally absent. The local ξ coordinates are erected at a particular point in the spacetime, and are generally not the most convenient to calculate with, but all we need to know for present purposes is that they *exist.* They serve as a foothold to get to a formulation valid for arbitrary coordinates.

In the ξ frame, the equation of motion for a particle is

$$\frac{d^2\xi^\alpha}{d\tau^2} = 0, \tag{103}$$

with

$$c^2 d\tau^2 = -\eta_{\alpha\beta} d\xi^\alpha d\xi^\beta \tag{104}$$

serving as the invariant time interval. If we are doing light, then $d\tau = 0$, but this is less of a problem than it may seem. One may use another scalar parameter, say σ, and set $d\tau = \epsilon\, d\sigma$ letting $\epsilon \to 0$ for light. We will return to this point below, but for the present we note that photon orbits are generally described in terms of coordinate time or angular variables, and all limits are smooth and well-behaved.

Next, we write equation (103) in any other set of coordinates you like, x^μ. Our inertial coordinates ξ^α will be some mathematical function or other of the x^μ so

$$0 = \frac{d^2\xi^\alpha}{d\tau^2} = \frac{d}{d\tau}\left(\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{dx^\mu}{d\tau}\right) \tag{105}$$

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

$$0 = \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{d^2x^\mu}{d\tau^2} + \frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} \tag{106}$$

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

$$\frac{\partial x^\lambda}{\partial\xi^\alpha}\frac{\partial\xi^\alpha}{\partial x^\mu} = \delta^\lambda_\mu \tag{107}$$

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 then becomes

$$\frac{d^2x^\lambda}{d\tau^2} + \Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0, \tag{108}$$

where

$$\Gamma^\lambda_{\mu\nu} = \frac{\partial x^\lambda}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu} \tag{109}$$

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 inverse identity for the second derivatives of ξ^α that we will use shortly:

$$\frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu} = \frac{\partial\xi^\alpha}{\partial x^\lambda}\Gamma^\lambda_{\mu\nu} \tag{110}$$

(How does this work out when used in equation [106]?)

No need to worry, despite all the intimidating notation. (Early editions of relativity texts liked to use **gothic font** $\mathfrak{T}^{\lambda}_{\mu\nu}$ 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,θ) 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, $\Gamma^r_{\theta\theta}=-r$; in the second, $\Gamma^\theta_{r\theta}=1/r$. (What happened to the 2?)

Exercise. Prove the last statements using $\xi^x=r\cos\theta, \xi^y=r\sin\theta$. (In an ordinary, non-distorted flat space, the "local" ξ coordinates are valid everywhere.)

Exercise. On the surface of a unit-radius sphere, choose any point as your North Pole, work in colatitude θ and azimuth ϕ 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 ξ^α coordinates are tied to a local region of the *space* near the North Pole point. In our freely-falling inertial coordinate system, the local coordinates are tied to a point in *spacetime,* and in general are valid only in the neighbourhood of this point: in a distorted space, local coordinates are just that, valid only locally. Of course, "local" can seem pretty extensive to tiny human beings like us. If you want to experience the effects of a local inertial coordinate system in the Earth's gravitational field, you just have to hop. (Sitting in your chair you may feel quite inert, but your frame, alas, is noninertial!)

Returning now to the question of an alternate formulation of equation (108), we may formally pose the question of this expression changes when we change parameters from τ to some other variable σ, assuming there is some functional relation between the two. This is, once again, simply a case of repeated application of the chain rule. We leave it as a straighforward exercise for the reader to show that the resulting equation is

$$\frac{d^2x^\lambda}{d\sigma^2}+\Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}=-\left(\frac{d\tau}{d\sigma}\right)^2\left(\frac{d^2\sigma}{d\tau^2}\right)\frac{dx^\lambda}{d\sigma}. \tag{111}$$

Therefore, the form of the geodesic equation remains invariant if, and only if, τ and σ are linearly related.

3.3 The metric tensor

In our locally inertial coordinates, the invariant spacetime interval is

$$c^2d\tau^2=-\eta_{\alpha\beta}d\xi^\alpha d\xi^\beta, \tag{112}$$

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

$$c^2d\tau^2=-\eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu}dx^\mu dx^\nu\equiv-g_{\mu\nu}dx^\mu dx^\nu \tag{113}$$

where

$$g_{\mu\nu}=\eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu} \tag{114}$$

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 $\Gamma^\lambda_{\mu\nu}$. 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 for now our key goal 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 their explicit reliance on the *local* freely falling inertial coordinates ξ^α, the $g_{\mu\nu}$ and $\Gamma^\lambda_{\mu\nu}$ quantities are awkward to use in their present formulation. Fortunately, there is a direct relationship between $\Gamma^\lambda_{\mu\nu}$ and the first derivatives of $g_{\mu\nu}$ that will free us of local bondage, permitting us to dispense with the ξ^α altogether. As we have emphasised, while their *existence* is crucial to formulate the mathematical structure of curved space, the practical necessity of the ξ's to carry out calculations is minimal.

Differentiate equation (114):

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} = \eta_{\alpha\beta}\frac{\partial^2\xi^\alpha}{\partial x^\lambda \partial x^\mu}\frac{\partial \xi^\beta}{\partial x^\nu} + \eta_{\alpha\beta}\frac{\partial \xi^\alpha}{\partial x^\mu}\frac{\partial^2\xi^\beta}{\partial x^\lambda \partial x^\nu} \tag{115}$$

Now use (110) for the second derivatives of ξ:

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} = \eta_{\alpha\beta}\frac{\partial \xi^\alpha}{\partial x^\rho}\frac{\partial \xi^\beta}{\partial x^\nu}\Gamma^\rho_{\lambda\mu} + \eta_{\alpha\beta}\frac{\partial \xi^\alpha}{\partial x^\mu}\frac{\partial \xi^\beta}{\partial x^\rho}\Gamma^\rho_{\lambda\nu} \tag{116}$$

All remaining ξ derivatives may be absorbed as part of the metric tensor, leading to

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} = g_{\rho\nu}\Gamma^\rho_{\lambda\mu} + g_{\mu\rho}\Gamma^\rho_{\lambda\nu} \tag{117}$$

It remains only to unweave the Γ's from the cloth of indices. This is done by first adding $\partial g_{\lambda\nu}/\partial x^\mu$ to the above (switch λ and μ), then subtracting the same but with indices μ and ν reversed:

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\lambda\mu}}{\partial x^\nu} = g_{\rho\nu}\Gamma^\rho_{\lambda\mu} + \cancel{g_{\mu\rho}\Gamma^\rho_{\lambda\nu}} + g_{\rho\nu}\Gamma^\rho_{\mu\lambda} + \cancel{g_{\rho\lambda}\Gamma^\rho_{\mu\nu}} - \cancel{g_{\rho\mu}\Gamma^\rho_{\nu\lambda}} - \cancel{g_{\rho\lambda}\Gamma^\rho_{\nu\mu}}. \tag{118}$$

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

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\lambda\mu}}{\partial x^\nu} = 2g_{\rho\nu}\Gamma^\rho_{\mu\lambda} \tag{119}$$

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

$$g^{\nu\sigma}g_{\rho\nu} = \delta^\sigma_\rho, \tag{120}$$

leaving us with the pretty result

$$\Gamma^\sigma_{\mu\lambda} = \frac{g^{\nu\sigma}}{2}\left(\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\lambda\mu}}{\partial x^\nu}\right). \tag{121}$$

All mention of the ξ's has vanished. The affine connection is completely specified by $g^{\mu\nu}$ and the derivatives of $g_{\mu\nu}$ in whatever coordinates you like. In practical applications, the inverse matrix is not difficult to find, as we will usually work with metric tensors whose off-diagonal terms either vanish or are sparse.

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},

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

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

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

Exercise. Derive the geodesic equations in spherical (r, θ, ϕ) coordinates, taking $dx^i/d\tau \equiv \dot{x}^i$ to be the ordinary time derivative, and use of course three-dimensional geometry for g_{ij}:

$$\ddot{r} - r(\dot{\phi}^2 \sin^2\theta + \dot{\theta}^2) = 0, \quad r\ddot{\theta} + 2\dot{r}\dot{\theta} - r\dot{\phi}^2 \sin\theta\cos\theta = 0, \quad r\ddot{\phi} + 2\dot{r}\dot{\phi} + 2r\dot{\phi}\dot{\theta}\cot\theta = 0.$$

3.5 Variational calculation of the geodesic equation

The physical significance of the relationship between the metric tensor and affine connection may be further elucidated by a variational calculation. Of all possible paths in our spacetime from some point A to another point B, which path leaves the proper time an extremum (in this case, a maximum)?

The motivation for investigating this formulation is clear. "The shortest distance between two points is a straight line," and in Cartesian coordinates the equations for this line-geodesic are $d^2\xi_i/ds^2 = 0$. This is an elementary property of Euclidian space. We may also pose the question of what is the shortest distance between two points in a more general curved space as well, and this naturally lends itself to a variational approach. The mathematical machinery that was fashioned for generalising the spacelike straight line geodesic 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 use an invariant *timelike* interval for geodesics embedded in distorted Minkowski geometries. Just as ds depends upon the particular path chosen in our Euclidian space, $d\tau$ will depend upon the chosen path in our generalised spacetime geometry. As before, photon orbits are a smooth limit of the finite $d\tau$ dynamics.

Let $\dot{x}^\mu = dx^\mu/d\tau$. The proper time from coordinate location A to location B may be formally written:

$$T_{AB} = \int_A^B d\tau = \frac{1}{c}\int_A^B (-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu)^{1/2}\, d\tau. \tag{122}$$

Next, we allow small but otherwise arbitrary departures from our path by varying x^λ and $\dot{x}^\lambda$. These departures could in principle violate the formal constancy of $-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$ to first order, but via the variational calculation itself, we are enforcing the condition that the contribution from the leading linear order departures in fact vanishes. In effect, we are adopting the final expression as our definition of T_{AB}, and then imposing the constancy of the integrand *after* the first order variational caculation is done. (See Jackson, pp. 583-4 for further discussion on this point.) We then find

$$\delta T_{AB} = \frac{1}{2c}\int_A^B (-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu)^{-1/2}\,\delta\left(-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu\right) d\tau = \frac{1}{2c^2}\;\delta\int_A^B (-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu)\, d\tau \tag{123}$$

since the leading inverse square root in the first equality is just $1/c$ for the zeorth-order solution path. Our problem is now equivalent to an Euler-Lagrange problem with lagrangian

$$\mathcal{L} = g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}. \tag{124}$$

The associated variational equation is:

$$0 = \frac{d}{d\tau}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^\rho}\right) - \frac{\partial \mathcal{L}}{\partial x^\rho} = 2g_{\rho\kappa}\ddot{x}^\kappa + \dot{x}^\mu \dot{x}^\nu (2\partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}) = 2g_{\rho\kappa}\ddot{x}^\kappa + \dot{x}^\mu \dot{x}^\nu (\partial_\nu g_{\rho\mu} + \partial_\mu g_{\rho\nu} - \partial_\rho g_{\mu\nu}). \tag{125}$$

Multiplication by $g^{\lambda\rho}/2$ then leads directly to the standard form of the geoodesic equation:

$$0 = \ddot{x}^\lambda + \frac{g^{\lambda\rho}}{2}(\partial_\nu g_{\rho\mu} + \partial_\mu g_{\rho\nu} - \partial_\rho g_{\mu\nu})\dot{x}^\mu \dot{x}^\nu = \ddot{x}^\lambda + \Gamma^\lambda_{\mu\nu}\dot{x}^\mu \dot{x}^\nu. \tag{126}$$

Thus, if the geodesic equation (108) 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.*

Finally, we re-emphasize that the very practical side to working with the variational method: it is much easier to obtain the equations of motion for a given $g_{\mu\nu}$ this way, as opposed to first calculating the $\Gamma^\lambda_{\mu\nu}$ explicitly. Simply use equations (124) and (125). A by-product of the variational method is that it automatically generates all the nonvanishing affine connection components: just read them off as the coefficients of $(dx^\mu/d\tau)(dx^\nu/d\tau)$. The $\Gamma^\lambda_{\mu\nu}$ are then available for whatever task is at hand.

3.6 The Newtonian limit

In classical mechanics, we 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 does not seem very geometrical. What is the connection with the relativistic Lagrangian (124)? More generally, how can we make contact with Newtonian mechanics from the geodesic equation?

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 $cdt \gg |d\boldsymbol{x}|$, the geodesic equation greatly simplifies:

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{00}\left(\frac{cdt}{d\tau}\right)^2 = 0. \tag{127}$$

Now

$$\Gamma^\mu_{00} = \frac{1}{2}g^{\mu\nu}\left(\frac{\partial g_{0\nu}}{\partial(cdt)} + \frac{\partial g_{0\nu}}{\partial(cdt)} - \frac{\partial g_{00}}{\partial x^\nu}\right) \tag{128}$$

In the static Newtonian limit (meaning the time variations of the field are nonrelativistic), the largest of the g derivatives are contained in the final term. This leads to:

$$\Gamma^\mu_{00} \simeq -\frac{1}{2}g^{\mu\nu}\frac{\partial g_{00}}{\partial x^\nu}. \tag{129}$$

Since the gravitational field is weak, $g_{\mu\nu}$ differs very little from its Minkoswki limiting value:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad h_{\mu\nu} \ll 1. \tag{130}$$

Neglecting the difference between t and τ in this limit, the spatial components of the geodesic equation are therefore

$$\frac{d^2\boldsymbol{x}}{dt^2} - \frac{c^2}{2}\boldsymbol{\nabla} h_{00} = 0. \tag{131}$$

Isaac Newton would say:

$$\frac{d^2\boldsymbol{x}}{dt^2} + \boldsymbol{\nabla}\Phi = 0, \tag{132}$$

with Φ being the classical gravitational potential. The two views are perfectly consistent if

$$h_{00} \simeq -\frac{2\Phi}{c^2}, \quad g_{00} \simeq -\left(1 + \frac{2\Phi}{c^2}\right) \tag{133}$$

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 for orbital mechanics is all in the time component, h_{00}. It is now possible to measure this term directly by 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. We have chosen the constant by making the geometry Minkowskian at large distances from any matter.

For self-consistency of the $t \simeq \tau$ assumption, let us check the $\mu = 0$ geodesic equation in the static Newtonian limit. Only the $\Gamma^0_{0i}, \Gamma^0_{i0}$ terms contribute, where i is a spatial coordinate. This leads to

$$\frac{d^2t}{d\tau^2} - \frac{\partial h_{00}}{\partial x^i}\frac{dx^i}{d\tau}\frac{dt}{d\tau} = 0, \tag{134}$$

summed over i. Since

$$\frac{\partial h_{00}}{\partial x^i}\frac{dx^i}{d\tau} = \frac{dh_{00}}{d\tau},$$

this is just a first order ordinary differential equation for $dt/d\tau$. The solution is

$$\frac{dt}{d\tau} = C\exp(h_{00}),$$

where C is a constant of integration. Assuming that h_{00} vanishes at large distances from the source, C may be chosen to be 1. With $h_{00} \ll 1$, the leading order correction term is

$$dt \simeq (1 + h_{00})d\tau. \tag{135}$$

This $\sim v^2/c^2$ correction justifies our approximation of setting $d\tau = dt$ in the spatial geodesic equation of motion.

At the surface of a spherical object of mass M and radius R,

$$h_{00} \simeq \frac{2GM}{Rc^2} \simeq 4.2 \times 10^{-6}\left(\frac{M}{M_\odot}\right)\left(\frac{R_\odot}{R}\right) \tag{136}$$

where $M_\odot$ is the mass of the Sun (about 2×10^{30} kg) and $R_\odot$ is the radius of the Sun (about 7×10^8 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 with a one metre separation in height on the Earth's surface?

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

$$\int \left[c^2(1+2\Phi/c^2)dt^2 - d|\boldsymbol{x}|^2\right]^{1/2} dt \tag{137}$$

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

$$\int c\left(1 + \frac{\Phi}{c^2} - \frac{v^2}{2c^2} + ...\right) dt \tag{138}$$

where $v^2 \equiv (d|\boldsymbol{x}|/dt)^2$. 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

$$c^2 d\tau^2 = -g_{\mu\nu} dx^\mu dx^\nu \tag{139}$$

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

$$d\tau(x) = [-g_{00}(x)]^{1/2} dt \tag{140}$$

where dt is the time interval registered at infinity, where $-g_{00} \to 1$. (Compare: the "proper length" on the unit sphere for an interval at constant θ 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,

$$\frac{d\tau(y)}{d\tau(x)} = \left[\frac{g_{00}(y)}{g_{00}(x)}\right]^{1/2}, \tag{141}$$

and in particular

$$\frac{d\tau(R)}{dt} = \frac{\nu(\infty)}{\nu} = [-g_{00}(R)]^{1/2} \tag{142}$$

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 from a source at rest at $r = R$ is

$$\nu_\infty = \left(1 - \frac{2GM}{Rc^2}\right)^{1/2} \nu \tag{143}$$

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

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," as the leading astrophysicist of his time, A.S. Eddington remarked. For it proved not only 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 spectral redshift of light due to a gravitational field.

Alas, the modern consensus is that these measurements were flawed. Adams knew what he was looking for, and he found it, a case of "confirmation bias." Though premature, the activity this apparently positive observation imparted to the study of white dwarfs and 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. In 1976, an experiment using rocket-borne hydrogen masers measured the gravitational redshift by comparing clocks at an altitude of 10,000 km with ground-based clocks, and found agreement with Einstein's prediction to one part in 10^4.

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 Δt from when the photon escapes is

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

where δm 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 δx 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 δx, 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 gravitational 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." Equation (127), which may be written (with $\dot{x} \equiv dx/d\tau$)

$$\ddot{x}^i + \Gamma^i_{00}(\dot{x}^0)^2 = 0,$$

should be compared with the standard flat space, ordinary time, radial centripetal force equation in affine connection form:

$$\ddot{r} + \Gamma^r_{\phi\phi}(\dot{\phi})^2 = 0$$

The analogy is unmistakable. Newtonian gravity is a centripetal force arising from the flow of time going around a bend in space. Close to a source, the effect of weak gravity appears as a distortion predominantly in the structure of time. Far from a source, weak gravity generally takes the form of gravitational radiation, and this, we shall see, may be thought of (in the most natural coordinates) as a distortion of the structure 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, $x \to x'$.

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 $\Gamma^{\lambda}_{\mu\nu}$. 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^μ. If $x'^\mu = x'^\mu(x)$, there is no doubt how the dx'^μ are related to the dx^μ. It is called the chain rule, and it is by now very familiar:

$$dx'^\mu = \frac{\partial x'^\mu}{\partial x^\nu}\, dx^\nu \tag{144}$$

Be careful to distinguish between the coordinates x^μ, which are almost always mere vector imposters and can be pretty much anything, and their differentials dx^μ, which are true vectors. Equation (144) implies that the x'^μ coordinates may be constructed from knowledge of dx^μ in a manner that is independent of the path chosen, i.e., the dx'^μ are exact differentials. Any set of quantities V^μ that transforms in this way is known as a *contravariant vector*:

$$V'^\mu = \frac{\partial x'^\mu}{\partial x^\nu}\, V^\nu \tag{145}$$

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

$$V'_\mu = \frac{\partial x^\nu}{\partial x'^\mu}\, V_\nu \tag{146}$$

"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 Λ^α_β and $\tilde{\Lambda}^\beta_\alpha$ 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 Φ. Once again, the chain rule tells us how to transform from one set of coordinates to another—we've no choice:

$$\frac{\partial \Phi}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial \Phi}{\partial x^\nu} \tag{147}$$

The generalisation to tensor transformation laws is immediate. A contravariant tensor $T^{\mu\nu}$ transforms as though it were a vector direct product $V^\mu V^\nu$:

$$T'^{\mu\nu} = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma}T^{\rho\sigma}. \tag{148}$$

A covariant tensor $T_{\mu\nu}$ then transforms as

$$T'_{\mu\nu} = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu}T_{\rho\sigma}, \tag{149}$$

and a mixed tensor T^μ_ν as

$$T'^\mu_\nu = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x'^\nu}T^\rho_\sigma. \tag{150}$$

The direct product formalism ($V^\mu V^\nu$, $V^\mu V_\nu$, etc.) is a useful way to think of tensorial indices. The generalisation of the transformation laws to mixed tensors of arbitrary rank is self-evident.

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

$$g'_{\mu\nu} \equiv \eta_{\alpha\beta}\frac{\partial \xi^\alpha}{\partial x'^\mu}\frac{\partial \xi^\beta}{\partial x'^\nu} = \eta_{\alpha\beta}\frac{\partial \xi^\alpha}{\partial x^\lambda}\frac{\partial \xi^\beta}{\partial x^\rho}\frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial x^\rho}{\partial x'^\nu} \equiv g_{\lambda\rho}\frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial x^\rho}{\partial x'^\nu} \tag{151}$$

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:

$$\frac{\partial A'^\lambda}{\partial x'^\mu} = \frac{\partial}{\partial x'^\mu}\left(\frac{\partial x'^\lambda}{\partial x^\nu}A^\nu\right) = \frac{\partial x'^\lambda}{\partial x^\nu}\frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial A^\nu}{\partial x^\rho} + \frac{\partial^2 x'^\lambda}{\partial x^\rho \partial x^\nu}\frac{\partial x^\rho}{\partial x'^\mu}A^\nu \tag{152}$$

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 alas 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. We have noted that direct products of tensors always transform as tensors:

$$W^{\mu\nu}_{\rho\sigma} = T^{\mu\nu}S_{\rho\sigma}. \tag{153}$$

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

$$T'_\mu{}^\rho \equiv g'_{\mu\nu}T'^{\nu\rho} = \frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial x^\lambda}{\partial x'^\nu}\frac{\partial x'^\nu}{\partial x^\kappa}\frac{\partial x'^\rho}{\partial x^\tau}g_{\sigma\lambda}T^{\kappa\tau} = \frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial x'^\rho}{\partial x^\tau}g_{\sigma\kappa}T^{\kappa\tau} = \frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial x'^\rho}{\partial x^\tau}T_\sigma{}^\tau \tag{154}$$

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 raised by multiplication with $g^{\rho\mu}$.)

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

$$T'^\mu_{\ \mu} = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\rho}{\partial x'^\mu}T^\nu_{\ \rho} = \delta^\rho_\nu T^\nu_{\ \rho} = T^\nu_{\ \nu} \tag{155}$$

i.e., $T^\mu_{\ \mu}$ is a scalar T (just like $V^\mu V_\mu$). Exactly the same type of calculation shows that $T^{\mu\nu}_{\ \ \ \mu}$ is a vector T^ν, and so on. Remember always to contract "up–down:" $T^\mu_{\ \mu} = T$, *never* $T^{\mu\mu} = T$.

The generalisation of the familiar scalar dot product between vectors A^μ and B^μ 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

$$\cos\Delta\theta = \frac{A^i B_i}{(A^j A_j\ B^k B_k)^{1/2}} = \frac{g_{ij}A^i B^j}{(g_{kl}A^k A^l\ g_{mn}B^m B^n)^{1/2}} \tag{156}$$

the analogue of $\boldsymbol{A}\cdot\boldsymbol{B}/(|\boldsymbol{A}||\boldsymbol{B}|)$. 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

4.2.1 Contravariant vectors

Knowing the general form of the transformation law for a vector allows us to appreciate more fully the content of the geodesic equation. With $V'^\alpha = d\xi^\alpha/d\tau$, the geodesic equation directly 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 more clearly, consider the derivative of any vector $dA'^\alpha/d\tau$, expressed here in local inertial coordinates ξ^α. In these coordinates this is the correct form of the derivative. Then, upon going to arbitrary coordinates x^μ and vector A^μ, exactly the same manipulations we used to derive the geodesic equation yield,

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

(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

$$\frac{\partial x^\lambda}{\partial\xi^\alpha}\frac{dA'^\alpha}{d\tau} = \frac{dA^\lambda}{d\tau} + \Gamma^\lambda_{\mu\nu}A^\mu\frac{dx^\nu}{d\tau}. \tag{157}$$

The left hand side is now precisely the formal transformation law of $dA'^\alpha/d\tau$ from local inertial to general coordinates. The right side tells us what that expression is, expressed

in these general coordinates. Evidently, the combination on the right of this equation is a general vector quantity unto itself!

We will return to this very useful 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 (157) may be written as

$$V^\nu \left[\frac{\partial A^\lambda}{\partial x^\nu} + \Gamma^\lambda_{\mu\nu} A^\mu \right] \tag{158}$$

Since V^ν is a vector and A^λ is arbitrary, the expression in square brackets must be a mixed tensor of rank two: it contracts with a contravariant vector V^ν to produce another contravariant vector. Recall that $\Gamma^\lambda_{\mu\nu}$ 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

$$\frac{\partial A^\lambda}{\partial x^\mu} + \Gamma^\lambda_{\mu\nu} A^\nu \equiv A^\lambda_{;\mu} \tag{159}$$

the so called *covariant derivative.* Following convention, we shall use a semicolon to denote covariant differentiation. (Some authors use a comma for partial derivatives (e.g $A^\nu_{,\,\mu}$), but it is more clear to use full partial derivative notation [or ∂_μ], and we shall abide by this practise in these notes.) We now have a generalisation of the standard partial derivative to tensor form[4].

Note that the standard geodesic equation may be written entirely in terms of the 4-velocity V^μ:

$$V^\nu V^\mu_{;\nu} = 0, \tag{160}$$

the relativistic analogue of $\partial_t v^i + v^j \partial_j v^i = 0$. Moreover, if (160) is written in terms of the 4-momentum p^μ,

$$p^\nu p^\mu_{;\nu} = 0, \tag{161}$$

we obtain a geodesic equation valid at once for photons as well as massive particles, without the need to appeal to a second affine parameter for the photon equation. Since a photon's wavenumber k^μ is directly proportional to its 4-momentum, we may replace p^μ in (160) with k^μ, thereby obtaining an important covariant equation for geometrical optics.

4.2.2 Transformation of the affine connection

Note: This section is off-syllabus and non-examinable.

Let us examine our result in more detail, and in the process learn how the non-tensor $\Gamma^\lambda_{\mu\nu}$ changes under general coordinate transformations. We shall also see how to construct the covariant derivative of covariant vectors, and of yet more general tensors.[5]

We begin with the transformation law for $\Gamma^\lambda_{\mu\nu}$. This is just repeated application of the chain rule:

$$\Gamma'^\lambda_{\mu\nu} \equiv \frac{\partial x'^\lambda}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x'^\mu \partial x'^\nu} = \frac{\partial x'^\lambda}{\partial x^\rho} \frac{\partial x^\rho}{\partial \xi^\alpha} \frac{\partial}{\partial x'^\mu} \left(\frac{\partial x^\sigma}{\partial x'^\nu} \frac{\partial \xi^\alpha}{\partial x^\sigma} \right) \tag{162}$$

[4] Gauge-covariant derivatives in quantum field theories are a closely related concept.

[5] Notice the use of the word "covariant" in two very different senses. This is awkward, but completely standard, mathematical nomenclature.

Carrying through the derivative,

$$\Gamma'^{\lambda}_{\mu\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial \xi^{\alpha}} \left(\frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial^2 \xi^{\alpha}}{\partial x^{\tau} \partial x^{\sigma}} + \frac{\partial^2 x^{\sigma}}{\partial x'^{\mu} \partial x'^{\nu}} \frac{\partial \xi^{\alpha}}{\partial x^{\sigma}} \right) \tag{163}$$

Cleaning up, and recognising an affine connection when we see one, helps get rid of these meddlesome ξ's:

$$\Gamma'^{\lambda}_{\mu\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma^{\rho}_{\tau\sigma} + \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\mu} \partial x'^{\nu}} \tag{164}$$

Because of the presence of the final term, this is not the transformation law of any sort of tensor. But consider now a *differential* change in the affine connection, $\delta\Gamma^{\lambda}_{\mu\nu}$, which arises from a variation in the metric tensor itself, $\delta g_{\mu\nu}$. This might be physically caused by a propagating wave or it may just be part of a formal mathematical calculation. The point is that the coordinate systems x^{μ} and x'^{μ} in equation (164) are unchanged by the variation $\delta g_{\mu\nu}$. Under these conditions the transformation law for $\delta\Gamma^{\lambda}_{\mu\nu}$ is

$$\delta\Gamma'^{\lambda}_{\mu\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \delta\Gamma^{\rho}_{\tau\sigma}, \tag{165}$$

because the final term involves only the unchanged coordinates, and thus vanishes when the differential is calculated. Hence, $\delta\Gamma^{\lambda}_{\mu\nu}$ *is* a true tensor, with all the privileges that entails.

Equation (164) may also be written

$$\Gamma'^{\lambda}_{\mu\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma^{\rho}_{\tau\sigma} - \frac{\partial x^{\rho}}{\partial x'^{\nu}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \frac{\partial^2 x'^{\lambda}}{\partial x^{\sigma} \partial x^{\rho}} \tag{166}$$

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

$$\frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial x'^{\nu}} = \delta^{\lambda}_{\nu}.)$$

Hence

$$\Gamma'^{\lambda}_{\mu\nu} A'^{\nu} = \left(\frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma^{\rho}_{\tau\sigma} - \frac{\partial x^{\rho}}{\partial x'^{\nu}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \frac{\partial^2 x'^{\lambda}}{\partial x^{\rho} \partial x^{\sigma}} \right) \frac{\partial x'^{\nu}}{\partial x^{\eta}} A^{\eta}, \tag{167}$$

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

$$\Gamma'^{\lambda}_{\mu\nu} A'^{\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \Gamma^{\rho}_{\tau\sigma} A^{\sigma} - \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \frac{\partial^2 x'^{\lambda}}{\partial x^{\rho} \partial x^{\sigma}} A^{\rho} \tag{168}$$

Finally, adding this to (152), the unwanted terms cancel just as they should. We thus obtain

$$\frac{\partial A'^{\lambda}}{\partial x'^{\mu}} + \Gamma'^{\lambda}_{\mu\nu} A'^{\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\nu}} \frac{\partial x^{\rho}}{\partial x'^{\mu}} \left(\frac{\partial A^{\nu}}{\partial x^{\rho}} + \Gamma^{\nu}_{\rho\sigma} A^{\sigma} \right), \tag{169}$$

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

4.2.3 Covariant vectors

It is now a one-step process to deduce how covariant derivatives work for covariant vectors. With V^λ arbitrary, consider the contraction

$$V_\lambda A^\lambda_{;\mu} = V_\lambda \frac{\partial A^\lambda}{\partial x^\mu} + \Gamma^\lambda_{\mu\nu} A^\nu V_\lambda \tag{170}$$

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

$$\frac{\partial (A^\lambda V_\lambda)}{\partial x^\mu} - A^\lambda \left[\frac{\partial V_\lambda}{\partial x^\mu} - \Gamma^\nu_{\mu\lambda} V_\nu \right]. \tag{171}$$

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

$$V_{\lambda;\mu} = \frac{\partial V_\lambda}{\partial x^\mu} - \Gamma^\nu_{\mu\lambda} V_\nu \tag{172}$$

That this really *is* a vector may also be directly verified via a calculation exactly similar to our previous one for the covariant derivative of a contravariant vector. From equation (172), we may easily write the covariant form of the derivative $dV_\lambda/d\tau$, valid in general coordinates:

$$\frac{dx^\mu}{d\tau} V_{\lambda;\mu} = \frac{dV_\lambda}{d\tau} - \Gamma^\nu_{\mu\lambda} \frac{dx^\mu}{d\tau} V_\nu \tag{173}$$

I certainly hope that after all these examples you will not make the cardinal sin of believing that if V^μ is a vector, then so must be its mathematical differential dV^μ. Can you see how to instantly destroy that belief using the transformation law for a vector? By contrast, coordinate differentials dx^μ certainly *are* vectors; it is the coordinates x^μ themselves that are not. They are simply a group of well-posed mathematical functions.

4.2.4 Tensors of arbitrary rank

Covariant derivatives of *tensors* are now simple to deduce, since they have exactly the transformation properties of direct products of vectors. We have noted that a tensor $T^{\lambda\kappa}$, for example, must formally transform like a direct product $A^\lambda V^\kappa$. Well, in local inertial coordinates the covariant derivative of this is just the ordinary derivative

$$\frac{\partial (A^\lambda V^\kappa)}{\partial x^\mu} = A^\lambda \frac{\partial V^\kappa}{\partial x^\mu} + \frac{\partial A^\lambda}{\partial x^\mu} V^\kappa.$$

But we already know how to generalise the partial vector derivatives on the right to arbitrary coordinates! This leads immediately to

$$(A^\lambda V^\kappa)_{;\mu} = \frac{\partial (A^\lambda V^\kappa)}{\partial x^\mu} + \Gamma^\kappa_{\nu\mu} A^\lambda V^\nu + \Gamma^\lambda_{\mu\nu} A^\nu V^\kappa \tag{174}$$

We know that this will transform as a mixed tensor of rank 3, which means that we may replace the direct product $A^\lambda V^\kappa$ with any second rank contrariant tensor $T^{\lambda\kappa}$ in the above formula:

$$T^{\lambda\kappa}_{;\mu} = \frac{\partial T^{\lambda\kappa}}{\partial x^\mu} + \Gamma^\lambda_{\mu\nu} T^{\nu\kappa} + \Gamma^\kappa_{\nu\mu} T^{\lambda\nu} \tag{175}$$

Identical reasoning leads to

$$T_{\lambda\kappa;\mu} = \frac{\partial T_{\lambda\kappa}}{\partial x^\mu} - \Gamma^\nu_{\lambda\mu} T_{\nu\kappa} - \Gamma^\nu_{\kappa\mu} T_{\lambda\nu} \tag{176}$$

and of course

$$T^\lambda_{\kappa;\mu} = \frac{\partial T^\lambda_\kappa}{\partial x^\mu} + \Gamma^\lambda_{\nu\mu} T^\nu_\kappa - \Gamma^\nu_{\mu\kappa} T^\lambda_\nu \tag{177}$$

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 drill until it is second-nature.

4.2.5 Going from special to general relativity

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 scalars, 4-vectors and 4-tensors. Replace $\eta_{\alpha\beta}$ with $g_{\mu\nu}$. Replace standard partial derivatives with covariant derivatives.* Voilà: your equation is set for the presence of gravitational fields. This, by the way, is not merely a presecription to do general relativity: it is the way to turn any set of properly framed mathematical equations written in Cartesian coordinates into the corresponding equations valid in any coordinates you fancy.

It will not have escaped your attention, I am sure, that applying (176) to $g_{\mu\nu}$ produces

$$g_{\mu\nu;\lambda} = \frac{\partial g_{\mu\nu}}{\partial x^\lambda} - g_{\rho\nu}\Gamma^\rho_{\mu\lambda} - g_{\mu\rho}\Gamma^\rho_{\nu\lambda} = 0 \tag{178}$$

where equation (117) 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 tensorial covariant derivative of $g_{\mu\nu}$ must also 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 $\boldsymbol{V}$ may be expanded in a set of basis vectors,

$$\boldsymbol{V} = V^a \boldsymbol{e}_a \tag{179}$$

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:

$$\boldsymbol{V} = V_b \boldsymbol{e}^b \tag{180}$$

Same $\boldsymbol{V}$ mind you, just different ways of representing its components. If the $\boldsymbol{e}$'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,

$$\boldsymbol{e}_a \cdot \boldsymbol{e}^{\boldsymbol{b}} = \delta_a^b. \tag{181}$$

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

$$\boldsymbol{e'}_a = \frac{\partial x^b}{\partial x'^a} \boldsymbol{e_b}, \quad \boldsymbol{e'}^a = \frac{\partial x'^a}{\partial x^b} \boldsymbol{e}^{\boldsymbol{b}}. \tag{182}$$

Note that the dot product rule gives

$$\boldsymbol{V} \cdot \boldsymbol{V} = V^a V_b \boldsymbol{e}_a \cdot \boldsymbol{e}^{\boldsymbol{b}} = V^a V_b \delta_a^b = V^a V_a, \tag{183}$$

as we would expect. On the other hand, expanding the differential line element $\boldsymbol{ds}$,

$$ds^2 = \boldsymbol{e_a} dx^a \cdot \boldsymbol{e_b} dx^b = \boldsymbol{e_a} \cdot \boldsymbol{e_b} dx^a dx^b \tag{184}$$

so that we recover the metric tensor

$$g_{ab} = \boldsymbol{e_a} \cdot \boldsymbol{e_b} \tag{185}$$

Exactly the same style calculation gives

$$g^{ab} = \boldsymbol{e^a} \cdot \boldsymbol{e^b} \tag{186}$$

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

$$\boldsymbol{e}_b = g_{ab} \boldsymbol{e}^a, \tag{187}$$

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

$$\boldsymbol{e}^b = g^{ab} \boldsymbol{e_a} \tag{188}$$

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

As with any vector field, basis vectors will generally change with coordinate position. We *define* a thrice-indexed object Γ^b_{ac} by

$$\frac{\partial \boldsymbol{e}_a}{\partial x^c} = \Gamma^b_{ac}\boldsymbol{e}_b \tag{189}$$

so that

$$\Gamma^b_{ac} = \boldsymbol{e}^b{\cdot}\partial_c\boldsymbol{e}_a \equiv \partial_c(\boldsymbol{e}_a{\cdot}\boldsymbol{e}^b) - \boldsymbol{e}_a{\cdot}\partial_c\boldsymbol{e}^b = -\boldsymbol{e}_a{\cdot}\partial_c\boldsymbol{e}^b. \tag{190}$$

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

$$\frac{\partial \boldsymbol{e}^b}{\partial x^c} = -\Gamma^b_{ac}\boldsymbol{e}^a \tag{191}$$

Consider $\partial_c g_{ab} = \partial_c(\boldsymbol{e}_a{\cdot}\boldsymbol{e}_b)$. Using (189),

$$\partial_c g_{ab} = (\partial_c\boldsymbol{e}_a){\cdot}\boldsymbol{e}_b + \boldsymbol{e}_a{\cdot}(\partial_c\boldsymbol{e}_b) = \Gamma^d_{ac}\boldsymbol{e}_d{\cdot}\boldsymbol{e}_b + \boldsymbol{e}_a{\cdot}\Gamma^d_{bc}\boldsymbol{e}_d, \tag{192}$$

or finally

$$\partial_c g_{ab} = \Gamma^d_{ac}g_{db} + \Gamma^d_{bc}g_{ad}, \tag{193}$$

exactly what we found in (117)! This leads, in turn, precisely to (121), the equation for the affine connection in terms of the g partial derivatives. We now have a more intuitive understanding of what the Γ'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 Γ^b_{ac} just go away.

Finally, consider

$$\partial_a(V^b\boldsymbol{e}_b) = (\partial_a V^b)\boldsymbol{e}_b + V^b\partial_a\boldsymbol{e}_b = (\partial_a V^b)\boldsymbol{e}_b + V^b\Gamma^c_{ab}\boldsymbol{e}_c \tag{194}$$

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

$$\boldsymbol{e}^d{\cdot}\partial_a(V^b\boldsymbol{e}_b) = \partial_a V^d + V^b\Gamma^d_{ab} \equiv V^d_{\ ;a}, \tag{195}$$

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

$$\boldsymbol{e}_d{\cdot}\partial_a(V_b\boldsymbol{e}^b) = \partial_a V_d - V_b\Gamma^b_{ad} \equiv V_{d;a}, \tag{196}$$

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 $\boldsymbol{V}$ are not simply $\partial_a V^d$ and $\partial_a V_d$, but rather $\boldsymbol{e}^d{\cdot}\partial_a(V^b\boldsymbol{e}_b)$ and $\boldsymbol{e}_d{\cdot}\partial_a(V_b\boldsymbol{e}^b)$ respectively. The Γ-terms then emerge as derivatives of the $\boldsymbol{e}$ basis vectors. Our terse, purely coordinate notation avoids the use of the $\boldsymbol{e}$ 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:

$$g'_{\mu\nu} = \frac{\partial x^\kappa}{\partial x'^\mu} g_{\kappa\lambda} \frac{\partial x^\lambda}{\partial x'^\nu} \tag{197}$$

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

$$g' = \left| \frac{\partial x}{\partial x'} \right|^2 g \tag{198}$$

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 $x \to x'$. The significance of this result is that there is another quantity that also transforms with a Jacobian factor: the volume element d^4x.

$$d^4x' = \left| \frac{\partial x'}{\partial x} \right| d^4x. \tag{199}$$

This means

$$\sqrt{-g'}\, d^4x' = \sqrt{-g} \left| \frac{\partial x}{\partial x'} \right| \left| \frac{\partial x'}{\partial x} \right| d^4x = \sqrt{-g}\, d^4x. \tag{200}$$

In other words, $\sqrt{-g}\, d^4x$ 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^4x 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 $dx\, dy\, dz = RdR\, dz\, d\phi = r^2 \sin\theta\, dr\, d\theta\, d\phi$. 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

Ordinary partial derivatives of a scalar transform as covariant vectors, so in the case of a simple gradient there is no distinction between a covariant and standard partial derivative. Another result that is easily adpated is the covariant curl,

$$V_{\mu;\nu} - V_{\nu;\mu} = \frac{\partial V_\mu}{\partial x^\nu} - \frac{\partial V_\nu}{\partial x^\mu}. \tag{201}$$

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

$$V^\mu_{\ ;\mu} = \frac{\partial V^\mu}{\partial x^\mu} + \Gamma^\mu_{\mu\lambda} V^\lambda \tag{202}$$

where by definition

$$\Gamma^\mu_{\mu\lambda} = \frac{g^{\mu\rho}}{2} \left(\frac{\partial g_{\rho\mu}}{\partial x^\lambda} + \frac{\partial g_{\rho\lambda}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\rho} \right). \tag{203}$$

Now the lead factor $g^{\mu\rho}$ is symmetric in its indices, whereas the last two g derivatives combined are antisymmetric in the same indices, so that combination disappears when summed over $\mu\rho$. We are left with

$$\Gamma^{\mu}_{\mu\lambda} = \frac{g^{\mu\rho}}{2}\frac{\partial g_{\rho\mu}}{\partial x^{\lambda}}. \tag{204}$$

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,

$$\Gamma^{\mu}_{\mu\lambda} = \frac{1}{2}\frac{\partial \ln|g|}{\partial x^{\lambda}}, \tag{205}$$

where g is as usual the determinant of $g_{\mu\nu}$, here just the product of the diagonal elements. Though this result has only been derived for diagonal $g_{\mu\nu}$, W72 pp. 106-7, shows that this result is true for *any* $g_{\mu\nu}$.[6]

The covariant divergence (202) becomes

$$V^{\mu}_{;\mu} = \frac{1}{\sqrt{|g|}}\frac{\partial(\sqrt{|g|}V^{\mu})}{\partial x^{\mu}} \tag{206}$$

a neat and tidy result. Note that the vanishing of the scalar integral

$$\int \sqrt{|g|}d^4x\, V^{\mu}_{;\mu} = 0, \tag{207}$$

if V^{μ} vanishes sufficiently rapidly at infinity.

Exercise. Apply our divergence formula (206) to the case of standard spherical coordinates in three dimensions and compare with the usual formula for the divergence. What do you notice about the terms involving V^{θ} and V^{ϕ}? What needs to be done to get agreement? Explain.

Another Γ-contraction of interest is

$$\Gamma^{\sigma} \equiv g^{\mu\lambda}\Gamma^{\sigma}_{\mu\lambda} = g^{\mu\lambda}\frac{g^{\nu\sigma}}{2}\left(\frac{\partial g_{\mu\nu}}{\partial x^{\lambda}} + \frac{\partial g_{\lambda\nu}}{\partial x^{\mu}} - \frac{\partial g_{\lambda\mu}}{\partial x^{\nu}}\right) \tag{208}$$

The first two sums on the right are identical because of the $\mu\lambda$ symmetry, so focus on the first one. Using (justify!):

$$g^{\mu\lambda}\frac{\partial g_{\mu\nu}}{\partial x^{\lambda}} = -g_{\mu\nu}\frac{\partial g^{\mu\lambda}}{\partial x^{\lambda}},$$

we multiply by $g^{\nu\sigma}$, so that a Kronecker δ^{σ}_{μ} conveniently appears on the right. This gives $-\partial_{\lambda}g^{\sigma\lambda}$ as the sum of the first two terms on the right side of (208). The final term on the right side of this equation is now recognisable as our new friend (focus on $\mu\lambda$):

$$-g^{\mu\lambda}\frac{g^{\nu\sigma}}{2}\frac{\partial g_{\lambda\mu}}{\partial x^{\nu}} = -g^{\nu\sigma}\Gamma^{\mu}_{\mu\nu} = -\frac{g^{\nu\sigma}}{2}\frac{\partial \ln|g|}{\partial x^{\nu}},$$

[6] Sketchy proof for the mathematically inclined: For matrix M, trace Tr, differential δ, to first order in δ we have $\delta \ln\det M = \ln\det(M+\delta M) - \ln\det M = \ln\det M^{-1}(M+\delta M) = \ln\det(1+M^{-1}\delta M) = \ln(1+\mathrm{Tr}\, M^{-1}\delta M) = \mathrm{Tr}\, M^{-1}\delta M$. Can you supply the missing details?

to whom we have been introduced in (205). Adding this to $-\partial_\lambda g^{\sigma\lambda}$, we finally have

$$\Gamma^\sigma = g^{\mu\lambda}\Gamma^\sigma_{\mu\lambda} = -\frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^\nu}\left(\sqrt{|g|}g^{\nu\sigma}\right) \tag{209}$$

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

$$T^{\mu\nu}_{\ \ ;\mu} = \frac{\partial T^{\mu\nu}}{\partial x^\mu} + \Gamma^\mu_{\mu\lambda}T^{\lambda\nu} + \Gamma^\nu_{\mu\lambda}T^{\mu\lambda}, \text{ or } \quad T^\mu_{\ \nu;\mu} = \frac{\partial T^\mu_{\ \ \nu}}{\partial x^\mu} + \Gamma^\mu_{\mu\lambda}T^\lambda_{\ \ \nu} - \Gamma^\lambda_{\mu\nu}T^\mu_{\ \ \lambda} \tag{210}$$

and using (205), we may condense this to

$$T^{\mu\nu}_{\ \ ;\mu} = \frac{1}{\sqrt{|g|}}\frac{\partial(\sqrt{|g|}T^{\mu\nu})}{\partial x^\mu} + \Gamma^\nu_{\mu\lambda}T^{\mu\lambda}, \quad \text{or} \quad T^\mu_{\ \nu;\mu} = \frac{1}{\sqrt{|g|}}\frac{\partial(\sqrt{|g|}T^\mu_{\ \ \nu})}{\partial x^\mu} - \Gamma^\lambda_{\mu\nu}T^\mu_{\ \ \lambda}. \tag{211}$$

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

$$A^{\mu\nu}_{\ \ ;\mu} = \frac{1}{\sqrt{|g|}}\frac{\partial(\sqrt{|g|}A^{\mu\nu})}{\partial x^\mu} \quad \text{if } A^{\mu\nu} \text{ antisymmetric.} \tag{212}$$

Exercise. Don't give up if $T^{\mu\nu}$ is symmetric. Prove that if $T^{\mu\nu}$ is *symmetric* in its indices *and* the metric tensor is independent of coordinate x^ν, then

$$T^\mu_{\ \nu;\mu} = \frac{1}{\sqrt{|g|}}\frac{\partial(\sqrt{|g|}T^\mu_{\ \ \nu})}{\partial x^\mu}.$$

Black holes, for example, have spacetime geometries independent of ϕ, and an ideal gas has a symmetric stress tensor $T^{\mu\nu}$, so this equation is very useful for astrophysicists.

Exercise. Define $\Gamma_{\sigma\mu\nu} = g_{\sigma\lambda}\Gamma^\lambda_{\mu\nu}$. Find an explicit expression for $\Gamma_{\sigma\mu\nu}$ and prove that

$$\sum_{\sigma\mu\nu}\Gamma_{\sigma\mu\nu} = \sum_{\sigma\mu\nu}\partial_\sigma g_{\mu\nu}$$

where the sum is over the three even permutations of $\sigma\mu\nu$.

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

$$T^\mu_{\ \nu;\mu} = 0$$

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

$$T^{\mu}{}_{\nu} = P\delta^{\mu}{}_{\nu} + (\rho + P/c^2)U^{\mu}U_{\nu} \tag{213}$$

(The expression for the mixed $T^{\mu}{}_{\nu}$ is actually unchaged from special relativity. Otherwise, one simply changes η's to g's.) Our conservation equation is

$$0 = T^{\mu}{}_{\nu;\mu} = \frac{\partial P}{\partial x^{\nu}} + \left[(\rho + P/c^2)U^{\mu}U_{\nu}\right]_{;\mu} \tag{214}$$

where we have made use of the Leibniz rule for the covariant derivative of a product, and the fact that the δ^{μ}_{ν} covariant derivative vanishes. Using (211):

$$0 = \frac{\partial P}{\partial x^{\nu}} + \frac{1}{|g|^{1/2}}\frac{\partial}{\partial x^{\mu}}\left[|g|^{1/2}(\rho + P/c^2)U^{\mu}U_{\nu}\right] - \Gamma^{\mu}_{\nu\lambda}(\rho + P/c^2)U_{\mu}U^{\lambda} \tag{215}$$

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

$$U^{\mu}U_{\mu} = -c^2, \tag{216}$$

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

$$U_0U^0 = -c^2, \tag{217}$$

and with

$$\Gamma^0_{\nu 0} = \frac{1}{2}\frac{\partial \ln|g_{00}|}{\partial x^{\nu}}, \tag{218}$$

a vanishing second term on the right of equation (215), we then find

$$0 = \frac{\partial P}{\partial x^{\nu}} + \left(\rho c^2 + P\right)\frac{\partial \ln|g_{00}|^{1/2}}{\partial x^{\nu}} \tag{219}$$

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

$$\boldsymbol{\nabla} P + \rho\boldsymbol{\nabla}\Phi = 0 \tag{220}$$

The difference for a static problem is the replacement of ρ by $\rho + P/c^2$ for the inertial mass density, and the appearance of $\ln|g_{00}|^{1/2}$ as the relativistic generalisation of the potential (note the agreement in the Newtonian limit).

If $P = P(\rho)$ (a function of energy density only), then with $P' \equiv dP/d\rho$, equation (219) may be formally integrated:

$$\int \frac{P'(\rho)\,d\rho}{P(\rho) + \rho c^2} + \ln|g_{00}|^{1/2} = \text{constant}. \tag{221}$$

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 $\gamma \geq 1$.

Exercise. Show that equation (215) may be quite generally written

$$0 = \frac{\partial P}{\partial x^{\nu}} + \frac{1}{|g|^{1/2}}\frac{\partial}{\partial x^{\mu}}\left[|g|^{1/2}(\rho + P/c^2)U^{\mu}U_{\nu}\right] + (\rho + P/c^2)U_{\mu}\frac{\partial U^{\mu}}{\partial x^{\nu}} \tag{222}$$

Show that this is consistent with equation (219). If you do this exercise now, you will save yourself some effort on a Problem Set!

4.7 Covariant differentiation and parallel transport

Recall the geodesic equation,

$$\frac{d^2x^\lambda}{d\tau^2} + \Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0. \tag{223}$$

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

$$\frac{dV^\lambda}{d\tau} + \Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{d\tau}V^\nu = 0, \tag{224}$$

which, following the work of this chapter, we recognise as a covariant formulation of the statement that the vector V^λ is unchanged along a geodesic path. We now formally designate the upgrade of $d/d\tau$ with its own notation:

$$\frac{dV^\alpha}{d\tau} \to \frac{dV^\lambda}{d\tau} + \Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{d\tau}V^\nu \equiv \frac{DV^\lambda}{D\tau}. \tag{225}$$

$DV^\lambda/D\tau$ is by construction a true vector. For those who wish an explicit demonstration, the transformation

$$\frac{DV'^\lambda}{D\tau} = \frac{\partial x'^\lambda}{\partial x^\mu}\frac{DV^\mu}{D\tau} \tag{226}$$

may of course be verified directly. The inhomogeneous contributions from the Γ transformation and the derivatives of the derivatives of the coordinate transformation coefficients cancel in a manner exactly analogous to our original covariant partial derviative calculation.

Exactly the same reasoning is used to define the covariant derivative for a covariant vector, which follows from (172):

$$\frac{dV_\lambda}{d\tau} - \Gamma^\nu_{\mu\lambda}\frac{dx^\mu}{d\tau}V_\nu \equiv \frac{DV_\lambda}{D\tau}. \tag{227}$$

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

$$\frac{dT^\sigma_\lambda}{d\tau} + \Gamma^\sigma_{\mu\nu}\frac{dx^\nu}{d\tau}T^\mu_\lambda - \Gamma^\mu_{\lambda\nu}\frac{dx^\nu}{d\tau}T^\sigma_\mu \equiv \frac{DT^\sigma_\lambda}{D\tau}. \tag{228}$$

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 necessary.

We have earlier seen that the covariant derivative $g_{\mu\nu;\lambda} = 0$. The same is true for the derivative of the inverse $g^{\mu\nu}{}_{;\lambda} = 0$, a fact most easily proven by using $(g^{\mu\rho}g_{\rho\nu})_{;\lambda} = 0$ and applying the chain rule for covariant derivatives. It then follows that $D/D\tau$ must vanish when applied either to the metric tensor $g_{\mu\nu}$ or to its inverse $g^{\mu\nu}$. In particular, the geodesic equation of motion for $V_\mu = dx_\mu/d\tau$ is simply

$$\frac{DV_\mu}{D\tau} = \frac{dV_\mu}{d\tau} - \Gamma^\rho_{\mu\nu}V^\rho V_\nu = 0, \tag{229}$$

since this is simply a matter of lowering the index on the standard geodesic equation $DV^\mu/D\tau = 0$. The geodesic equation of the covariant 4-velocity then becomes

$$\frac{dV_\mu}{d\tau} = \frac{1}{2}(\partial_\mu g_{\rho\nu})V^\rho V^\nu, \tag{230}$$

the demonstration of which we shall leave to the reader. Note that this leads to the important conclusion that if $g_{\rho\nu}$ is independent of x^μ, then $V_\mu = dx_\mu/d\tau$ is conserved along a geodesic orbit.

When a vector or tensor is evaluated along a geodesic path in spacetime, if the quantity in question remains unchanged in a locally inertial reference frame ($d/d\tau = 0$), this condition becomes, in arbitrary coordinates, $D/D\tau = 0$. We are simply expressing the same physical result in a covariant language. (Once again, this works because the two statements agree in the inertial coordinate frame, and this zero difference remains zero in any coordinate frame.) The condition $D/D\tau = 0$ is more generally known as *parallel transport*. For example, a vector may always point along the y axis as we move it around in the xy plane, and is in this sense "constant," but its r and θ components *will* have to change in order to keep this true! The prescription for how those components must generally change, when in local inertial coordinates they remain fixed, is the content of the parallel transport equation.

Suppose that we execute a complete round trip following the rule of parallel transport $D/D\tau = 0$, 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

Let none ignorant of geometry enter my door.

— Inscription over the entrance to Plato's Academy

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.

$$V^{\mu}_{\ ;\sigma} = \frac{\partial V^{\mu}}{\partial x^{\sigma}} + \Gamma^{\mu}_{\nu\sigma} V^{\nu} \equiv T^{\mu}_{\ \sigma} \tag{231}$$

Then

$$T^{\mu}_{\ \sigma;\tau} = \frac{\partial T^{\mu}_{\ \sigma}}{\partial x^{\tau}} + \Gamma^{\mu}_{\nu\tau} T^{\nu}_{\ \sigma} - \Gamma^{\nu}_{\sigma\tau} T^{\mu}_{\ \nu}, \tag{232}$$

or

$$T^{\mu}_{\ \sigma;\tau} = \frac{\partial^2 V^{\mu}}{\partial x^{\tau} \partial x^{\sigma}} + \frac{\partial}{\partial x^{\tau}} \left(\Gamma^{\mu}_{\lambda\sigma} V^{\lambda} \right) + \Gamma^{\mu}_{\nu\tau} \left(\frac{\partial V^{\nu}}{\partial x^{\sigma}} + \Gamma^{\nu}_{\lambda\sigma} V^{\lambda} \right) - \Gamma^{\nu}_{\sigma\tau} \left(\frac{\partial V^{\mu}}{\partial x^{\nu}} + \Gamma^{\mu}_{\lambda\nu} V^{\lambda} \right) \tag{233}$$

The first term and the last group (proportional to $\Gamma^{\nu}_{\sigma\tau}$) are manifestly symmetric in σ and τ, 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 any partial derivative of V^{μ}. The residual terms from taking the covariant derivative commutator are then

$$T^{\mu}_{\ \sigma;\tau} - T^{\mu}_{\ \tau;\sigma} = \left[\frac{\partial \Gamma^{\mu}_{\lambda\sigma}}{\partial x^{\tau}} - \frac{\partial \Gamma^{\mu}_{\lambda\tau}}{\partial x^{\sigma}} + \Gamma^{\mu}_{\nu\tau} \Gamma^{\nu}_{\lambda\sigma} - \Gamma^{\mu}_{\nu\sigma} \Gamma^{\nu}_{\lambda\tau} \right] V^{\lambda}, \tag{234}$$

which we may write as

$$V^{\mu}_{\ ;\sigma;\tau} - V^{\mu}_{\ ;\tau;\sigma} = T^{\mu}_{\ \sigma;\tau} - T^{\mu}_{\ \tau;\sigma} = R^{\mu}_{\ \lambda\sigma\tau} V^{\lambda} \tag{235}$$

Now, both sides of this equation must be tensors, and V^{λ} is itself some arbitrary vector. (Good thing those partial derivatives of V^{λ} 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

$$R^{\mu}_{\ \lambda\sigma\tau} = \frac{\partial \Gamma^{\mu}_{\lambda\sigma}}{\partial x^{\tau}} - \frac{\partial \Gamma^{\mu}_{\lambda\tau}}{\partial x^{\sigma}} + \Gamma^{\mu}_{\nu\tau} \Gamma^{\nu}_{\lambda\sigma} - \Gamma^{\mu}_{\nu\sigma} \Gamma^{\nu}_{\lambda\tau} \tag{236}$$

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_λ whose covariant derivative along a curve $x(\tau)$ vanishes. Then,

$$\frac{dV_\lambda}{d\tau} = \Gamma^{\mu}_{\lambda\nu}\frac{dx^\nu}{d\tau}V_\mu \tag{237}$$

Consider next a tiny round trip journey over a closed path on which V_λ is changing by the above prescription. If we remain in the neighbourhood of some point X^ρ, with x^ρ passing through X^ρ at some instant τ_0, $x^\rho(\tau_0) = X^\rho$, we Taylor expand as follows:

$$\Gamma^{\mu}_{\lambda\nu}(x) = \Gamma^{\mu}_{\lambda\nu}(X) + (x^\rho - X^\rho)\frac{\partial\Gamma^{\mu}_{\lambda\nu}}{\partial X^\rho} + ... \tag{238}$$

$$V_\mu[x(\tau)] = V_\mu(X) + dV_\mu + ... = V_\mu(X) + (x^\rho - X^\rho)\Gamma^{\sigma}_{\mu\rho}(X)V_\sigma(X) + ... \tag{239}$$

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

$$\Gamma^{\mu}_{\lambda\nu}(x)V_\mu(x) = \Gamma^{\mu}_{\lambda\nu}V_\mu + (x^\rho - X^\rho)V_\sigma\left(\frac{\partial\Gamma^{\sigma}_{\lambda\nu}}{\partial X^\rho} + \Gamma^{\sigma}_{\mu\rho}\Gamma^{\mu}_{\lambda\nu}\right) + ... \tag{240}$$

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

$$dV_\lambda = \Gamma^{\mu}_{\lambda\nu}(x)V_\mu(x)\,dx^\nu \tag{241}$$

around a tiny closed path $\oint$, and using (241) and (240), we find that there is a change in the starting value ΔV_λ arising from the term linear in x^ρ given by

$$\Delta V_\lambda = \left(\frac{\partial\Gamma^{\sigma}_{\lambda\nu}}{\partial X^\rho} + \Gamma^{\sigma}_{\mu\rho}\Gamma^{\mu}_{\lambda\nu}\right)V_\sigma \oint x^\rho\,dx^\nu \tag{242}$$

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 ρ and ν. (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 (ρ, ν).

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

$$\Delta V_\lambda = \frac{1}{2}R^\sigma{}_{\lambda\nu\rho}V_\sigma \oint x^\rho\, dx^\nu \tag{243}$$

where

$$R^\sigma{}_{\lambda\nu\rho} = \left(\frac{\partial\Gamma^\sigma_{\lambda\nu}}{\partial X^\rho} - \frac{\partial\Gamma^\sigma_{\lambda\rho}}{\partial X^\nu} + \Gamma^\sigma_{\mu\rho}\Gamma^\mu_{\lambda\nu} - \Gamma^\sigma_{\mu\nu}\Gamma^\mu_{\lambda\rho}\right) \tag{244}$$

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 $dx^\rho dx^\nu$:

$$\Delta V_\lambda = \frac{1}{2}R^\sigma{}_{\lambda\nu\rho}V_\sigma dx^\rho\, dx^\nu. \tag{245}$$

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° meridian through Greenwich till you hit the north pole. Now, once again parallel to itself, move the pencil down the 90°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[7], 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 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!

[7] "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.*

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 orbit about the Earth, don't seem very exciting any more to 21st century physicists.

The curvature tensor also appears when we calculate the difference δx^μ in geodesic spacetime paths executed by two closely-spaced curves, $x^\mu(t)$ and $x'^\mu(t) = x^\mu(t) + \delta x^\mu(t)$. The covariant derivative of δx^μ satisfies (cf. [225]:

$$\frac{D\delta x^\mu}{D\tau} = \frac{d\delta x^\mu}{d\tau} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\delta x^\lambda \tag{246}$$

Applying the $D/D\tau$ operator to $D\delta x^\mu/D\tau$ itself then gives

$$\frac{D^2\delta x^\mu}{D\tau^2} = \frac{d^2\delta x^\mu}{d\tau^2} + \frac{\partial\Gamma^\mu_{\nu\lambda}}{\partial x^\rho}\frac{dx^\rho}{d\tau}\frac{dx^\nu}{d\tau}\delta x^\lambda + ... \tag{247}$$

where the ... notation indicates terms with at least one power of a Γ present. We will work in local inertial coordinates so that these Γ's vanish, but we must be careful to retain the Γ-derivatives, which do not!

To evaluate $d^2\delta x^\mu/d\tau^2$, we use the fact that both $x^\mu(\tau)$ and $x'^\mu(\tau) = x^\mu(\tau) + \delta x^\mu(\tau)$ satisfy the geodesic equation. The difference between these two separate geodesic equations is, to first order in δx^μ,

$$0 = \frac{d^2\delta x^\mu}{d\tau^2} + \frac{\partial\Gamma^\mu_{\nu\lambda}}{\partial x^\rho}\delta x^\rho\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} + ... \tag{248}$$

where once again ... denotes the additional terms that are proportional to Γ factors and therefore vanish in inertial coordinates. If we now swap the ρ and λ indices in (248) and combine this with (247), we obtain

$$\frac{D^2\delta x^\mu}{D\tau^2} = \left(\frac{\partial\Gamma^\mu_{\nu\lambda}}{\partial x^\rho} - \frac{\partial\Gamma^\mu_{\nu\rho}}{\partial x^\lambda}\right)\frac{dx^\rho}{d\tau}\frac{dx^\nu}{d\tau}\delta x^\lambda + ... = R^\mu{}_{\nu\lambda\rho}\frac{dx^\rho}{d\tau}\frac{dx^\nu}{d\tau}\delta x^\lambda \tag{249}$$

The final equality, which is exact, follows because the surviving Γ derivatives combine to form precisely the curvature tensor in these inertial coodinates. But this is a vector equation, and so this must be true in any other coordinate system as well! The presence of curvature implies that nearby geodesics will deviate from another, or conversely, that the existence of such deviations mean that true spacetime curvature is present. The principle of equivalence tells us that while we have no way of measuring the gravitational field *locally* as we are all falling along our common geodesic, the very small, higher order, geodesic deviations can in fact reveal an ambient gravitational field. In the Newtonian limit, these are just the familiar tidal forces! We will see this more explicitly in the next section, but only if you work the *Exercise.*

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 Γ, and quadratic products of Γ. 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 Γ, since there is no self-contracted Γ 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:

$$R^{a}{}_{bcd} = \frac{\partial \Gamma^{a}_{bc}}{\partial x^{d}} - \frac{\partial \Gamma^{a}_{bd}}{\partial x^{c}} + \Gamma^{*}_{bc}\Gamma^{a}_{d*} - \Gamma^{*}_{bd}\Gamma^{a}_{c*} \tag{250}$$

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 Γs, with no Γ 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

$$R_{\lambda\mu\nu\kappa} = g_{\lambda\sigma}R^{\sigma}{}_{\mu\nu\kappa} \tag{251}$$

or

$$R_{\lambda\mu\nu\kappa} = g_{\lambda\sigma}\left[\frac{\partial \Gamma^{\sigma}_{\mu\nu}}{\partial x^{\kappa}} - \frac{\partial \Gamma^{\sigma}_{\mu\kappa}}{\partial x^{\nu}} + \Gamma^{\eta}_{\mu\nu}\Gamma^{\sigma}_{\kappa\eta} - \Gamma^{\eta}_{\mu\kappa}\Gamma^{\sigma}_{\nu\eta}\right] \tag{252}$$

Remembering the definition of the affine connection (121), the right side of (252) is

$$\frac{g_{\lambda\sigma}}{2}\frac{\partial}{\partial x^{\kappa}}\left[g^{\sigma\rho}\left(\frac{\partial g_{\rho\mu}}{\partial x^{\nu}} + \frac{\partial g_{\rho\nu}}{\partial x^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\rho}}\right)\right] - \frac{g_{\lambda\sigma}}{2}\frac{\partial}{\partial x^{\nu}}\left[g^{\sigma\rho}\left(\frac{\partial g_{\rho\mu}}{\partial x^{\kappa}} + \frac{\partial g_{\rho\kappa}}{\partial x^{\mu}} - \frac{\partial g_{\mu\kappa}}{\partial x^{\rho}}\right)\right]$$
$$+ g_{\lambda\sigma}\left(\Gamma^{\eta}_{\mu\nu}\Gamma^{\sigma}_{\kappa\eta} - \Gamma^{\eta}_{\mu\kappa}\Gamma^{\sigma}_{\nu\eta}\right) \tag{253}$$

The x^{κ} and x^{ν} 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 $g_{\lambda\sigma}g^{\sigma\rho} = \delta^{\rho}_{\lambda}$, the terms that are linear in the second order g-derivatives are

$$\frac{1}{2}\left(\frac{\partial^2 g_{\lambda\nu}}{\partial x^{\kappa}\partial x^{\mu}} - \frac{\partial^2 g_{\mu\nu}}{\partial x^{\kappa}\partial x^{\lambda}} - \frac{\partial^2 g_{\lambda\kappa}}{\partial x^{\nu}\partial x^{\mu}} + \frac{\partial^2 g_{\mu\kappa}}{\partial x^{\nu}\partial x^{\lambda}}\right) \tag{254}$$

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^κ and x^ν 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:

$$g_{\lambda\sigma}\frac{\partial g^{\sigma\rho}}{\partial x^\kappa} = -g^{\sigma\rho}\frac{\partial g_{\lambda\sigma}}{\partial x^\kappa} = -g^{\sigma\rho}\left(\Gamma^\eta_{\kappa\lambda}g_{\eta\sigma} + \Gamma^\eta_{\kappa\sigma}g_{\eta\lambda}\right) \tag{255}$$

where in the final equality, equation (176) 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 (253) are now of the form $g\Gamma\Gamma$:

$$-\left(\Gamma^\eta_{\kappa\lambda}g_{\eta\sigma} + \cancel{\Gamma^\eta_{\kappa\sigma}g_{\eta\lambda}}\right)\Gamma^\sigma_{\mu\nu} + \left(\Gamma^\eta_{\nu\lambda}g_{\eta\sigma} + \cancel{\Gamma^\eta_{\nu\sigma}g_{\eta\lambda}}\right)\Gamma^\sigma_{\mu\kappa} + g_{\lambda\sigma}\left(\cancel{\Gamma^\eta_{\mu\nu}\Gamma^\sigma_{\kappa\eta}} - \cancel{\Gamma^\eta_{\mu\kappa}\Gamma^\sigma_{\nu\eta}}\right), \tag{256}$$

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:

$$g_{\eta\sigma}\left(\Gamma^\eta_{\nu\lambda}\Gamma^\sigma_{\mu\kappa} - \Gamma^\eta_{\kappa\lambda}\Gamma^\sigma_{\mu\nu}\right) \tag{257}$$

Adding together the terms in (254) and (257), we arrive at

$$R_{\lambda\mu\nu\kappa} = \frac{1}{2}\left(\frac{\partial^2 g_{\lambda\nu}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\kappa \partial x^\lambda} - \frac{\partial^2 g_{\lambda\kappa}}{\partial x^\nu \partial x^\mu} + \frac{\partial^2 g_{\mu\kappa}}{\partial x^\nu \partial x^\lambda}\right) + g_{\eta\sigma}\left(\Gamma^\eta_{\nu\lambda}\Gamma^\sigma_{\mu\kappa} - \Gamma^\eta_{\kappa\lambda}\Gamma^\sigma_{\mu\nu}\right) \tag{258}$$

Exercise. What is $R_{\lambda\mu\nu\kappa}$ in local inertial coordinates? Why is this the same form as emerges in the Newtonian limit? Show that for a static potential Φ, the equation of geodesic deviation (249) for a Cartesian spatial component in the Newtonian limit reduces to

$$\frac{d^2\delta x^i}{dt^2} = -\frac{\partial^2\phi}{\partial x^i \partial x^j}\delta x^j$$

(Hint: Note that you may raise indicies with $\eta^{\alpha\beta}$ [why?], and that only the $R_{\lambda 0\nu 0}$ term survives from $R_{\lambda\mu\nu\kappa}$ [why?].) Give a physical interpretation of this equation as a tidal force manifestation.

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 Γ vanish. These results may then be verified just from the terms linear in the g derivatives in (258):

$$R_{\lambda\mu\nu\kappa} = R_{\nu\kappa\lambda\mu} \qquad \text{(symmetry)} \tag{259}$$

$$R_{\lambda\mu\nu\kappa} = -R_{\mu\lambda\nu\kappa} = -R_{\lambda\mu\kappa\nu} = R_{\mu\lambda\kappa\nu} \qquad \text{(antisymmetry)} \tag{260}$$

$$R_{\lambda\mu\nu\kappa} + R_{\lambda\kappa\mu\nu} + R_{\lambda\nu\kappa\mu} = 0 \qquad \text{(cyclic)} \tag{261}$$

5.3.3 How many independent components does the curvature have?

Note: This section is off-syllabus and non-examinable.

To count the number of independent components of $R_{\lambda\mu\nu\kappa}$ in n dimensions, start with the simple observation that an antisymmetric tensor $A_{\mu\nu}$ has $n(n-1)/2$ independent components in n dimensions: there are n choices for μ, only $n-1$ choices for ν (since the tensor vanishes if $\mu = \nu$), and we must divide by 2 since $A_{\mu\nu} = -A_{\nu\mu}$. There are $n(n+1)/2$ independent components of a symmetric tensor, say $S_{\mu\nu}$, since we simply add back in the n diagonal elements which are now present, but that we neglected in the component counting of $A_{\mu\nu}$. Next, we mentally group the indicies of the covariant form of the curvature tensor as $R_{(\lambda\mu)(\nu\kappa)}$. By equation (260), there are then $n(n-1)/2$ combinations allowed each for $\lambda\mu$ or $\kappa\nu$. With $(\lambda\mu)$ or $(\kappa\nu)$ each viewed as a sort of mega-index unto itself, the symmetry of equation (259) now tells us there are

$$\frac{1}{2}\left[\frac{n(n-1)}{2}\right]\left[\frac{n(n-1)}{2}+1\right] \equiv N \tag{262}$$

ways to list the components of $R_{(\lambda\mu)(\nu\kappa)}$.

Wait. We are not yet finished, because we have not yet taken into account equation (261), the cyclic sum. How many constraints does this represent? In $n = 4$ dimensions it is only one constraint. To see this, set $\lambda\mu\nu\kappa = 1234$. If we start with any other combination, the constraint equation is either identical to the first, or just multiplied by an overall factor of -1. In $n > 4$ dimensions, we therefore have exactly one equation for each unique combination of $\lambda\mu\nu\kappa$ with no index repeats. This is just the number of combinations of 4 items, taken from a larger set of n items with distinct members. This is

$$\frac{n!}{(n-4)!4!} = \frac{1}{24}n(n-1)(n-2)(n-3). \tag{263}$$

The total number of independent components of $R_{\lambda\mu\nu\kappa}$ is therefore

$$N - \frac{1}{24}n(n-1)(n-2)(n-3) = \frac{1}{12}n^2(n^2-1). \tag{264}$$

In one dimension there are zero components: R_{1111} vanishes by antisymmetry. Is a curved line really not curved? Indeed it is not, because the casual use of the word "curved" is not the same as the mathematical sense of a distortion from Euclidian (or Minkowski!) spatial structure. You can use a piece of string to measure distances accurately along flat or "curved" surfaces equally well. There is simply no escape from a Euclidian metric in one dimension, it is all just a change of variable: the metric $f(x)dx \equiv dX$, where $X = \int f(x)\,dx$.

In two dimensions, there is one component of the curvature tensor, R_{1212}. This was first identified by Gauss in his study of the mathematics of curved surfaces.

In three dimensions, there are 6 components, but we shall later see that this is not enough for a theory of gravity! It is only with at least four dimensions, with its 20 different components of the curvature tensor, that we have enough freedom to support a gravitational field from an external source (a tidal force actually) in empty space.

5.4 The Ricci tensor

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

$$R_{\mu\kappa} = g^{\lambda\nu} R_{\lambda\mu\nu\kappa} = g^{\lambda\nu} R_{\nu\kappa\lambda\mu} \text{ (by symmetry) } = g^{\nu\lambda} R_{\nu\kappa\lambda\mu} = R_{\kappa\mu} \tag{265}$$

so that the Ricci tensor is symmetric. From equation (236) we have explicitly:

$$R_{\mu\kappa} = \frac{\partial \Gamma^\lambda_{\mu\lambda}}{\partial x^\kappa} - \frac{\partial \Gamma^\lambda_{\mu\kappa}}{\partial x^\lambda} + \Gamma^\eta_{\mu\lambda}\Gamma^\lambda_{\kappa\eta} - \Gamma^\eta_{\mu\kappa}\Gamma^\lambda_{\lambda\eta}. \tag{266}$$

The Ricci tensor is of fundamental importance 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, the advance of a planet's orbital perihelion, or the spacetime geometry of a black hole. 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 $T_{\mu\nu}$. (That 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. If it did, there would be no gravitational field in empty space.

Exercise. The expression given with the "Really Useful Numbers" on page (7) of these notes has a somewhat different form of the Ricci tensor to the one above. With $g = |\det g_{\mu\nu}|$, this is

$$R_{\mu\kappa} = \frac{1}{2}\frac{\partial^2 \ln g}{\partial x^\kappa \partial x^\mu} - \frac{1}{\sqrt{g}}\frac{\partial \sqrt{g}\,\Gamma^\lambda_{\mu\kappa}}{\partial x^\lambda} + \Gamma^\eta_{\mu\lambda}\Gamma^\lambda_{\kappa\eta}.$$

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 also form the Ricci (or curvature) scalar

$$R \equiv R^\mu{}_\mu, \tag{267}$$

another fundamentally 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.

Finally, it will be of interest later, in the course of a variational calculation, to have at hand a formula for $\delta R_{\mu\kappa}$, the change in the Ricci tensor due to a change $\delta g_{\mu\nu}$ in the metric tensor. This is given by:

$$\delta R_{\mu\kappa} = (\delta\Gamma^\lambda_{\mu\lambda})_{;\kappa} - (\delta\Gamma^\lambda_{\mu\kappa})_{;\lambda} \tag{268}$$

where the $\delta\Gamma$ are the corresponding changes in the affine connections. The easiest way to see this is to note that in local inertial coordinates (with vanishing affine connection), we need retain only the first two partial derivative terms on the right side of equation (266). The result for general coordinates then follows by simply turning the partial derivatives into covariant derivatives, since both $\delta R_{\mu\kappa}$ and $\delta\Gamma^{\lambda}_{\mu\nu}$ are true tensors.

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:

$$\nabla^2\Phi = 4\pi G\rho \tag{269}$$

where Φ is the usual gravitational potential, and ρ the mass density. Second derivatives of functions are generally associated with curvature. 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 ∇^2 operator directly to the curvature of an ordinary 2D surface:

$$\frac{\partial^2\zeta}{\partial x^2} + \frac{\partial^2\zeta}{\partial y^2} = \frac{1}{R_1} + \frac{1}{R_2} \tag{270}$$

where ζ 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):

$$\gamma\left(\frac{\partial^2\zeta}{\partial x^2} + \frac{\partial^2\zeta}{\partial y^2}\right) = \Delta P, \tag{271}$$

where γ 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 Φ as the displacement of a surface membrane and ρ as the applied pressure difference. In this analogy, that makes $1/(4\pi G)$ equal to 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 **div(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,

$$R_{\lambda\mu\nu\kappa;\eta} = \frac{1}{2}\frac{\partial}{\partial x^\eta}\left(\frac{\partial^2 g_{\lambda\nu}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\kappa \partial x^\lambda} - \frac{\partial^2 g_{\lambda\kappa}}{\partial x^\mu \partial x^\nu} + \frac{\partial^2 g_{\mu\kappa}}{\partial x^\nu \partial x^\lambda}\right) \tag{272}$$

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

$$R_{\lambda\mu\nu\kappa;\eta} + R_{\lambda\mu\eta\nu;\kappa} + R_{\lambda\mu\kappa\eta;\nu} = 0 \tag{273}$$

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 ∂x^c indices downstairs are unambiguous, since as coordinate derivatives their order is immaterial. The first term in (273) 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 (273) into the form of a *single* vanishing covariant tensor divergence, for reasons that will soon become very clear. Toward this goal, contract λ with ν, remembering the symmetries in (260). (E.g., in the second term on the left side of [273], swap ν and η before contracting, changing the sign.) We find

$$R_{\mu\kappa;\eta} - R_{\mu\eta;\kappa} + R^{\nu}{}_{\mu\kappa\eta;\nu} = 0. \tag{274}$$

Next, contract μ with κ:

$$R_{;\eta} - R^{\mu}{}_{\eta;\mu} - R^{\nu}{}_{\eta;\nu} = 0 \tag{275}$$

(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 μ and contract with κ:

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

Cleaning things up, our contracted identity (275) becomes:

$$(\delta^{\mu}_{\eta} R - 2R^{\mu}{}_{\eta})_{;\mu} = 0. \tag{276}$$

Raising η (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:

$$\left(R^{\mu\nu} - g^{\mu\nu} \frac{R}{2} \right)_{;\mu} = 0 \tag{277}$$

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 (277) 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 from the source vector: why does the displacement current need to be added to the equation $\boldsymbol{\nabla} \times \boldsymbol{B} = \mu_0 \boldsymbol{J}$? 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," $(1/c^2)\partial \boldsymbol{E}/\partial t$, adding it to the right side of the equation, because $\boldsymbol{\nabla \cdot J}$ 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.

5.7 The Einstein field equations

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, ρ, is 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

$$G_{\mu\nu} = CT_{\mu\nu} \tag{278}$$

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 derivative of $g_{00} \simeq -1 - 2\Phi/c^2$ enters directly into the dynamics; ii) in the same limit, we need to recover the Poisson equation, which involves a Laplacian of g_{00}; and iii) we are assuming that our theory of gravity does not change its behaviour with length scale: it has no characteristic length associated with it where the field properties change fundamentally. In other words, 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 $T^{\mu\nu}_{\ \ ;\mu} = 0$. 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

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

(As always, there is no difficulty shifting indices up or down as considerations demand: the index shifters $g_{\mu\nu}$ and $g^{\mu\nu}$ all have vanishing covariant derivatives.) We have therefore

determined the field equations of gravity up to an overall normalisation:

$$R_{\mu\nu} - \frac{g_{\mu\nu}R}{2} = CT_{\mu\nu} \tag{279}$$

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 Φ 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 (279) reads (raise μ, contract with ν):

$$R - \frac{4 \times R}{2} = -R = CT. \tag{280}$$

Substituting this for R back in the original equation leads to

$$R_{\mu\nu} = C\left(T_{\mu\nu} - \frac{g_{\mu\nu}T}{2}\right) \equiv CS_{\mu\nu} \tag{281}$$

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

$$S_{\mu\nu} = T_{\mu\nu} - g_{\mu\nu}T/2. \tag{282}$$

The 00 component of (281) is

$$R_{00} = C\left(T_{00} - \frac{g_{00}T}{2}\right) \tag{283}$$

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 $g_{\mu\nu}$.

$$R_{00} = C\left(T_{00} - \frac{\eta_{00}T^0_{\ 0}}{2}\right) = C\left(T_{00} - \frac{T_{00}}{2}\right) = C\frac{T_{00}}{2} = C\frac{\rho c^2}{2}, \tag{284}$$

where ρ is the Newtonian mass density. Calculating R_{00} explicitly,

$$R_{00} = R^{\nu}_{\ 0\nu 0} = \eta^{\lambda\nu}R_{\lambda 0\nu 0} \tag{285}$$

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

$$R_{\lambda\mu\nu\kappa} = \frac{1}{2}\left(\frac{\partial^2 g_{\lambda\nu}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\kappa \partial x^\lambda} - \frac{\partial^2 g_{\lambda\kappa}}{\partial x^\nu \partial x^\mu} + \frac{\partial^2 g_{\mu\kappa}}{\partial x^\nu \partial x^\lambda}\right), \tag{286}$$

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

$$R_{\lambda 0\nu 0} = \frac{1}{2}\frac{\partial^2 g_{00}}{\partial x^\nu \partial x^\lambda}. \tag{287}$$

Finally,

$$R_{00} = \eta^{\lambda\nu} R_{\lambda 0 \nu 0} = \frac{1}{2}\eta^{\lambda\nu}\frac{\partial^2 g_{00}}{\partial x^\lambda \partial x^\nu} = \frac{1}{2}\nabla^2 g_{00} = -\frac{1}{c^2}\nabla^2\Phi = \frac{C\rho c^2}{2} \tag{288}$$

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*:

$$\boxed{G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -\frac{8\pi G}{c^4}T_{\mu\nu}} \tag{289}$$

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 μ and ν,

$$R = \frac{8\pi G}{c^4}T, \tag{290}$$

and the field equations become

$$R_{\mu\nu} = -\frac{8\pi G}{c^4}\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right) \equiv -\frac{8\pi G}{c^4}S_{\mu\nu} \tag{291}$$

where as before,

$$S_{\mu\nu} = T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T, \tag{292}$$

a "Bianchified form" of the stress tensor. In vacuo, the field equations reduce to the analogue of the Laplace Equation:

$$R_{\mu\nu} = 0. \tag{293}$$

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 Λ 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 Λ term to the right side of the field equations:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -\frac{8\pi G}{c^4}T_{\mu\nu} + \Lambda g_{\mu\nu}, \tag{294}$$

and dubbed Λ 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.[8] With the historical discovery of an expanding univese, Einstein retracted the Λ term, calling it "the biggest mistake of my life."

[8]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 Λ 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. Λ is otherwise completely negligible. The so-called biggest mistake of Einstein's life was, therefore, quadratic in amplitude: one factor of error for introducing Λ for the wrong reason, the second factor for retracting Λ for the wrong reason! Except for cosmological problems, we will always assume $\Lambda = 0$.

Always remember that because they have been constructed to admit a new solution by a transformation of coordinates, the field equations will generally admit more than one solution, unless there are additional constraints. 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.[9] We have the usual time t coordinate for index 0, and let us say α, β, γ 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, ϕ, z), or spherical (r, θ, ϕ) 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 favoured set of 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. In numerical relativity, for example in determining the metric tensor for merging black holes, one often has to solve for the coordinates as part of the problem itself.

5.8 An alternative formulation of the field equations

Note: This section is off-syllabus and non-examinable.

Let us write the general metric tensor $g_{\mu\nu}$ in the form

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{295}$$

where $\eta_{\mu\nu}$ is the Minkowski metric. This is a useful and obvious decomposition when $h_{\mu\nu}$ is small compared with unity, as would be the case for either the Newtonian or weak radiation limits, but here we shall not restrict ourselves to this assumption. We shall assume only that $h_{\mu\nu} \to 0$ at large distances from the sources under consideration. Nevertheless, the linear (in $h_{\mu\nu}$) components of $G_{\mu\nu}$ are of interest in their own right.

The metric tensor $g_{\mu\nu}$ behaves, of course, as a tensor under any proper coordinate transformation. While $h_{\mu\nu}$ is not itself a true tensor, beause $\eta_{\mu\nu}$ transforms identically under Lorentz transformations, $h_{\mu\nu}$ must be Lorentz covariant. This implies that spacetime partial derivatives of $h_{\mu\nu}$ are Lorentz tensors as well. The indices of $h_{\mu\nu}$ and its derivatives are raised and lowered by $\eta_{\mu\nu}$, preserving Lorentz covariance in all their forms.

We now expand the left side of the field equation in powers of $h_{\mu\nu}$, retaining the linear terms, whilst placing quadratic and higher order terms on the right side of the equation.

[9] Don't smirk. If we're using awkward coordinates, it can be *very* hard to tell. You'll see.

Formally, we then have

$$R^{(1)}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R^{(1)} = -\frac{8\pi G T_{\mu\nu}}{c^4} - \sum_{n=2}^{\infty}\left(R^{(n)}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R^{(n)}\right), \tag{296}$$

where the superscript (n) notation denotes nth powers of $h_{\mu\nu}$ grouped together. The field equations then take the form

$$R^{(1)}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R^{(1)} = -\frac{8\pi G \tau_{\mu\nu}}{c^4}, \tag{297}$$

with

$$\tau_{\mu\nu} = T_{\mu\nu} + \frac{c^4}{8\pi G}\sum_{n=2}^{\infty}\left(R^{(n)}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R^{(n)}\right) \equiv T_{\mu\nu} + t_{\mu\nu}. \tag{298}$$

Equations (297) and (298) may appear to be just a formal reworking of the field equations, with all of the nonlinear complexity hidden in the artificial construct $t_{\mu\nu}$. But $t_{\mu\nu}$ has physical significance. First, as the difference between two Lorentz covariant quantities, it must be Lorentz covariant. (This holds for $\tau_{\mu\nu}$ as well, of course.) Second, $t_{\mu\nu}$ captures all nonlinear terms of the field equations that are quadratic and higher order in the derivatives of the metric tensor, and appears as a simple additive partner to $T_{\mu\nu}$. We may therefore identify $t_{\mu\nu}$ as the gravitational stress-energy "pseudo tensor" generated by the gravitational field *itself*. This means, in turn, that the pseudo tensor $\tau_{\mu\nu}$ represents the total stress-energy from both material and gravitational sources. Note that indices on both $t_{\mu\nu}$ and $\tau_{\mu\nu}$ are governed by $\eta_{\mu\nu}$, whereas only $g_{\mu\nu}$ can perform this role for the true tensor $T_{\mu\nu}$.

Following the linearised Bianchi identities, the ordinary divergence of $\tau_{\mu\nu}$ vanishes exactly:

$$\partial^\nu \tau_{\mu\nu} = \partial^\nu(R^{(1)}_{\mu\nu} - \eta_{\mu\nu}R^{(1)}/2) = 0. \tag{299}$$

Since the divergence of $\tau^{0\nu}$ vanishes within any spatial volume V, we have

$$\frac{d}{dt}\int_V \tau^{0\nu}\, d^3x = -\int_V \partial_i \tau^{i\nu}\, d^3x = -\int_S \tau^{i\nu} n_i dS \to -\int_S t^{i\nu} n_i dS, \tag{300}$$

where we have used the standard form of the divergence theorem in the second equality. The final term on the right relies on the assumption that $T_{\mu\nu}$ vanishes at finite radius (or at least falls off more rapidly than $1/r^2$), whereas a contribution from $t^{\mu\nu}$ may be present in the form of time-dependent gravitational radiation. For static geometries however, the $\tau_{i\nu}$ in the final surface integral on the right will generally fall off sufficiently rapidly (typically $(\partial_\mu h^{\nu\lambda})^2 \sim 1/r^4$), that the $\tau^{0\nu}$ volume integral is time-independent and convergent. When this condition is met, then by analogy with the ordinary stress tensor, the integral

$$P^\nu = \int_V \tau^{0\nu} d^3x \tag{301}$$

may be interpreted as the total energy-momentum pseudo vector of matter, radiation *and gravitation.* In spite of this remarkable inclusiveness (actually, *because* of it), the integral in (301) is taken over ordinary Euclidian space.

One may also define the antisymmetric pseudo tensor

$$J^{\mu\nu} = \int_V (x^\mu \tau^{0\nu} - x^\nu \tau^{0\mu})\, d^3x, \tag{302}$$

whose spatial J^{ij} components now correspond to the combined angular momentum of gravitation, radiation, and matter. Further discussion and applications of P^ν and $J^{\mu\nu}$ to black hole geometries may be found in §11.2.

If we turn our problem around by directly evaluating the divergence of the Ricci tensor in local inertial coordinates (a very simple task), we may construct a one-line proof of the contracted Bianchi Identities, and thereby an alternative derivation of the field equations themselves. The key point is that in local inertial coordinates erected about a particular point X^ρ in spacetime, $h_{\mu\nu}$ is quadratically small in the coordinate difference $x^\rho - X^\rho$. In these local coordinates, *the only Ricci curvature derivatives that appear are linear in* $h_{\mu\nu}$. In particular,

$$R_{\mu\nu} \text{ (local inertial coodinates)} = \mathrm{R}^{(1)}_{\mu\nu}, \tag{303}$$

with equality holding for the divergence as well. Defining $h \equiv h^\mu{}_\mu$, to linear order in $h_{\mu\nu}$ we have,

$$\det g_{\mu\nu} = -1 + h, \qquad \Gamma^\lambda_{\mu\nu} = \frac{1}{2}(\partial_\mu h^\lambda{}_\nu + \partial_\nu h^\lambda{}_\mu - \partial^\lambda h_{\mu\nu}). \tag{304}$$

Using the form of the Ricci tensor from page (7) appropriate to local inertial coordinates, we readily find

$$2R^{\mu\nu} = \partial^\mu\partial^\nu h - \partial^\mu\partial_\lambda h^{\lambda\nu} - \partial^\nu\partial_\lambda h^{\lambda\mu} + \partial_\lambda\partial^\lambda h^{\mu\nu} = 2R^{(1)\mu\nu}, \tag{305}$$

$$R = \partial_\mu\partial^\mu h - \partial_\mu\partial_\nu h^{\mu\nu} = R^{(1)}. \tag{306}$$

Hence,

$$2\partial_\mu R^{\mu\nu} = \partial^\nu(\partial_\lambda\partial^\lambda h - \partial_\mu\partial_\lambda h^{\lambda\mu}) = \eta^{\mu\nu}\partial_\mu(\partial_\lambda\partial^\lambda h - \partial_\lambda\partial_\rho h^{\lambda\rho}) = \eta^{\mu\nu}\partial_\mu R^{(1)}, \tag{307}$$

which leads immediately to the contracted Bianchi Identities,

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

after transforming the vanishing difference tensor to general coordinates. With the Bianchi Identities in hand, we have seen that it is but a very short step from the Newtonian limit to the general field equations. By examining in isolation the linear terms of the field equations in the formulation of equation (297), we may also physically understand the vanishing of the *covariant* divergence of $T_{\mu\nu}$ in the presence of spacetime curvature. This is really an interchange *between* the material stress-energy and gravitational stress-energy components. By contrast, the vanishing of the *ordinary* divergence of $\tau_{\mu\nu}$ expresses a true local conservation of the combined material *plus* gravitational stress-energy.

As you see, the war treated me kindly enough, in spite of the heavy gunfire, to allow me to get away from it all and take this walk in the land of your ideas.

— Karl Schwarzschild writing to Albert Einstein, 1915

6 Schwarzschild black holes

6.1 Schwarzschild metric

The first exact solution to the nonlinear equations of general relativity was found by Karl Schwarzschild in 1915, a month after the appearance of Einstein's epochal paper. This is the relativistic analgoue of the Newtonian "$1/r$" point mass solution for the potential function. The Schwarzschild solution has been central to the development of general relativity. Its first order departure from Newtonian theory gives rise to a readily measurable perhelion advance of Mercury's orbit, while the strong field form is the foundation of our understanding of black holes.

The mathematical problem is to determine the form of the metric tensor $g_{\mu\nu}$ for the spacetime surrounding a nonrotating point mass M by solving the field equations in the Ricci form $R_{\mu\nu} = 0$, subject to the appropriate boundary conditions. We will choose to work with coordinates that look as much as possible like standard spherical coordinates, going over to the standard r, θ, ϕ variables at large distances from the source.

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

$$-c^2 d\tau^2 = -B\, c^2 dt^2 + A\, dr^2 + C\, d\Omega^2 \tag{308}$$

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 everywhere. (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:"

$$-c^2 d\tau^2 = -B(r)\, c^2 dt^2 + A(r)\, dr^2 + r^2\,(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{309}$$

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

$$g_{tt} = -B(r) \qquad g_{rr} = A(r) \qquad g_{\theta\theta} = r^2 \qquad g_{\phi\phi} = r^2 \sin^2\theta \tag{310}$$

and its inverse $g^{\mu\nu}$, the reciprocals of $g_{\mu\nu}$:

$$g^{tt} = -B^{-1}(r) \qquad g^{rr} = A^{-1}(r) \qquad g^{\theta\theta} = r^{-2} \qquad g^{\phi\phi} = r^{-2}(\sin\theta)^{-2} \tag{311}$$

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

$$g = r^4 AB \sin^2\theta \tag{312}$$

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

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

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

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

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

$$\begin{gathered}
\Gamma^t_{tr} = \Gamma^t_{rt} = \frac{B'}{2B} \\
\Gamma^r_{tt} = \frac{B'}{2A} \quad \Gamma^r_{rr} = \frac{A'}{2A} \quad \Gamma^r_{\theta\theta} = -\frac{r}{A} \quad \Gamma^r_{\phi\phi} = -\frac{r\sin^2\theta}{A} \\
\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r} \qquad \Gamma^\theta_{\phi\phi} = -\sin\theta\,\cos\theta \\
\Gamma^\phi_{\phi r} = \Gamma^\phi_{r\phi} = \frac{1}{r} \qquad \Gamma^\phi_{\phi\theta} = \Gamma^\phi_{\theta\phi} = \cot\theta
\end{gathered} \tag{313}$$

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

Next, we need the Ricci Tensor:

$$R_{\mu\kappa} \equiv R^\lambda{}_{\mu\lambda\kappa} = \frac{\partial\Gamma^\lambda_{\mu\lambda}}{\partial x^\kappa} - \frac{\partial\Gamma^\lambda_{\mu\kappa}}{\partial x^\lambda} + \Gamma^\eta_{\mu\lambda}\Gamma^\lambda_{\kappa\eta} - \Gamma^\eta_{\mu\kappa}\Gamma^\lambda_{\lambda\eta} \tag{314}$$

Remembering equation (205), this may be written

$$R_{\mu\kappa} = \frac{1}{2}\frac{\partial^2 \ln g}{\partial x^\kappa \partial x^\mu} - \frac{\partial\Gamma^\lambda_{\mu\kappa}}{\partial x^\lambda} + \Gamma^\eta_{\mu\lambda}\Gamma^\lambda_{\kappa\eta} - \frac{\Gamma^\eta_{\mu\kappa}}{2}\frac{\partial \ln g}{\partial x^\eta} \equiv \frac{\partial^2 \ln\sqrt{g}}{\partial x^\kappa \partial x^\mu} - \frac{1}{\sqrt{g}}\frac{\partial(\sqrt{g}\,\Gamma^\lambda_{\mu\kappa})}{\partial x^\lambda} + \Gamma^\eta_{\mu\lambda}\Gamma^\lambda_{\kappa\eta} \tag{315}$$

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

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

This gives

$$R_{tt} = -\frac{B''}{2A} + \frac{B'}{4A}\left(\frac{B'}{B} + \frac{A'}{A}\right) - \frac{B'}{rA} \tag{316}$$

Next, R_{rr}:

$$R_{rr} = \frac{1}{2}\frac{\partial^2 \ln g}{\partial r^2} - \frac{\partial \Gamma^r_{rr}}{\partial r} + \Gamma^\eta_{r\lambda}\Gamma^\lambda_{r\eta} - \frac{\Gamma^r_{rr}}{2}\frac{\partial \ln g}{\partial r}$$

$$= \frac{1}{2}\frac{\partial}{\partial r}\left(\cancel{\frac{A'}{A}} + \frac{B'}{B} + \frac{4}{r}\right) - \cancel{\frac{\partial}{\partial r}\left(\frac{A'}{2A}\right)} + \Gamma^\eta_{r\lambda}\Gamma^\lambda_{r\eta} - \frac{A'}{4A}\left(\frac{A'}{A} + \frac{B'}{B} + \frac{4}{r}\right)$$

$$= \frac{B''}{2B} - \frac{1}{2}\left(\frac{B'}{B}\right)^2 - \frac{2}{r^2} + \left(\Gamma^t_{rt}\right)^2 + \left(\Gamma^r_{rr}\right)^2 + \left(\Gamma^\theta_{r\theta}\right)^2 + \left(\Gamma^\phi_{r\phi}\right)^2 - \frac{1}{4}\left(\frac{A'}{A}\right)^2 - \frac{A'B'}{4AB} - \frac{A'}{rA}$$

$$= \frac{B''}{2B} - \frac{1}{2}\left(\frac{B'}{B}\right)^2 - \cancel{\frac{2}{r^2}} + \frac{B'^2}{4B^2} + \cancel{\frac{A'^2}{4A^2}} + \cancel{\frac{1}{r^2}} + \cancel{\frac{1}{r^2}} - \frac{1}{4}\cancel{\left(\frac{A'}{A}\right)^2} - \frac{A'B'}{4AB} - \frac{A'}{rA}$$

so that finally we arrive at

$$R_{rr} = \frac{B''}{2B} - \frac{1}{4}\frac{B'}{B}\left(\frac{A'}{A} + \frac{B'}{B}\right) - \frac{A'}{rA} \tag{317}$$

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!

$$R_{\theta\theta} = \frac{\partial \Gamma^\lambda_{\theta\lambda}}{\partial \theta} - \frac{\partial \Gamma^\lambda_{\theta\theta}}{\partial x^\lambda} + \Gamma^\eta_{\theta\lambda}\Gamma^\lambda_{\theta\eta} - \Gamma^\eta_{\theta\theta}\Gamma^\lambda_{\lambda\eta}$$

$$= \frac{1}{2}\frac{\partial^2 \ln g}{\partial \theta^2} - \frac{\partial \Gamma^r_{\theta\theta}}{\partial r} + \Gamma^\eta_{\theta\lambda}\Gamma^\lambda_{\theta\eta} - \Gamma^r_{\theta\theta}\Gamma^\lambda_{\lambda r}$$

$$= \frac{d(\cot\theta)}{d\theta} + \frac{d}{dr}\left(\frac{r}{A}\right) + \Gamma^\eta_{\theta\lambda}\Gamma^\lambda_{\theta\eta} + \frac{r}{2A}\frac{\partial \ln g}{\partial r}$$

$$= -\frac{1}{\sin^2\theta} + \frac{1}{A} - \frac{rA'}{A^2} + \Gamma^r_{\theta\lambda}\Gamma^\lambda_{\theta r} + \Gamma^\theta_{\theta\lambda}\Gamma^\lambda_{\theta\theta} + \Gamma^\phi_{\theta\lambda}\Gamma^\lambda_{\theta\phi} + \frac{r}{2A}\left(\frac{A'}{A} + \frac{B'}{B} + \frac{4}{r}\right)$$

$$= -\frac{1}{\sin^2\theta} + \frac{3}{A} - \frac{rA'}{2A^2} + \Gamma^r_{\theta\theta}\Gamma^\theta_{\theta r} + \Gamma^\theta_{\theta r}\Gamma^r_{\theta\theta} + \left(\Gamma^\phi_{\theta\phi}\right)^2 + \frac{rB'}{2AB}$$

$$= -\frac{1}{\sin^2\theta} + \frac{3}{A} - \frac{rA'}{2A^2} - \frac{2}{A} + \cot^2\theta + \frac{rB'}{2AB}$$

The trigonometric terms add to -1. We finally obtain

$$R_{\theta\theta} = -1 + \frac{1}{A} + \frac{r}{2A}\left(-\frac{A'}{A} + \frac{B'}{B}\right) \tag{318}$$

$R_{\phi\phi}$ is the last nonvanishing Ricci component. No whining now! The first term in (314) vanishes, since nothing in the metric depends on ϕ. Then,

$$R_{\phi\phi} = -\frac{\partial \Gamma^\lambda_{\phi\phi}}{\partial x^\lambda} + \Gamma^\eta_{\phi\lambda}\Gamma^\lambda_{\phi\eta} - \frac{\Gamma^\eta_{\phi\phi}}{2}\frac{\partial \ln|g|}{\partial x^\eta}$$

$$= -\frac{\partial \Gamma^r_{\phi\phi}}{\partial r} - \frac{\partial \Gamma^\theta_{\phi\phi}}{\partial \theta} + \Gamma^r_{\phi\lambda}\Gamma^\lambda_{\phi r} + \Gamma^\theta_{\phi\lambda}\Gamma^\lambda_{\phi\theta} + \Gamma^\phi_{\phi\lambda}\Gamma^\lambda_{\phi\phi} - \frac{1}{2}\Gamma^r_{\phi\phi}\frac{\partial \ln|g|}{\partial r} - \frac{1}{2}\Gamma^\theta_{\phi\phi}\frac{\partial \ln|g|}{\partial \theta}$$

$$= \frac{\partial}{\partial r}\left(\frac{r\sin^2\theta}{A}\right) + \frac{\partial}{\partial\theta}(\sin\theta\cos\theta) + \Gamma^r_{\phi\phi}\Gamma^\phi_{\phi r} + \Gamma^\theta_{\phi\phi}\Gamma^\phi_{\phi\theta} + \Gamma^\phi_{\phi r}\Gamma^r_{\phi\phi} + \Gamma^\phi_{\phi\theta}\Gamma^\theta_{\phi\phi}$$

$$+\frac{1}{2}\sin\theta\cos\theta\frac{\partial\ln\sin^2\theta}{\partial\theta} + \frac{1}{2}\left(\frac{r\sin^2\theta}{A}\right)\left(\frac{A'}{A}+\frac{B'}{B}+\frac{4}{r}\right)$$

$$= \cancel{\frac{\sin^2\theta}{A}} - \frac{rA'\sin^2\theta}{A^2} + \cancel{\cos^2\theta} - \sin^2\theta - \cancel{\frac{\sin^2\theta}{A}} - \cancel{\cos^2\theta} - \frac{\sin^2\theta}{A} - \cancel{\cos^2\theta} + \cancel{\cos^2\theta} + \frac{r\sin^2\theta}{2A}\left(\frac{A'}{A}+\frac{B'}{B}+\frac{4}{r}\right)$$

$$= \sin^2\theta\left[\frac{r}{2A}\left(-\frac{A'}{A}+\frac{B'}{B}\right)+\frac{1}{A}-1\right] = \sin^2\theta R_{\theta\theta}$$

The fact that $R_{\phi\phi} = \sin^2\theta R_{\theta\theta}$ and that $R_{\mu\nu} = 0$ if μ and ν 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 (291):

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

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

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

where U_μ 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 [216]) 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

$$\frac{R_{rr}}{A} + \frac{R_{tt}}{B} = -\frac{1}{rA}\left(\frac{A'}{A}+\frac{B'}{B}\right) = 0 \tag{319}$$

whence we find

$$AB = \text{constant} = 1 \tag{320}$$

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 (316) simply

$$B'' + \frac{2B'}{r} = 0, \tag{321}$$

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 Chapter 3 that the next order term at large distances must be $2\Phi/c^2$ in order to recover the Newtonian limit. Hence,

$$B = 1 - \frac{2GM}{rc^2}, \qquad A = \left(1 - \frac{2GM}{rc^2}\right)^{-1} \tag{322}$$

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

$$\boxed{-c^2 d\tau^2 = -\left(1 - \frac{2GM}{rc^2}\right) c^2 dt^2 + \left(1 - \frac{2GM}{rc^2}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2} \tag{323}$$

This remarkable, simple and critically important exact solution of the Einstein field equations was obtained in 1915 by Karl Schwarzschild from the trenches of World War I. Tragically, Schwarzschild did not survive the war,[10] 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.

Exercise. **The Tolman-Oppenheimer-Volkoff Equation.** Let us strike again while the iron is hot. (If you do this now, you will be ahead of the game when we discuss neutron stars.) Referring back to the previous exercise, we revisit a portion of our Schwarzschild calculation, but with the source terms $S_{\mu\nu}$ retained. Form a familiar combination once again:

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

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?

[10] 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.

Finally, use the equation $R_{\theta\theta} = -8G\pi S_{\theta\theta}/c^4$ together with hydrostatic equilibrium (219) (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 of hydrostatic equilibrium equation for a Newtonian star, with three correction terms. The final factor on the right is purely geometrical, the g_{rr} metric term A. The corrective replacement of ρ 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!*

6.2 Schwarzschild metric in Cartesian coordinates

Note: This section is off-syllabus and non-examinable.

It is useful for a number of applications to have a form of the Schwarzschild metric in Cartesian coordinates $x^i = (x, y, z)$. We shall be particularly interested in the difference between the Schwarzschild and Minkowski metrics, $h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu}$. It is not difficult to switch from spherical to Cartesian coordinate differentials in equation (323). This leads directly to:

$$-c^2 d\tau^2 = \eta_{\mu\nu}dx^\mu dx^\nu + \left(\frac{2GM}{rc^2}\right)c^2dt^2 + \left(\frac{2GM/c^2}{r-2GM/c^2}\right)n_i n_j dx^i dx^j, \tag{324}$$

where

$$n_i = x_i/r \quad \text{and} \quad x_i dx^i = r dr. \tag{325}$$

(Note that the sub/superscript position is unimportant in Cartesian x_i coordinates.) In particular for $r \gg 2GM/c^2$, the leading order departures from Minkowski spacetime are

$$h_{00} = \frac{2GM}{rc^2}, \quad h_{0i} = 0, \quad h_{ij} = \frac{2GMn_in_j}{rc^2}, \quad h \equiv \eta^{\mu\nu}h_{\mu\nu} = 0. \tag{326}$$

There is another set of coordinates with a modified time that will also be of interest when we consider black holes with rotation, so-called *Kerr* black holes. If we make a formal t-coordinate change

$$c d\bar{t} = cdt - \frac{2GM/c^2}{r-2GM/c^2}dr, \tag{327}$$

then, after some algebra, the metric becomes

$$-c^2 d\tau^2 = \eta_{\mu\nu}d\bar{x}^\mu d\bar{x}^\nu + \frac{2GM}{rc^2}\left(c d\bar{t} - n_i d\bar{x}^i\right)^2, \tag{328}$$

where it is now understood that $\eta_{00}dx^0dx^0 = -(cd\bar{t})^2$ in the first term on the right, and $d\bar{x}^i = dx^i$. A little further rearranging of terms allows us to write that this in the form

$$-c^2d\tau^2 = (\eta_{\mu\nu} + \lambda l_\mu l_\nu)d\bar{x}^\mu d\bar{x}^\nu, \tag{329}$$

where

$$\lambda = \frac{2GM}{rc^2}, \quad l^\mu = (1, n^i), \quad l_\nu = \eta_{\mu\nu}l^\mu, \quad dx^0 = cd\bar{t}, \quad d\bar{x}^i = dx^i, \quad \eta_{\mu\nu}l^\mu l^\nu = 0. \tag{330}$$

Note that the h_{ij}, h_{00}, and h Minkowski-departure coefficients are unchanged in these coordinates from their form in equation (326), but that now *off-diagonal* components $h_{0i} = -\lambda n_i$ are present.

In equation (329), the Schwarzschild metric tensor has been decomposed into the Minkowski metric tensor $\eta_{\mu\nu}$ plus a second tensor component $l_\mu l_\nu$ which is Minkowski-orthogonal: l^μ is a null vector with respect to $\eta_{\mu\nu}$. In 1963, the New Zealand mathematician Roy Kerr was investigating the question of whether the were any vacuum metrics, other than Schwarzschild, with this particular Minkowski-plus-null-tensor form of equation. This formulation simplifies (somewhat!) the otherwise extremely complicated nonlinar partial differential equations that emerge from the general field equations. The happy end to this story is that Kerr was indeed able to find a second exact, analytic solution of the vacuum field equations, which imbued the spacetime with rotation! This remarkable achievement ultimately allows the construction of explicit solutions for the most general vacuum black hole solutions possible, including the effects of angular momentum and electrical charge as well as mass. The uniqueness of this solution has been proved. The Kerr metric is the topic of Chapter 7.

6.3 The Schwarzschild Radius

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

$$r = \frac{2GM}{c^2} \equiv R_S \tag{331}$$

the Schwarzschild metric in standard form becomes singular in appearance. R_S is known as the Schwarzschild radius. Numerically, with M measured in units of one solar mass $M_\odot$,

$$R_S = 2.95\,(M/M_\odot)\ \text{km}, \tag{332}$$

which is *well* inside any normal star, and hence not a physical singularity at all. The Schwarzschild radius is part of the external vacuum spacetime only for black holes. As we shall see, 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.

We have been careful to write "singular in appearance" because in fact, the spacetime is perfectly well behaved at $r = R_S$, even for a true black hole. 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. They are literally being stretched to their limits. The coordinate-independent curvature scalar R remains zero without so much as a ripple as we

pass through $r = R_S$. (Just contract the field equations $R_{\mu\nu} = 0$ with $g^{\mu\nu}$.) We can see a similar 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 ϕ is a bad coordinate at the poles, $\theta = 0$ or $\theta = \pi$, which is to say $x = 0$. Bad coordinates happen to good spacetimes. Get over it.

The physical interpretation of the first two terms of the metric (323) is that the proper time interval at a fixed spatial location *external to* R_S, is given by

$$cdt\left(1 - \frac{R_S}{r}\right)^{1/2} \qquad \text{(proper time interval at fixed location, } r > R_S\text{)}. \tag{333}$$

The proper radial distance interval at a fixed angular location and time, also external to R_S, is

$$dr\left(1 - \frac{R_S}{r}\right)^{-1/2} \qquad \text{(proper radial distance interval at fixed time \& angles, } r > R_S\text{)}. \tag{334}$$

But a strange thing happens to the r and t coordinates inside of $r = R_S$. The signs of their contributing terms in the metric $g_{\mu\nu}dx^\mu dx^\nu$ flip. The distinction between space and time is not defined by the r and t characters of the Roman alphabet, it is defined by which of these coordinate variables collaborates with the spacelike solid angle differential $r^2 d\Omega^2$ by adding to it positively, and which stands alone, flaunting its hyperbolic minus sign. The physical interpretation of the first two terms of the metric is now,

$$dr\left(\frac{R_S}{r} - 1\right)^{-1/2} \qquad \text{(proper time interval at fixed location, } r < R_S\text{)}. \tag{335}$$

$$cdt\left(\frac{R_S}{r} - 1\right)^{1/2} \qquad \text{(proper radial distance interval at fixed time \& angles, } r < R_S\text{)}. \tag{336}$$

While nonsingular, the location $r = R_S$ is evidently not without interest! We explore the nature of the Schwarzschild radius further in §6.5.

Exercise. **Getting rid of the Schwarzschild coordinate singularity. A challenge problem for the adventurous student.** Consider the rather unusual coordinate transformation found by Martin Kruskal. Start with our standard spherical coordinates t, r, θ, ϕ and introduce new coordinates u and v as follows. When $r > R_S$,

$$u = (r/R_S - 1)^{1/2}\exp(r/2R_S)\cosh(ct/2R_S)$$

$$v = (r/R_S - 1)^{1/2}\exp(r/2R_S)\sinh(ct/2R_S)$$

and when $r < R_S$,

$$u = (1 - r/R_S)^{1/2} \exp(r/2R_S) \sinh(ct/2R_S)$$
$$v = (1 - r/R_S)^{1/2} \exp(r/2R_S) \cosh(ct/2R_S)$$

Show that in all regions the Schwarzschild metric transforms to

$$-c^2 d\tau^2 = \frac{4R_S^3}{r} e^{-r/R_S}(-dv^2 + du^2) + r^2 d\Omega^2$$

The coordinate singularity at $r = R_S$ is gone, while the true singularity at $r = 0$ remains. Here, r should be regarded as a function of u and v given implicity by

$$(r/R_S - 1)e^{r/R_S} = u^2 - v^2.$$

(An explicit form for $r(u, v)$ may be found in terms of the Lambert W-function, $r/R_S = 1 + W[(u^2 - v^2)/e]$, available in most standard software packages.) This means that $v^2 - u^2 = 1$ is the $r = 0$ singularity. But that can happen at two locations, $v = \pm(1 + u^2)^{1/2}$. *Two* singularities, not one. Moreover, with t allowed to be positive or negative, we see that u and v can both be positive, or one positive and the other negative, but apparently not both negative. That is strange. What are we excluding? The relationship between Schwarzschild coordinates and Kruskal coordinates turns out to be more subtle than you might have guessed; more care needs to be taken with negative square roots in u and v coordinates. In the end, not only do we need to keep all negative roots, if we are to fully understand the true nature of Schwarzschild spacetime, we also need to be careful about what is spacelike and what is timelike when $r < R_S$. The classical Schwarzschild solution in spherical coordinates turns out to be only part of the story: standard Schwarzschild coordinates don't cover all of the spacetime geometry! That was a great surprise when it was discovered by Kruskal in 1960, and in many ways it marked the beginning of our modern approach to black hole theory.

Kruskal coordinates are also strange in the limit $R_S \to 0$, i.e. when the central mass vanishes. The metric does not docilely cooperate by going over to Minkowski space. Instead, it becomes defiantly singular: all u and v structure is lost! Kruskal coordinates make sense only in the spacetime geometry of a Schwarzschild black hole, and not at all in Minkowski space. By contrast, standard Schwarzschild spherical coordinates work fine in the limit of vanishing M, but (we now know) not quite so well with finite M. If you would like to pursue these topics further, see the very nice discussion in MTW, pp. 823–836 for advanced but fascinating reading.

6.4 Why does the determinant of $g_{\mu\nu}$ not change with M?

Note: This section is off-syllabus and non-examinable.

The reader will have noticed that the metric coefficients A and B are reciprocals, so that the absolute value of the metric tensor determinant $ABr^4 \sin^2\theta$ is independent of the mass M, even as the mass completely vanishes. In other words, the 4-volume element $dt\, dr\, r^2 \sin^2\theta\, d\theta d\phi$ remains unchanged when a central mass is added. This is very interesting. Had we known this in advance, we could have carried out our calculation for the Schwarzschild metric, not in the two unknown functions A and B, but with just B alone, using $A = 1/B$. That would have been much easier.

The answer to our posed section question lies with the curvature tensor $R_{\lambda\mu\nu\kappa}$, and its relation to the Ricci tensor $R_{\mu\kappa}$. The Ricci tensor is the linear superposition $g^{\lambda\nu}R_{\lambda\mu\nu\kappa}$. If there are at least as many independent components of $R_{\mu\kappa}$ as there are of $R_{\lambda\mu\nu\kappa}$, then it is possible to invert this expansion, and express the full curvature tensor as a superposition of Ricci terms. For example, from equation (264) we note that in three dimensions there are six independent components of $R_{\lambda\mu\nu\kappa}$, and since there are also six independent components of $R_{\mu\kappa}$ we can indeed express the curvature tensor as a superposition of the Ricci tensor components. It is given explicitly by:

$$R_{\lambda\mu\nu\kappa} = g_{\lambda\nu}R_{\lambda\kappa} - g_{\lambda\kappa}R_{\mu\nu} - g_{\mu\nu}R_{\lambda\kappa} + g_{\mu\kappa}R_{\lambda\nu} - \frac{R}{2}(g_{\lambda\nu}g_{\mu\kappa} - g_{\lambda\kappa}g_{\mu\nu}) \tag{337}$$

where, as usual $R = g^{\mu\nu}R_{\mu\nu}$. It is easy to verify that this expression has all the correct symmetry properties of the curvature tensor indicies, and that contraction of λ with ν returns the Ricci tensor $R_{\mu\kappa}$ on both sides of the equation. It is not very much more difficult to show that this formula is explicitly correct in locally orthogonal coordinates (for which $g_{\mu\nu}$ is diagonal), and being a tensor, must then hold in any coordinate frame. (See W72, page 144.)

Now, if a simple Ricci superposition like equation (337) also held in our Universe of four dimensions, we would be in trouble, since Schwarzschild geometry satisfies $R_{\mu\kappa} = 0$, and there would be no curvature tensor at all! We would be back to Minkowski spacetime. Instead, the general expression for the curvaure tensor in $n \geq 3$ dimensions is

$$R_{\lambda\mu\nu\kappa} = \frac{1}{n-2}(g_{\lambda\nu}R_{\lambda\kappa} - g_{\lambda\kappa}R_{\mu\nu} - g_{\mu\nu}R_{\lambda\kappa} + g_{\mu\kappa}R_{\lambda\nu}) - \frac{R}{(n-1)(n-2)}(g_{\lambda\nu}g_{\mu\kappa} - g_{\lambda\kappa}g_{\mu\nu}) + C_{\lambda\mu\nu\kappa} \tag{338}$$

where $C_{\lambda\mu\nu\kappa}$ is now an additional tensor component known as the *Weyl tensor*. The two coefficients depending on n ensure that contracting on $\lambda\nu$ returns the proper Ricci tensor, because the Weyl tensor is defined to be *traceless* (i.e. it vanishes) on contraction of any two indices. When $n = 4$, as in our beloved Universe, it is OK that $R_{\mu\kappa} = 0$: the curvature tensor then consists entirely of a nonvanishing Weyl tensor, so curvature *does* exist. But now the curvature tensor is traceless! *That means it cannot be associated with any change in volume associated with distortion in spacetime.* All volume distortions are contained in the Ricci tensor components, whereas the Weyl tensor embodies only volume conserving distortions.

The unchanging determinant of the standard Schwarzschild metric is a reflection of this property of the curvature tensor. The 4-volume element $r^2 \sin\theta \, dr \, d\theta d\phi$ is preserved for these particularly simple and symmetric coordinates. Yet more striking, the metric determinant also does not change for a Kerr (spinning) black hole in its standard coordinates, despite the great complexity of the metric. Both geometries satisfy $R_{\mu\kappa} = 0$, and have only the Weyl contribution to the full curvature tensor. But be careful. Because the determinant is not itself a coordinate invariant, we cannot conclude the inverse is true, that a change in the determinant with mass necessarily implies a nonvanishing Ricci tensor. Look at the Schwarzschild metric in Kruskal coordinates from the above Exercise for a nice counter example.

6.5 Working with Schwarzschild spacetime.

6.5.1 Radial photon geodesic

The gravitational redshift recorded by an observer at infinity relative to a location r in the external Schwarzschild spacetime is given (we now know) *precisely* by

$$dt = \frac{d\tau}{(1 - R_S/r)^{1/2}} \quad \text{(Exact)}, \tag{339}$$

so that at $r \to R_S$, 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 to 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 also 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 inside of $r = R_S$. R_S is said to be the location of the *event horizon.* The singularity at $r = 0$ is present in the spacetime, but completely hidden from the outside world at $R = R_S$ within an event horizon. Roger Penrose has aptly named this "cosmic censorship."

The time coordinate change for light to travel from r_A to r_B along a radial geodesic path is given by setting

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

and then computing

$$t_{AB} = \int_A^B dt = \frac{1}{c}\int_{r_A}^{r_B} \frac{dr}{(1 - R_S/r)} = \frac{r_B - r_A}{c} + \frac{R_S}{c}\ln\left(\frac{r_B - R_S}{r_A - R_S}\right). \tag{340}$$

This 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 either path endpoint reaches R_S. When R_S may be considered small over the entire integration path, to leading order we have

$$t_{AB} \simeq \frac{r_B - r_A}{c} + \frac{R_S}{c}\ln\left(\frac{r_A}{r_B}\right) = \frac{r_B - r_A}{c}\left(1 + \frac{R_S \ln(r_A/r_B)}{r_B - r_A}\right). \tag{341}$$

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 less than 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

We now formally solve the geodesic equation for the orbits around a point mass. The geodesic equation is

$$\frac{d^2x^\lambda}{d\tau^2} + \Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0 \tag{342}$$

Although $d\tau = 0$ for photons, we shall see that photon orbits may be precisely computed by a limiting procedure from the finite τ analysis.

The explicit form of the orbital equations themselves is most easily derived by starting with the Euler-Lagrange equations using the Lagrangian

$$\mathcal{L} = g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = -B(r)c^2\dot{t}^2 + A(r)\dot{r}^2 + r^2\dot{\phi}^2 \tag{343}$$

where the dot represents $d/d\tau$, just as we learned at the end of section 3.5, page 38. Here, we have fixed the orbital plane to $\theta = \pi/2$, so that $\dot{\theta}$ vanishes identically. A and B depend explicitly on r and implicitly on τ via $r = r(\tau)$. Recall that for large r, A and B both approach unity. The Euler-Lagrange equation for time t is simple, since there is no $\partial\mathcal{L}/\partial t$:

$$\frac{d}{d\tau}\left(\frac{\partial\mathcal{L}}{\partial\dot{t}}\right) = -2c^2\frac{d}{d\tau}\left(B\,\frac{dt}{d\tau}\right) = 0, \tag{344}$$

and the same holds true analogously for the ϕ equation:

$$\frac{d}{d\tau}\left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\right) = 2\frac{d}{d\tau}\left(r^2\frac{d\phi}{d\tau}\right) = 0. \tag{345}$$

These equations are both very easy to solve. The solution of the time equation

$$\frac{dt}{d\tau} = \frac{\gamma_\infty}{B}, \tag{346}$$

where γ_∞ is a constant of integration. The notation reminds us that if the particle reaches infinity where $B = 1$, this constant is in fact the actual relativistic γ factor. In this case, $\gamma_\infty \geq 1$. If $\gamma_\infty < 1$, the particle remains in a bound orbit, never reaching infinity. In the nonrelativistic limit, γ_∞ is always very close to 1. The approach to unity is from above for an unbound orbit, from below for a bound orbit.

General relativity also conserves angular momentum for an axisymmetric spacetime, just like ordinary mechanics. The angular variable ϕ is determined by the simple, Newtonian-like, differential equation:

$$r^2\frac{d\phi}{d\tau} = \gamma_\infty J \qquad \text{(constant.)} \tag{347}$$

The constant J has dimensions of angular momentum. The factor γ_∞ has been inserted for later convenience, allowing for the difference between $d\tau$ and dt. (The photon limit corresponds to $\gamma_\infty \to \infty$, J finite.)

We could now proceed with the formal Euler-Lagrange equation for r, but there is a much more direct option available to us. Simply write down the line element, using the last two derived results for $dt/d\tau$ and $d\phi/d\tau$. This gives:

$$A\left(\frac{dr}{d\tau}\right)^2 + \frac{\gamma_\infty^2 J^2}{r^2} - \frac{c^2\gamma_\infty^2}{B} = -c^2, \tag{348}$$

Explicitly substituting for B in (348), we find that the extremal (maximum and minimum) values of orbital radius r for a bound orbit (when $dr/d\tau = 0$) are solutions of

$$\left(1 - \frac{R_S}{r}\right)\left(\frac{J^2}{r^2} + \frac{c^2}{\gamma_\infty^2}\right) - c^2 = 0 \tag{349}$$

for ordinary matter, and of

$$\left(1 - \frac{R_S}{r}\right)\frac{J^2}{r^2} - c^2 = 0 \tag{350}$$

for photons, $\gamma_\infty \to \infty$. (A bound photon orbit can only be circular, and the orbit is in fact unstable.)

The radial equation of motion may be written for either $dr/d\tau$, dr/dt, or $dr/d\phi$ respectively (we use $AB = 1$):

$$\left(\frac{dr}{d\tau}\right)^2 + B\left(c^2 + \frac{J^2\gamma_\infty^2}{r^2}\right) = c^2\gamma_\infty^2 \tag{351}$$

$$\left(\frac{dr}{dt}\right)^2 + B^3\left(\frac{c^2}{\gamma_\infty^2} + \frac{J^2}{r^2}\right) = B^2c^2 \tag{352}$$

$$\left(\frac{dr}{d\phi}\right)^2 + r^2B\left(1 + \frac{c^2r^2}{\gamma_\infty^2 J^2}\right) = \frac{c^2r^4}{J^2} \tag{353}$$

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, practical analytic progress is possible.

Exercise. A case where it is possible to solve the orbital equations explicitly is that of radial infall, $J = 0$. Show the solution to equation (351) for the comoving time τ just reduces to the classical Newtonian problem! For the simple case of zero velocity at infinite distance ($\gamma_\infty = 1$), show that the infall solution is

$$\frac{c\tau}{R_S} = \text{constant} \; - \; \frac{2}{3}\left(\frac{r}{R_S}\right)^{3/2},$$

the constant determined by where we choose the zero of τ. The key point is that an infalling radial particle reaches R_S in a finite comoving time from a finite external radius. By contrast, show that the solution to the t-equation (352) is given by

$$\frac{ct}{R_S} = \text{constant} \; - \; \frac{2}{3}\left(\frac{r}{R_S}\right)^{3/2} - 2\left(\frac{r}{R_S}\right)^{1/2} + \ln\left|\frac{(r/R_S)^{1/2} + 1}{(r/R_S)^{1/2} - 1}\right|.$$

Take the limit $r \to R_S$ and demonstrate that

$$r/R_S - 1 \to \text{constant} \; \times \; \exp(-ct/R_S),$$

so that an *infinite* amount of coordinate time—the time on a distant observer's clock—must pass before the test mass reaches R_S from any finite external radius.

Exercise. Another case of interest is that of circular orbits. Here, $U^0 = cdt/d\tau$ and $U^\phi = d\phi/d\tau$ may be solved directly from

$$g_{\mu\nu}U^\mu U^\nu = -c^2, \quad (\partial_r g_{\mu\nu})U^\mu U^\nu = 0$$

The first of these equations is merely the 4-velocity normalisation from the fundamental line element. Explain why the second equation is true. (Hint: see eq. [230].) Show that

$$U^0 = \frac{c}{(1-3r_g/r)^{1/2}}, \quad U^\phi = \sqrt{\frac{GM/r^3}{1-3r_g/r}}, \quad U_0 = -\frac{c(1-2r_g/r)}{(1-3r_g/r)^{1/2}}, \quad U_\phi = \sqrt{\frac{GMr}{1-3r_g/r}},$$

where $r_g = GM/c^2$. Expand U_0 to first order in r_g/r. Relate your result to the Newtonian energy of a circular orbit. Finally, explain how U_0 and U_ϕ are related to the constants γ_∞ and $\gamma_\infty J$. We will revisit this problem when we study Kerr spacetime in Chapter 7.

Exercise. Consider next a small departure from circular orbit, which conserves the value of the angular momentum U_ϕ of that orbit. (You should have found that $U_\phi = \gamma_\infty J$ in the above exercise.) Show that differentiation of equation (351) with respect to τ leads to an equation of the form

$$\ddot{r} + F(r) = 0.$$

Determine the function F(r), which should depend on the constants GM, c, and $\gamma_\infty J$.

Next, let r_c be a circular orbit value satisfying $F(r_c) = 0$. For small departures $x = r - r_c \ll r_c$ which conserve the angular momentum of the r_c circular orbit, show that

$$\ddot{x} + \kappa^2 x = 0,$$

where

$$\kappa^2 = \frac{GM}{r_c^3}\left(\frac{1-6r_g/r_c}{1-3r_g/r_c}\right), \quad r_g = \frac{GM}{c^2}$$

Why does this mean that circular orbits with $3r_g < r_c < 6r_g$ are unstable?

6.6 The deflection of light by an intervening body.

Oh leave the Wise our measures to collate
One thing at least is certain, light has weight
One thing is certain and the rest debate
Light rays, when near the Sun, do not go straight.

— A. S. Eddington, 1919, after-dinner speech at the Royal Astronomical Society

The first measurable, new-physics prediction made by general relativity was that starlight passing by the limb of the Sun would be deflected by the gravitational field. This type of optical 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 happened to occur on 29 May 1919. The path of totality extended along a narrow band starting in South America and continuing to central Africa. An expedition headed by Sir Arthur Eddington observed

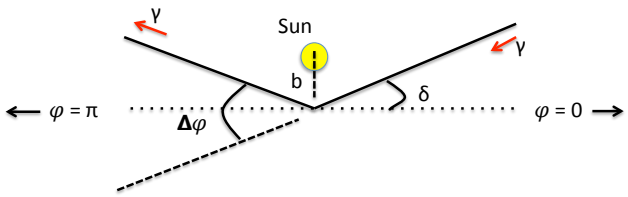


Figure 2: Bending of light by the gravitational field of the Sun. In flat spacetime the photon γ travels the straight line from $\varphi = 0$ to $\varphi = \pi$ along the path $r \sin \varphi = b$. In the presence of spacetime curvature, a photon starting at angle $\varphi = -\delta$ would finish its passage at $\varphi = \pi + \delta$. The deflection angle is thus $\Delta\varphi = 2\delta$.

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 affected the propagation of light, but that it did so by an amount in much better accord with general relativity theory than with an *ad hoc* Newtonian "corpuscular theory," one with a test mass velocity set equal to c. (The latter gives a deflection angle only half as large as GR, in essence because the R_S/r 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 and international athletes. *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 used 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 a powerful diagnostic tool, and the motivaton behind the *Euclid* space-based telescope mission, to be launched in 2022.

Back to the classic 1919 light-deflection test. As in Newtonian dynamics, it turns out to be easier to work with $u \equiv 1/r$, in which case

$$\left(\frac{du}{d\phi}\right)^2 = \frac{1}{r^4}\left(\frac{dr}{d\phi}\right)^2. \tag{354}$$

Equation (353) with $\gamma_\infty \to \infty$ for a photon may be written

$$\frac{1}{r^4}\left(\frac{dr}{d\phi}\right)^2 + \frac{B}{r^2} = \frac{c^2}{J^2} = \text{constant}. \tag{355}$$

In terms of u:

$$\left(\frac{du}{d\phi}\right)^2 + u^2\left(1 - \frac{2GMu}{c^2}\right) = \frac{c^2}{J^2}. \tag{356}$$

Differentiating with respect to ϕ ($du/d\phi \equiv u'$) leads quickly to

$$u'' + u = \frac{3GM}{c^2}u^2 \equiv 3\epsilon u^2. \tag{357}$$

Treating $\epsilon \equiv GM/c^2$ as a small parameter, we expand u as $u = u_0 + u_1$, with $u_1 = \mathrm{O}(\epsilon u_0^2) \ll u_0$. (Read "$u_1$ is of order ϵ times u_0^2 and much smaller than u_0". The attentive reader will note that ϵ and r_g are identical; we use the ϵ notation as a reminder that this is a small perturbative quantity.) Then, the leading order larger terms must obey the equation

$$u_0'' + u_0 = 0, \tag{358}$$

and the smaller terms of order ϵ must obey the equation

$$u_1'' + u_1 = 3\epsilon u_0^2. \tag{359}$$

To leading order ($u = u_0$), nothing happens: the photon moves in a straight line. If the radius of closest approach is the impact parameter b, then the equation for a straight line is $r\sin\phi = b$, or

$$u_0 = \frac{\sin\phi}{b} \tag{360}$$

which is the unique solution to equation (358) with boundary conditions $r = \infty$ at $\phi = 0$ and $\phi = \pi$. (Note that b is the radius of the body for a grazing encounter.)

At order ϵ, there is a deflection of the orbit from a straight line due to the presence of u_1:

$$u_1'' + u_1 = 3\epsilon u_0^2 = \frac{3\epsilon}{b^2}\sin^2\phi = \frac{3\epsilon}{2b^2}(1 - \cos 2\phi) \tag{361}$$

We therefore search for solutions of the form $u_1 = U + V\cos 2\phi$, where U and V are constants. Substituting this form into (361), we straightforwardly find $U = 3\epsilon/2b^2$ and $V = \epsilon/2b^2$. Our solution is then then

$$\frac{1}{r} = u_0 + u_1 = \frac{\sin\phi}{b} + \frac{3\epsilon}{2b^2} + \frac{\epsilon\cos 2\phi}{2b^2} \tag{362}$$

With $\epsilon = 0$, the solution describes the straight line $r\sin\phi = b$. The first order effects of including ϵ 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 π, but at the slightly corrected values $\phi = -\delta$ and $\phi = \pi + \delta$ where

$$\delta = \frac{2\epsilon}{b} \tag{363}$$

(See figure [2].) In other words, there is now a total deflection angle $\Delta\phi$ from a straight line of 2δ, or

$$\Delta\phi = \frac{4GM}{bc^2} = \frac{2R_S}{b} = 1.75 \text{ arcseconds for the Sun.} \tag{364}$$

The numerical value assumes that $b = R_\odot$, one solar radius. In practise, the actual numbers would have been just under an arsecond. Happily, arcsecond deflections were 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.

In more recent times, very long baseline radio interferometry (VLBI) can measure the solar deflection of radiation from pointlike extragalactic radio sources. Agreement with general relativity is at the level of 0.02% accuracy.

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? (See figure [3].) 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. The fully self-consistent form of the field equations still eluded Einstein at this point, but he at last he knew he was almost certainly on the right track! 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."

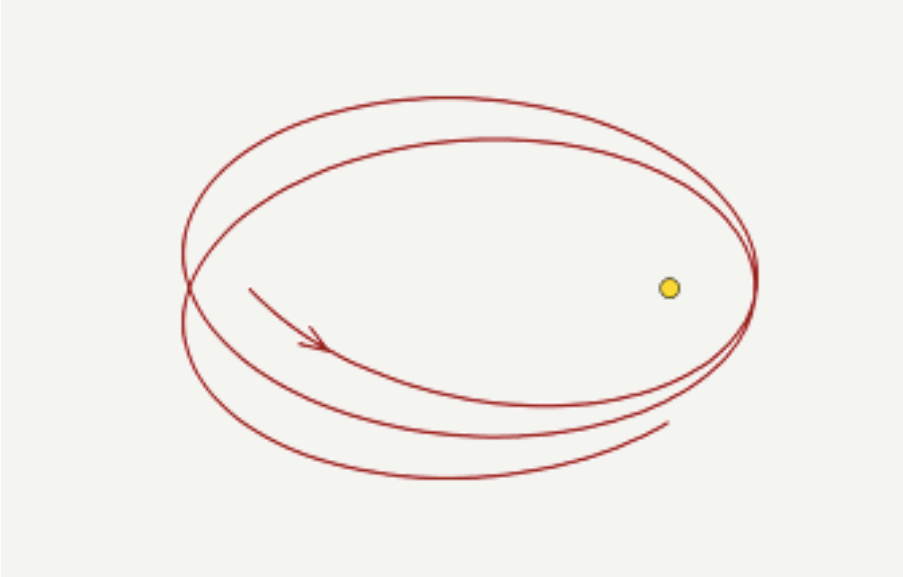


Figure 3: Departures from a $1/r$ gravitational potential cause elliptical orbits not to close. In the case of Mercury, the perihelion advances by 43 seconds of arc per century. The effect is shown here, *greatly* exaggerated.

6.7.1 Newtonian orbits

Interestingly, the first-order GR calculation of the perihelion advance is not much more difficult than a straight Newtonian perturbation calculation would be. We therefore begin with a detailed review of the Newtonian problem, and we will play off this solution for the GR perihelion advance.

Newtonian conservation of energy is

$$\frac{v_r^2}{2} + \frac{J^2}{2r^2} - \frac{GM}{r} = \mathcal{E}, \tag{365}$$

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, 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 (351), whose exact form we may write as

$$\frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 - \frac{GM}{r} + \frac{J^2\gamma_\infty^2}{2r^2}\left(1 - \frac{R_S}{r}\right) = \frac{c^2}{2}(\gamma_\infty^2 - 1). \tag{366}$$

This reverts to Newtonian form when $|\gamma_\infty - 1| \ll 1$, and the term in R_S/r is dropped. We may then define the Newtonian energy quantity $\mathcal{E}$ by

$$\mathcal{E} = c^2(\gamma_\infty^2 - 1)/2 \simeq c^2(\gamma_\infty - 1). \tag{367}$$

If the mass reaches infinity with finite kinetic energy $v_\infty^2/2$, this is agrees with the identification of γ_∞^2 with $1/(1 - v_\infty^2/c^2)$. On the other hand, if $\gamma_\infty < 1$, then $\mathcal{E} < 0$ and the particle is bound. In other words, in the Newtonian limit we may set $\gamma_\infty = 1$ everywhere, except when it appears as the difference $\gamma_\infty - 1$, when more care is needed.

The Newtonian energy equation may be solved for v_r:

$$v_r = \frac{dr}{d\phi}\frac{d\phi}{dt} = \frac{J}{r^2}\frac{dr}{d\phi} = \pm\left(2\mathcal{E} + \frac{2GM}{r} - \frac{J^2}{r^2}\right)^{1/2} \tag{368}$$

and thence separated:

$$\int \frac{J\,dr}{r^2\left(2\mathcal{E} + \frac{2GM}{r} - \frac{J^2}{r^2}\right)^{1/2}} = \pm\phi \tag{369}$$

With $u = 1/r$,

$$\int \frac{du}{\left(\frac{2\mathcal{E}}{J^2} + \frac{2GMu}{J^2} - u^2\right)^{1/2}} = \mp\phi \tag{370}$$

or

$$\int \frac{du}{\left[\frac{2\mathcal{E}}{J^2} + \frac{G^2M^2}{J^4} - \left(u - \frac{GM}{J^2}\right)^2\right]^{1/2}} = \mp\phi \tag{371}$$

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:

$$\cos^{-1}\left[\frac{u - \frac{GM}{J^2}}{\left(\frac{2\mathcal{E}}{J^2} + \frac{G^2M^2}{J^4}\right)^{1/2}}\right] = \pm\phi. \tag{372}$$

The integration constant, a mere orbital phase, may be set to zero without loss of generality. With $r = 1/u$, this rather cumbersome equation unfolds and simplifies down to

$$r = \frac{J^2/GM}{1 + \epsilon\cos\phi}, \qquad \epsilon^2 \equiv 1 + \frac{2\mathcal{E}J^2}{G^2M^2} \tag{373}$$

For $\mathcal{E} < 0$, we see that $\epsilon < 1$, and that (373) is just the equation for a an elliptical orbit of eccentricity ϵ. More generally, this is the classic equation for a conic section, with hyperbolic ($\epsilon > 1$), parabolic ($\epsilon = 1$) and elliptical ($\epsilon < 1$) solutions.

For the elliptical orbits, we may identify the *semilatus rectum* (r at $\phi = \pi/2$):

$$L = J^2/GM, \tag{374}$$

the perihelion (radius of closest approach) r_- and aphelion[11] (radius of farthest extent) r_+,

$$r_- = \frac{L}{1+\epsilon}, \qquad r_+ = \frac{L}{1-\epsilon}, \qquad \frac{1}{L} = \frac{1}{2}\left(\frac{1}{r_+} + \frac{1}{r_-}\right), \tag{375}$$

[11]The terms *perihelion* and *aphelion* refer specifically to a planetary orbit about the Sun. More generally, one speaks of *periastron* and *apastron*.

and the semi-major axis

$$a = \frac{1}{2}(r_+ + r_-), \qquad \text{whence} \quad L = a(1-\epsilon^2). \tag{376}$$

Notice that the zeros of the denominator in the integral (371) occur at $u_- = 1/r_-$ and $u_+ = 1/r_+$, corresponding, in our arccosine function, to ϕ equals 0 and π respectively.

Exercise.) **A shortcut.** Show that with $u = 1/r$, the orbital equation is simply

$$u'' + u = GM/J^2$$

where $u' = dy/d\phi$. Up to an arbtrary zero for ϕ, the solution is clearly $u = \epsilon\cos\phi + GM/J^2$, where ϵ is an integration constant of the equation. This is completely consistent with (373), but without more work we don't know how ϵ depends upon the energy and angular momentum of the orbit! The method in the text provides us with this information.

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 πab. Show that the total energy of a two-body bound system (masses m_1 and m_2) is $-Gm_1m_2/2a$, independent of ϵ. 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 ϵ. (There is a very simple way to do the latter!)

6.7.2 The relativistic orbit of Mercury

Equation (353) may be written in terms of $u = 1/r$ as

$$\left(\frac{du}{d\phi}\right)^2 + \left(1 - \frac{2GMu}{c^2}\right)\left(u^2 + \frac{c^2}{\gamma_\infty^2 J^2}\right) = \frac{c^2}{J^2}. \tag{377}$$

Now differentiate with respect to ϕ and simplify. The resulting equation (see the above Exercise!) is:

$$u'' + u = \frac{GM}{\gamma_\infty^2 J^2} + \frac{3GMu^2}{c^2} \simeq \frac{GM}{J^2} + \frac{3GMu^2}{c^2}, \tag{378}$$

since γ_∞ is very close to 1 for nonrelativistic Mercury. (Mercury's precisely measured semilatus rectum is formally $J^2\gamma_\infty^2/GM$; allocating tiny differences between J and $\gamma_\infty \simeq 1$ does not affect the perihelion advance calculation at all.) The Newtonian limit ($u \equiv u_N$) corresponds to dropping the final term on the right side of the equation; the resulting solution is that of §(6.7.1):

$$u = u_N = \frac{GM}{J^2}(1 + \epsilon\cos\phi), \tag{379}$$

where ϵ is the constant eccentricity, not assumed to be small here!

If you are looking for something small, on the right side of equation (378), $3GMu^2/c^2$ is tiny. We therefore view this as a perturbative correction, and to leading order may use the known Newtonian solution u_N^2 in this term for u^2. Writing $u = u_N + \delta u$, the differential equation becomes

$$\frac{d^2\delta u}{d\phi^2} + \delta u = \frac{3GM}{c^2}u_N^2 = \frac{3(GM)^3}{c^2J^4}(1 + 2\epsilon\cos\phi + \epsilon^2\cos^2\phi). \tag{380}$$

In Problem Set 2, you will be asked to solve this equation. The resulting solution for $u = u_N + \delta u$ may be written

$$u \simeq \frac{GM}{J^2}\left(1 + \epsilon\cos[\phi(1-\alpha)]\right) \tag{381}$$

where $\alpha = 3(GM/Jc)^2$. Thus, the perihelion occurs not with a ϕ-period of 2π, but with a slightly longer period of

$$\frac{2\pi}{1-\alpha} \simeq 2\pi + 2\pi\alpha, \tag{382}$$

i.e. an advance of the perihelion by an amount

$$\Delta\phi = 2\pi\alpha = 6\pi\left(\frac{GM}{Jc}\right)^2 = 6\pi\left(\frac{GM}{c^2 L}\right) = 2.783\times 10^{-6}\ \left(\frac{10^{10}\mathrm{m}}{L}\right) \tag{383}$$

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 42.98 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 (383) agreed so precisely with observations, he felt as though something inside him actually "snapped."

From the year of its discovery in 1915, until the 1982 gravitational radiation measurement of the binary pulsar 1913+16, the very precise accord of this theoretical prediction with the observed Mercury perihelion advance was regarded as general relativity's greatest observational success. By using radar ranging techniques (see §6.10 below), it is now possible to measure the perihelion advance of Mercury to 0.002% accuracy. The measured value is $42.98(1 \pm 0.001)$.

Exercise. The general relativistic perihelion advance for Venus and the Earth have also by now been observed. Show that the predicted angular shifts are 8.62 and 3.83 arcseconds per century, for Venus and Earth respectively. Look up the planetary orbit data you need from online sources or elsewhere.

Exercise. Show that equation (378) has an *exact* solution of the form

$$u = \frac{c^2\alpha}{GM} + \frac{c^2\beta}{GM\phi^2}$$

provided that

$$\alpha = \frac{1}{6}, \quad \beta = 2, \quad \frac{GM}{\gamma_\infty^2 J^2} = \frac{c^2}{12GM}$$

The solution for r is then

$$r = \frac{6GM/c^2}{1 + 12/\phi^2}$$

This is no mere quirk, but a solution of astrophysical interest. It corresponds to a test mass spiraling inward from $r = 6GM/c^2$, a marginally stable circular orbit, to $r = 0$. The value of ϕ runs from $-\infty$ up to 0 at $r = 0$. We will return to this remarkable solution for further discussion in Chapter 7.

6.8 General solution for Schwarzschild orbits

Note: This section is off-syllabus and non-examinable.

6.8.1 Formulation

An explicit and general solution for the orbits in a Schwarzschild metric may be found in terms of what is known as an *elliptic integral,* or more precisely, an "elliptic integral of the first kind." Today, it is part of any modern software package of special functions, like MATLAB or *Mathematica,* and not especially complicated. The function is denoted as $F(\phi\backslash\alpha)$, and is given by

$$F(\phi\backslash\alpha) = \int_0^\phi \frac{d\theta}{\sqrt{1-(\sin^2\alpha)\sin^2\theta}}. \tag{384}$$

With all quantitites real, this is a perfectly well-behaved, monotonically increasing function for $0 \le \phi < \pi/2$. If $\sin^2\alpha < 1$ this is well-defined for any ϕ. It would all be rather dull, except that when $\alpha = \pi/2$, there is a logarithmic singularity for $\phi \to \pi/2$. An alternative formulation of this function, one that is particularly useful here, is obtained by changing variables to $t = \sin\theta$. It is customary when using this form to write $m = \sin^2\alpha$ and $x = \sin\phi$:

$$F(\phi|m) = \int_0^x \frac{dt}{\sqrt{(1-t^2)(1-mt^2)}} = F(\sin^{-1}x|m) \tag{385}$$

Note the use of a bar (|) here, rather than a backslash (\) . This is all purely a matter of convention. We have thus far followed the notation of the standard reference *Handbook of Mathematical Functions,* by Abramowitz and Stegun[12]. To avoid any confusion, from here on we will use the "$F(\sin^{-1}x|m)$" formulation, and reserve the ϕ notation exclusively for the orbital azimuthal angle.

The orbital equation in its most general form may be taken directly from equation (377):

$$(u')^2 = \frac{2GM}{c^2}u^3 - u^2 + \frac{2GMu}{\gamma_\infty^2 J^2}u + \frac{c^2}{J^2\gamma_\infty^2}(1-\gamma_\infty^2) \tag{386}$$

where $u = 1/r$. Writing $\tilde{u} = R_S u$, in dimensionless form this equation becomes

$$(\tilde{u}')^2 = \tilde{u}^3 - \tilde{u}^2 + \frac{c^2R_S^2}{\gamma_\infty^2 J^2}\,\tilde{u} + \underbrace{\frac{c^2(1-1/\gamma_\infty^2)R_S^2}{J^2}}_{2\mathcal{E}R_S^2/J^2}. \tag{387}$$

In the Newtonian limit, the final three terms on the right side of this equation are all comparable and of order $(R_S/r)^2$; the $\tilde{u}^3$ term is an asymptotic order in R_S/r yet smaller. (The final term is related to the Newtonian energy $\mathcal{E}$ by the underbraced expression.) For the parameter space of interest, the cubic polynomial in $\tilde{u}$ on the right side of equation (387) has three real roots, $p_1 \le p_2 \le p_3$. The p_1 root may be positive or negative according to the sign of $\mathcal{E}$; p_2 and p_3 are always positive. If $\mathcal{E} \ge 0$ then $p_1 \le 0$ and the test mass may escape to infinity with non-negative energy. In Newtonian systems, the root p_2 is also very small (i.e., it corresponds to a radius *large* compared with R_S), while p_3 is always of order unity.

[12]Now available as an App: search for *Book of Mathematical Functions.*

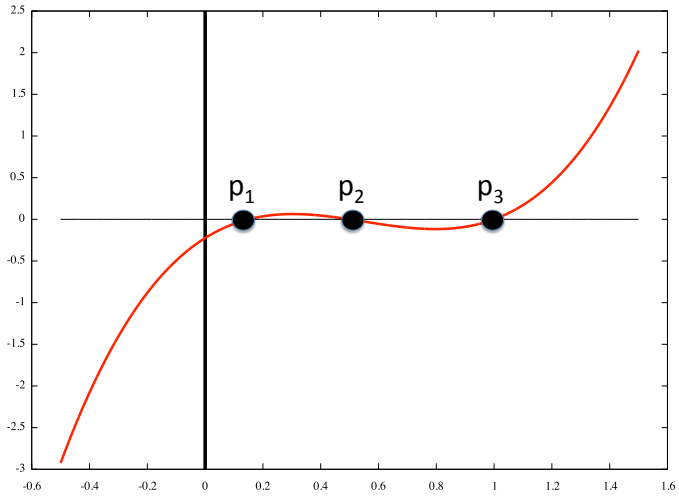


Figure 4: Representative graph of $\tilde{u}$ cubic polynomial from equation (387) for the case of a bound orbit. The dimensionless radius r/R_S varies between $1/p_2$ and $1/p_1$. (Numerical values chosen for ease of plotting.)

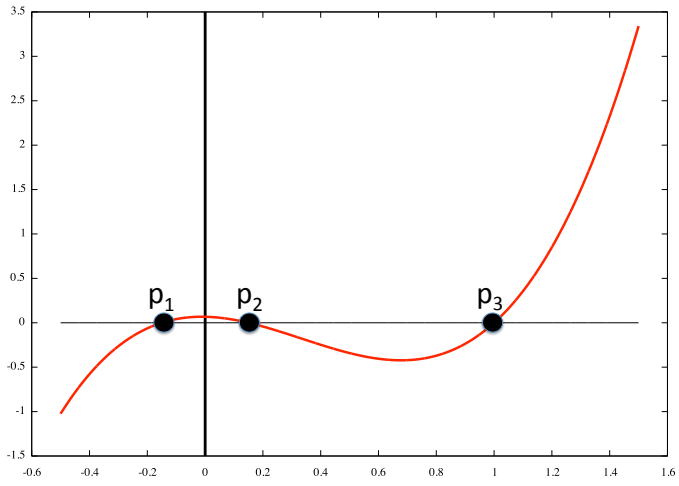


Figure 5: Similar to figure (4), but for the case of unbound orbits. In this case, $1/p_2$ corresponds to the radius of closest approach.

When $\tilde{u} > 1$, the mass is inside of R_S and headed unavoidably for $r = 0$. When $\mathcal{E} < 0$, the roots p_1 and p_2, both small and positive, and define the inner and outer orbital radii respectively.

For photon orbits, which also correspond to the asymptotic limit of any extreme relativistic motion at large distances from the origin, we work in the limit $\gamma_\infty \to \infty$. Our equation then reduces to:

$$(\tilde{u}')^2 = \tilde{u}^3 - \tilde{u}^2 + \frac{c^2 R_S^2}{J^2} \quad \text{(photon orbits)} \tag{388}$$

6.8.2 Solution

We shall assume that the equations (387) or (388) may be written with the cubic on the right side in factored form, with three real roots, $p_1 \le p_2 \le p_3$:

$$(\tilde{u}')^2 = (\tilde{u} - p_1)(\tilde{u} - p_2)(\tilde{u} - p_3) \tag{389}$$

Recall that $p_1 \le 0$ corresponds to unbound orbits, $p_1 > 0$ to bound orbits. (See figures [4] and [5].) This differential equation separates cleanly, and may be integrated:

$$\phi + \phi(0) = \int_{\tilde{u}_1}^{\tilde{u}_2} \frac{d\tilde{u}}{\sqrt{(\tilde{u} - p_1)(\tilde{u} - p_2)(\tilde{u} - p_3)}}, \tag{390}$$

where $\phi(0)$ is an initial phase constant which we retain here for later convenience (which is why we leave the integral on the right in generic form). The + sign for the square root has been selected, assuming that the orbital motion corresponds to ϕ increasing with time. Next, change variables in the integral to $y = \tilde{u} - p_1$. There results

$$\int_{\tilde{u}_1}^{\tilde{u}_2} \frac{d\tilde{u}}{\sqrt{(\tilde{u} - p_1)(\tilde{u} - p_2)(\tilde{u} - p_3)}} = \int_{y_1}^{y_2} \frac{dy}{\sqrt{y(y + p_1 - p_2)(y + p_1 - p_3)}} \tag{391}$$

Finally, set $y = (p_2 - p_1)t^2$. After simplification, the integral becomes

$$\frac{2}{\sqrt{p_3 - p_1}} \int_0^x \frac{dt}{\sqrt{(1 - t^2)(1 - mt^2)}}, \quad m \equiv \frac{p_2 - p_1}{p_3 - p_1}, \quad x = \sqrt{\frac{\tilde{u} - p_1}{p_2 - p_1}}. \tag{392}$$

Note that we have now made our choice for the integration bounds. The value of x shown follows straightforwardly from our change of variable path from $\tilde{u}$ through t. Note as well that $x, m \le 1$. It is not by accident that the above integral has exactly the same form as that of equation (385). Therefore, our final solution for the orbital ϕ angle as a function of r becomes

$$\phi + \pi = \frac{2}{\sqrt{p_3 - p_1}} F\left(\sin^{-1} \sqrt{\frac{R_S/r - p_1}{p_2 - p_1}} \,\middle|\, \frac{p_2 - p_1}{p_3 - p_1} \right). \tag{393}$$

We have chosen the phase $\phi(0) = \pi$ in order to reproduce the form of our Newtonian equation (373) for classical elliptical orbits in the low energy limit. This is the closed-form solution relating ϕ to r for orbits in a Schwarzschild geometry.

Exercise. Verify the Newtonian limit of (393), solving to leading order for p_1, p_2, p_3. Note that for small m, the integral in equation (392) reduces to $\sin^{-1} x$, and the F-function to $F(X|m) \simeq X$.

6.9 Gravitational Collapse

The relativistic theory of gravitational collapse begins with the calculation of a spherical distribution of matter evolving toward a black hole end state. This was first developed in a seminal study by Oppenheimer and Snyder (1939). They considered a spherical cloud of "dust," i.e., pressureless matter, evolving under the evolution of its own self-gravity. This work marks the first reckoning with the concept of the formation of a singularity in general relativity from nonsingular initial conditions.

Using what we have just learned, we may calculate the redshift observed by a distant stationary observer from any emission by the surface of the sphere. We make use the fact that, from the point-of-view of an unfortunate comoving observer perched upon the outer radius of the dust cloud, the entire underyling spherical distribution of matter may just as well be replaced with a Schwarzschild black hole right from the very start of the infall. This is no different from the Newtonian theorem that *an extended spherical distribution of matter exerts exactly the same force on an exterior test mass that it would as a singular mass point concentrated at the origin.* In a collapse calculation, the comoving observer is always just barely exterior to the sphere of matter, so the mass enclosed beneath their feet is constant, and may be regarded as acting at the centre of the sphere. This general relativistic upgrade of the classic Newtonian theorem was first proved by Birkhoff in 1923: any spherical distribution of matter, static or not, produces the Schwarzschild metric exterior to the mass. The Newtonian "empty shell theorem" also holds: any spherical distribution of matter *exterior* to an observer produces no spacetime distortion within the spherical cavity. These two results are collectively known as *Birkhoff's Theorem.* This theorem allows us to treat the redshift calculation as an exercise in Schwarzschild radial orbits. (It also justifies a Newtonian treatment of cosmology, as we shall later see.)

We begin by relating an infalling student's comoving (proper) time τ to the time $\bar{t}$ of their distant supervisor, who is safely off at infinity, filling out a lengthy online Risk Assessment Form for the project. (Here, we shall use the notation $\bar{t}$, instead of the more usual t, for compatibility with the metric calculation of Chapter 11.) The conversion between τ and $\bar{t}$ is simply a matter of solving the standard dynamical equations for a $J = 0$ radial orbit in Schwarzschild geometry. For reference, we note equation (351) for $dr/d\tau$ with $J = 0$ is:

$$\frac{1}{c^2}\left(\frac{dr}{d\tau}\right)^2 + 1 - \frac{R_S}{r} = \gamma_\infty^2. \tag{394}$$

Next, from equation (346), the ratio of time intervals for the distant observer and infalling observer is

$$\frac{d\bar{t}}{d\tau} = \frac{\gamma_\infty}{(1 - R_S/r)} = \frac{\sqrt{1 - R_S/a}}{(1 - R_S/r)} \tag{395}$$

where we have evaluated $\gamma_\infty = \sqrt{1 - R_S/a}$ by using equation (394) at the starting point of the orbit $r = a, dr/d\tau = 0$. We denote the solution of this equation by $\bar{t}(\tau)$.

We are not done yet; $d\bar{t}/d\tau$ is not quite the redshift factor. A photon emitted at time τ by the infalling clock at the surface of the sphere at radius $r(\tau)$, will not actually arrive until time $\bar{t}'$ by the clock of distant observer located at $\bar{r}'$, because there is an *additional* light transit time. The time $\bar{t}'$ is given by:

$$\bar{t}' = \bar{t}(\tau) + \frac{1}{c}\int_{r(\tau)}^{\bar{r}'} \frac{du}{1 - R_S/u}. \tag{396}$$

We must not neglect to add in this time delay along a Schwarzschild null geodesic *after* the photon is emitted, because the transit time is not just an additive constant, it changes from

one photon to the next. The desired redshift z is then defined by

$$1+z \equiv \frac{d\bar{t}'}{d\tau} = \frac{d\bar{t}}{d\tau} - \frac{dr/d\tau}{c(1-R_S/r)} = \frac{\sqrt{1-R_S/a} - (1/c)dr/d\tau}{(1-R_S/r)}. \tag{397}$$

We next use $dr/d\tau = -c\sqrt{(R_S/a)(a/r-1)}$ from (394); note the inward minus sign choice. Using this expression in the above, we obtain

$$1+z = (1-R_S/r)^{-1}\left(\sqrt{1-R_S/a} + \sqrt{(R_S/a)\,(a/r-1)}\right). \tag{398}$$

This is our formal redshift equation.

It is interesting to calculate the explicit time dependence of z as $r \to R_S$. If, as we shall assume, $a \gg R_S$, then

$$1+z \to \frac{2R_S}{r-R_S}, \quad \text{and} \quad dr/d\tau \to -c \qquad (\text{as } r \to R_S) \tag{399}$$

In this limit, the dependence of r on $\bar{t}'$ is simple, given by:

$$\frac{d\bar{t}'}{dr} = \frac{d\bar{t}'}{d\tau}\frac{d\tau}{dr} = (1+z) \times -\frac{1}{c} = -\frac{2R_S}{c(r-R_S)}. \tag{400}$$

Solving this simple ordinary differential equation, we have

$$r - R_S = \text{constant} \times \exp\left(-c\bar{t}'/2R_S\right), \quad r \to R_S. \tag{401}$$

Combining (399) with (401), we find that the redshift z explodes exponentially as $r \to R_S$:

$$z \propto \exp\left(c\bar{t}'/2R_S\right), \quad \text{as } r \to R_S. \tag{402}$$

The qualitative history of gravitational collapse unfolds as follows, once we are past the very initial stages, so that $R_S \ll r \ll a$. The small measured redshift z then grows as $\sqrt{R_S/r}$. This is simply the first order in v/c effect, in essence a *Newtonian* Doppler shift. But once the Schwarzschild radius R_S is approached, matters change abruptly. Then, in just a few light crossing times ($\sim R_S/c$), the redshift erupts exponentially. Using numbers for the mass and radius of the Sun, the total free fall time would be about 20 minutes, while the final exponential spurt is measured in tens of microseconds. From the point of view of a distant observer, the collapsing object suddenly just disappears, all emergent radiation explosively redshifted out of the bandpass of any detector!

6.10 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 either Mercury or Venus, and returns. One does the experiment when the inner planet 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 the planet 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 (352), using the "ordinary" time parameter t for an observer at infinity, with $\gamma_\infty \to \infty$ for radio waves:

$$\left(\frac{dr}{dt}\right)^2 + \frac{B^3 J^2}{r^2} = B^2 c^2 \tag{403}$$

It is convenient here to evaluate the constant J in terms of r_0, the point of closest approach to the Sun. With $dr/dt = 0$, we find

$$J^2 = \frac{r_0^2 c^2}{B_0} \tag{404}$$

where $B_0 \equiv B(r_0)$. Our 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

$$t(r, r_0) = \frac{1}{c}\int_{r_0}^{r} \frac{dr}{B\left(1 - \dfrac{B}{B_0}\dfrac{r_0^2}{r^2}\right)^{1/2}}. \tag{405}$$

Expanding to first order in R_S/r_0 with $B(r) = 1 - R_S/r$:

$$1 - \frac{B}{B_0}\frac{r_0^2}{r^2} \simeq 1 - \left[1 + R_S\left(\frac{1}{r_0} - \frac{1}{r}\right)\right]\frac{r_0^2}{r^2}. \tag{406}$$

This may be rewritten as:

$$1 - \frac{B}{B_0}\frac{r_0^2}{r^2} \simeq \left(1 - \frac{r_0^2}{r^2}\right)\left(1 - \frac{R_S r_0}{r(r + r_0)}\right) \tag{407}$$

Using this in our time integral for $t(r, r_0)$ and expanding,

$$t(r, r_0) = \frac{1}{c}\int_{r_0}^{r} dr\, \left(1 - \frac{r_0^2}{r^2}\right)^{-1/2}\left(1 + \frac{R_S}{r} + \frac{R_S r_0}{2r(r + r_0)}\right) \tag{408}$$

The required integrals are

$$\frac{1}{c}\int_{r_0}^{r} \frac{r\, dr}{(r^2 - r_0^2)^{1/2}} = \frac{1}{c}(r^2 - r_0^2)^{1/2} \tag{409}$$

$$\frac{R_S}{c}\int_{r_0}^{r} \frac{dr}{(r^2 - r_0^2)^{1/2}} = \frac{R_S}{c}\cosh^{-1}\left(\frac{r}{r_0}\right) = \frac{R_S}{c}\ln\left(\frac{r}{r_0} + \sqrt{\frac{r^2}{r_0^2} - 1}\right) \tag{410}$$

$$\frac{R_S r_0}{2c}\int_{r_0}^{r} \frac{dr}{(r + r_0)(r^2 - r_0^2)^{1/2}} = \frac{R_S}{2c}\sqrt{\frac{r - r_0}{r + r_0}} \tag{411}$$

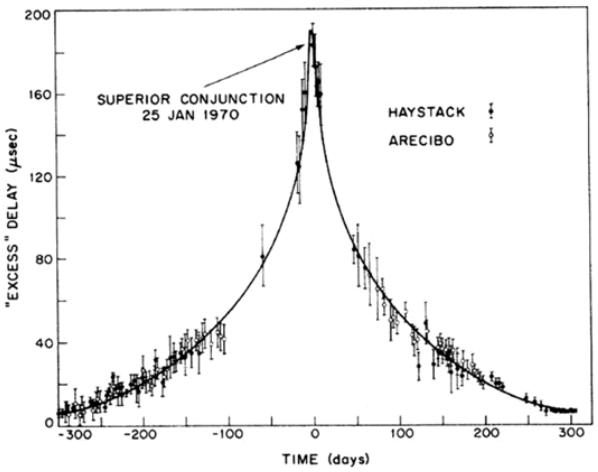


Figure 6: Radar echo delay from Venus as a function of time, fit with general relativistic prediction.

Thus,

$$t(r, r_0) = \frac{1}{c}(r^2 - r_0^2)^{1/2} + \frac{R_S}{c} \ln\left(\frac{r}{r_0} + \sqrt{\frac{r^2}{r_0^2} - 1}\right) + \frac{R_S}{2c}\sqrt{\frac{r - r_0}{r + r_0}} \tag{412}$$

The first term on the right is the leading order Newtonian result, the next two terms provide corrections of order R_S/c. 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 selection of the $\pm$ sign is dicatated by whether the signal passes through r_0 while enroute to the planet, i.e. on whether the planet is on the far or the near side of the Sun.

It may seem straightforward to plug in values appropriate to the Earth's and the planet's radial location, 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 would 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 (412) with $r = r(t)$. Figure (6) shows such a fit near the passage of superior conjunction (i.e. the planet on the distant side of its orbit relative to Earth, near the Sun seen in projection), in excellent agreement with theory. There is further discussion of this in W72 pp. 202–207, and an abundance of very topical information on the internet under "Shapiro delay." In 2003, it was reported that radio communication links with the Cassini spacecraft, on its voyage to Saturn, established the Shapiro delay with an agreement with general relativity to an accuracy of 0.0012%.

Modern applications of the Shapiro delay will use pulsars as signal probes, not terrestrial radar equipment. In this modern application, the time passage of a signal is altered by nothing so crude and dramatic as an entire star, but by the presence of low frequency gravitational waves sprinkled through the galaxy. Such is the accuracy of contemporary time measurement.

Black holes have no hair.

— John A. Wheeler

7 Kerr Black Holes

This chapter is off-syllabus and non-examinable. At present it is incomplete, a work in progress. Input from interested students is welcome; no obligation implied.

7.1 Rotating spacetime: Kerr-Schild metric

The remarkably prompt 1915 discovery of the Schwarzschild black hole solution was possible because one-dimensional spherical symmetry reduced the mathematical complexity of Einstein's field equations to a level at which the nonlinear complications became tractable. This naturally raised the question of whether there might be nonspherical but axisymmetric vacuum solutions as well, as one might perhaps expect for spacetimes exterior to an oblate rotating star. But the combination of the nonlinearity of the partial differential equations together with the absence of spherical symmetry made this a problem of ferocious difficulty.

It took nearly half a century, but in 1963, the mathematician Roy Kerr discovered precisely such a solution. Following Kerr's approach, we will introduce the Kerr metric using what are now known as Kerr-Schild coordinates, which are quasi-Cartesian. We begin with the standard equation for an oblate spheroid in familiar Euclidian x, y, z coordinates:

$$\frac{x^2+y^2}{r^2+a^2}+\frac{z^2}{r^2}=1. \tag{413}$$

Here, a is an arbitrary oblateness parameter with dimensions of length, and is to be regarded as a fixed constant. We treat r, however, as being defined by the equation itself, i.e., r is a function of the x, y, z coordinates and the oblateness constant a. As $a \to 0$, or $r \to \infty$, r becomes the standard Euclidian radius with corrections of order a^2/r^2. With these variables, the Kerr-Schild metric is just of the form (329), which we first encountered in our study of the Schwarzschild metric, but with the Kerr time coordinate differential dx^0 given by a slight modification of equation (327):

$$d\bar{x}^0 = c d\bar{t} = c dt - \frac{2GM/rc^2}{1-2GM/rc^2+a^2/r^2}dr. \tag{414}$$

The Kerr metric then becomes:

$$-c^2 d\tau^2 = (\eta_{\mu\nu}+\lambda l_\mu l_\nu)d\bar{x}^\mu d\bar{x}^\nu = \eta_{\mu\nu}d\bar{x}^\mu d\bar{x}^\nu + \lambda(l_\mu dx^\mu)^2, \tag{415}$$

now with

$$\lambda = \frac{2GM/rc^2}{1+a^2z^2/r^4}, \tag{416}$$

and l^μ is a Minkowski null vector $\eta_{\mu\nu}l^\mu l^\nu = 0$. By convention, $\eta_{\mu\nu}$ is used to raise and lower the l-indices, and $d\bar{x}^i = dx^i$.

We define the l^μ components with the aid of the following two-dimensional Cartesian vectors:

$$\boldsymbol{\sigma} = (r, a), \quad \boldsymbol{\sigma}_\perp = (-a, r), \quad \boldsymbol{R} = \frac{(x, y)}{\sqrt{r^2 + a^2}}. \tag{417}$$

Then, with orthogonal unit vectors $\hat{\boldsymbol{\sigma}}$, $\hat{\boldsymbol{\sigma}}_\perp$, l^μ is given by:

$$l^\mu = (1,\ \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{R},\ \hat{\boldsymbol{\sigma}}_\perp \cdot \boldsymbol{R},\ z/r), \tag{418}$$

where the two-dimensional spatial dot product is understood to be Euclidian. Combining equations (415)–(418) leads to:

$$-c^2 d\tau^2 = -c^2 d\bar{t}^2 + dx^2 + dy^2 + dz^2 + \frac{2GM/rc^2}{1 + a^2 z^2/r^4} \left[-cd\bar{t} + \frac{(rx + ay)dx + (ry - ax)dy}{r^2 + a^2} + \frac{z dz}{r} \right]^2 \tag{419}$$

This is certainly a remarkable generalisation of the Schwarzshild solution! To this day, there is no known elegant or enlightening derivation. Roy Kerr had no guidance. Nevertheless, he persisted. The Kerr metric components simply satisfy the equation $R_{\mu\nu} = 0$. The straightforward but lengthy task of verifying this we shall leave to the ambitious reader.

While the Kerr solution has proven to be an achievement of surpassing astrophysical importance, at the time it was discovered, the metric was not immediately appreciated as a singular black hole solution. In common with his predecessor Schwarzschild, Kerr assumed that his was a special vacuum solution that applied *outside* of a star—a rotating star certainly, but a normal, if perhaps idealised, star. An actual rotating star would undoubtedly have a messy ensemble of higher moments associated with its exerternal gravitational field, so for some time Kerr's "unthethered" solution was viewed as not much more than a fascinating curiosity. Unlike the Schwarzschild metric, which is an exact solution for the geometry outside a spherical mass distribution, the Kerr geometry is *not* the exact external solution outside of a star with finite angular momentum[13]. Moreover, in 1963, the notion that a *point-like* spacetime singularity could exist and harbour angular momentum as well as mass was far from understood. It took until the end of the decade, through the work of Brandon Carter, Penrose and others, before a profound milestone of general relativity was established: the Kerr metric represents *the* most general form of a rotating, electrically neutral, black hole to be found in nature. No stationary, stable modification of the axisymmetric Kerr vacuum metric is possible in conventional general relativity. In John Wheeler's memorable aphorism, "Black holes have no hair."

Schwarzschild versus Kerr: how important is rotation for black holes in a practical astronomical sense? The current astrophysical orthodoxy is that most, perhaps all, massive black holes in the centres of galaxies will have grown by copious accretion from a rotationally-supported, surrounding gaseous disc. In the process, the central hole will have also acquired a significant amount of angular momentum from the disc's inner edge. If so, for these objects, the finite angular momentum Kerr metric will be no small perturbation upon a classical Schwarzschild solution. This class of Kerr black hole must be treated on a proper mathematical footing right from the start. On the other hand, smaller black holes that have grown from the core of an imploding star, or perhaps even larger holes that may have grown from non disc-like accretion, may be well-represented by the low angular momentum limit of the Kerr metric, Schwarzschild's 1916 solution. In general, however, astrophysical black holes should be regarded as Kerr black holes.

[13] In solar system tests of general relativity, the Sun's angular momentum contributions to the spacetime geometry are far too small to be measured.

7.2 Boyer-Lindquist coordinates

For astrophysical applications, the most common form of the Kerr metric makes use of the quasi-spherical r, θ, ϕ *Boyer-Lindquist* coordinates:

$$-c^2 d\tau^2 = -\left(1 - \frac{2GMr}{\rho^2 c^2}\right) c^2 dt^2 - \frac{4GMar\sin^2\theta}{\rho^2 c^2} c dt\, d\phi + \frac{\rho^2}{\Delta} dr^2 + \rho^2 d\theta^2 + \left(r^2 + a^2 + \frac{2GMra^2\sin^2\theta}{\rho^2 c^2}\right)\sin^2\theta d\phi^2 \tag{420}$$

where

$$\rho^2 = r^2 + a^2\cos^2\theta, \qquad \Delta = r^2 - \frac{2GMr}{c^2} + a^2 \tag{421}$$

Note that we have returned to "distant observer standard time" t in (420), rather than $\bar{t}$. To relate the spatial Boyer-Lindquist coordinates to their Kerr-Schild counterparts, we first introduce the auxilliary variables

$$d\phi' = d\phi - (a/\Delta)dr, \qquad \hat{\boldsymbol{r}}' = (\sin\theta\cos\phi', \sin\theta\sin\phi', \cos\theta). \tag{422}$$

Then

$$x = \boldsymbol{\sigma}\cdot\hat{\boldsymbol{r}}', \quad y = \boldsymbol{\sigma}_\perp\cdot\hat{\boldsymbol{r}}', \quad z = r\cos\theta. \tag{423}$$

The dot products here are Euclidian, and x, y, z, r obey equation (413). In the limit of vanishing mass $M \to 0$, the metric (420) becomes

$$-c^2 d\tau^2 = -c^2 dt^2 + \frac{\rho^2}{r^2 + a^2} dr^2 + \rho^2 d\theta^2 + \left(r^2 + a^2\right)\sin^2\theta d\phi^2. \tag{424}$$

This is simply Minkowski space in oblate spheroidal coordinates, whose Cartesian counterparts obey equation (413) and are given by:

$$x = \sqrt{r^2 + a^2}\sin\theta\cos\phi, \quad y = \sqrt{r^2 + a^2}\sin\theta\sin\phi, \quad z = r\cos\theta. \tag{425}$$

(The $M \to 0$ limit taken directly from equations (422) and (423) requires an additional rotation of $\pi/2$ in the xy plane to bring the Cartesian axes into the above standard form, which one is free to do by axisymmetry.)

Apart from its evident complexity, the Kerr metric's most striking feature is the appearance of the off-diagonal components

$$g_{\phi 0} = g_{0\phi} = -\frac{2GMar\sin^2\theta}{\rho^2 c^2}. \tag{426}$$

The presence of such terms results in a truly novel property of Kerr spacetime: the "dragging" of inertial frames in the same sense as the rotation of the black hole. To see this most clearly, consider the fate of a test particle dropped into the hole, following a geodesic path with an initial 4-velocity component $U_\phi = 0$, i.e. the path is initially radial. In Problem Set 1, you showed that

$$\frac{dU_\rho}{d\tau} = \frac{1}{2}\frac{\partial g_{\mu\nu}}{\partial x^\rho} U^\mu U^\nu, \tag{427}$$

so that U_ϕ is a constant of the motion: the Kerr metric is axisymmetric. This, of course, is just the covariant angular momentum constant. Thus, following the geodesic, U_ϕ will remain

zero. However, in the presence of off-diagonal metric components, we cannot conclude that $d\phi/dt = 0$! Indeed, setting $U_\phi = 0$ we have:

$$\frac{d\phi}{d\tau} = U^\phi = g^{\phi 0} U_0, \quad \frac{cdt}{d\tau} = U^0 = g^{00} U_0, \tag{428}$$

or

$$\frac{1}{c}\frac{d\phi}{dt} = \frac{g^{\phi 0}}{g^{00}}. \tag{429}$$

The components of the (inverse) contravariant metric $g^{\mu\nu}$ are not difficult to calculate. It is always possible to rewrite the $g_{\mu\nu}$ matrix in block form, either by elementary row and column interchanges, or more simply by ordering the coordinates t, ϕ, r, θ from the start. Then, only a symmetric 2×2 matrix for the 0- and ϕ-indexed metric components need be inverted:

$$\begin{pmatrix} g_{00} & g_{0\phi} \\ g_{\phi 0} & g_{\phi\phi} \end{pmatrix}^{-1} \equiv \begin{pmatrix} g^{00} & g^{0\phi} \\ g^{\phi 0} & g^{\phi\phi} \end{pmatrix} = \frac{1}{D}\begin{pmatrix} g_{\phi\phi} & -g_{\phi 0} \\ -g_{0\phi} & g_{00} \end{pmatrix} \tag{430}$$

where D is the determinant

$$D = g_{00}g_{\phi\phi} - g_{\phi 0}g_{0\phi} = g_{00}g_{\phi\phi} - (g_{\phi 0})^2. \tag{431}$$

The geodesic equation becomes

$$\frac{1}{c}\frac{d\phi}{dt} = \frac{g^{\phi 0}}{g^{00}} = -\frac{g_{\phi 0}}{g_{\phi\phi}} \equiv \frac{\omega_d}{c}. \tag{432}$$

Dropped inward on an initially radial trajectory, a test particle acquires an azimuthal angular velocity in the same sense as the black hole angular momentum. In other words, a freely falling inertial frame, with a fixed vanishing angular momentum, finds itself dragged along by the spacetime geometry and forced to follow the rotation of the hole at the angular velocity ω_d given above. More generally, there will be a range of allowed angular velocities at a given radius, a point we explore below. A rotating, inertial, zero angular momentum frame certainly has no Newtonian counterpart!

To complete the calculation for $g^{\mu\nu}$, we evaluate the inversion formulae explicitly. One then finds:

$$D = -\Delta \sin^2\theta, \quad g^{\phi\phi} = \frac{\Delta - a^2\sin^2\theta}{\Delta\rho^2\sin^2\theta}, \quad g^{0\phi} = g^{\phi 0} = -\frac{2GMar}{\rho^2 c^2 \Delta}, \tag{433}$$

and

$$g^{00} = -\frac{\rho^2(r^2+a^2) + 2GMra^2\sin^2\theta/c^2}{\Delta\rho^2\sin^2\theta}. \tag{434}$$

The g^{rr} and $g^{\theta\theta}$ components are just simple reciprocals:

$$g^{rr} = \frac{1}{g_{rr}} = \frac{\Delta}{\rho^2}, \quad g^{\theta\theta} = \frac{1}{g_{\theta\theta}} = \frac{1}{\rho^2}. \tag{435}$$

Exercise. In Problem Set 1, you learned how to calculate the spatial metric γ_{ij} from a general metric $g_{\mu\nu}$. Show that the spatial interval dl^2 for the Kerr spacetime is given by

$$dl^2 = \frac{\rho^2}{\Delta}dr^2 + \rho^2\, d\theta^2 + \frac{\Delta\sin^2\theta}{1 - 2GMr/\rho^2c^2}d\phi^2$$

Verify that in the limit $a \to 0$, the Schwarzschild spatial interval is recovered.

7.3 Kerr geometry

7.3.1 Singularities

The Schwarzschild geometry features a coordinate singularity associated with an event horizon at $r = R_S = 2GM/c^2$, and a true singularity at $r = 0$. Kerr geometry is considerably richer. A cursory inspection of equation (420) reveals that there is singular structure in the Kerr metric when either Δ or ρ vanishes. The vanishing of Δ reduces to $r = R_S$ when a is small, a Schwarzschild coordinate singularity, and this suggests a corresponding Kerr coordinate singularity. This plausible claim may be directly investigated by calculating the curvature scalar $R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}$ and checking for singular behaviour; there is none. (Why is not possible to use the Ricci scalar R?) By contrast, the $r = 0$ true Schwarzschild singularity becomes $\rho = 0$ for the Kerr metric, or

$$r^2 + a^2 \cos^2\theta = 0. \tag{436}$$

This requires both $r = 0$ *and* $\theta = \pi/2$. Using the results of §(7.2), this dual combination of requirements corresponds to

$$x^2 + y^2 = a^2, \quad z = 0, \tag{437}$$

which, as we have noted, is a circular ring, not a dimensionless point. This question of what lies interior to the ring, we shall return to below.

The infinite g_{rr} coordinate singularity condition $\Delta = 0$ occurs when

$$r^2 - 2rr_g + a^2 = 0, \tag{438}$$

where we have defined $r_g = GM/c^2$, the gravitational radius. We denote the solutions r_+ and r_-, where

$$r_\pm = r_g \pm \sqrt{r_g^2 - a^2}. \tag{439}$$

Evidently, the r_+ and r_- singularities are present if and only if $a < r_g$. The physical significance of this limit is discussed further in §11.2, where it is demonstrated that the angular momentum J of a Kerr black hole is given by

$$J = Mac. \tag{440}$$

By way of example, were the Sun to somehow collapse to a black hole whilst retaining all of its angular momentum ($J = 1.92 \times 10^{41}$ in MKS), the result would be $a \simeq 0.2r_g$, a number also typical for a very rapidly rotating neutron star. In what follows, we shall assume that $a < r_g$.

There is another coordinate singularity where $\rho^2 \equiv r^2 + a^2\cos^2\theta = 2rr_g$, corresponding to $g_{00} = 0$. This occurs at radial coordinate locations $r'_\pm$ given by:

$$r'_\pm = r_g \pm \sqrt{r_g^2 - a^2\cos^2\theta}, \tag{441}$$

which become equal to their $r_\pm$ counterparts at $\theta = 0$ and $\theta = \pi$. In Kerr spacetime, what was the unique $r = R_S$ Schwarzschild-geometry coordinate singularity $g_{00} = 1/g_{rr} = 0$, corresponding to both an event horizon and a surface of infinite redshift, splits into no less than *four* distinct radii. To understand the physical properties of these locations requires some care.

7.3.2 Stationary limit surface: $g_{00} = 0$

Let us begin by examining the behaviour of purely azimuthal, null geodesics in Kerr geometry. These satisfy

$$g_{00}c^2dt^2 + 2g_{0\phi}cdt\, d\phi + g_{\phi\phi}d\phi^2 = 0. \tag{442}$$

This is a quadratic equation for $d\phi/dt$, with solutions:

$$\frac{1}{c}\frac{d\phi}{dt} = \frac{|g_{0\phi}|}{g_{\phi\phi}} \pm \sqrt{\frac{g_{0\phi}^2}{g_{\phi\phi}^2} - \frac{g_{00}}{g_{\phi\phi}}}. \tag{443}$$

or

$$\frac{d\phi}{dt} = \omega_d \pm \sqrt{\omega_d^2 - \frac{c^2 g_{00}}{g_{\phi\phi}}}, \tag{444}$$

where equation (432) has been used. In the limit $g_{00}/g_{\phi\phi} \to 0$, the two solutions may be written

$$\frac{d\phi_+}{dt} = 2\omega_d, \quad \frac{d\phi_-}{dt} = c\frac{g_{00}}{2g_{0\phi}}, \tag{445}$$

The (+) solution corresponds to a photon launched near the surface $g_{00} = 0$, which then takes off with twice the angular velocity of the dragged inertial frame, moving in the same sense. But facing *into* the fierce headwind of the dragged inertial frame, the (−) photon solution is barely changing its ϕ coordinate just outside the surface. At the vanishing g_{00} surface itself, the (−) azimuthal-headed photon remains at fixed ϕ, and inside the surface, it is dragged forward in ϕ as g_{00} becomes positve. Within the $g_{00} = 0$ surface, any material particle, no matter how powerfully it is driven, must be dragged in the sense of the inertial frame rotation arising from the hole's angular momentum. Recall that the 4-velocity of what would be a "stationary" object (i.e., a vanishing 3-velocity) must satisfy $g_{00}(U^0)^2 = -c^2$, but this is not possible when g_{00} passes through zero. The g_{00} surface is therefore known as the *stationary limit.*

The vanishing of g_{00} also means that signals sent from this surface become infinitely redshifted, as seen by a distant observer. However, in stark contrast to the Schwarzschild counterpart, this is *not* an event horizon. In Kerr geometry, one may cross the $g_{00} = 0$ surface, rotate with the hole, come back across the surface from the opposite direction, and return to describe the adventure. In Schwarzschild geometry, the time to reach the $g_{00} = 0$ surface is (logarithmically) infinite as measured by a distant observer. By contrast, in Kerr geometry, the time to reach the vanishing g_{00} stationary surface measured by a distant observer is *finite,* because the relevant integrand in time is only square root unbounded, and the integral itself is therefore finite.

7.3.3 Event horizon: $g^{rr} = 0$

To understand the physical nature of the two surfaces associated with the singular denominator of g_{rr} in the Kerr metric, we must first sharpen our understanding of what actually defines an event horizon. The key idea may be grasped from figure (7), which shows the light cones in a spacetime diagram near the surface $r = R_S$ for a Schwarzschild black hole. Future time is upward. This diagram displays the defining feature of the event horizon: the light cones have collapsed onto its surface, and lie tangentially along it in the spacetime. A better way to express this is that the collpase light cones *define* the surface: the event horizon is a *global null surface*, along which any differential coordinate vector dx^μ, caught between the

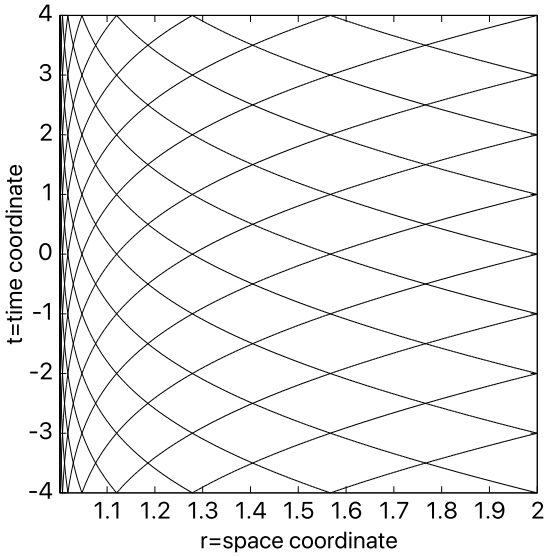


Figure 7: Spacetime structure of exterior null geodesics (light cones), near the surface $r = R_S = 1$ (in the displayed units) for a Schwarzschild black hole. The tangentially collapsed light cones define the surface by sheathing it. This is the hallmark of an event horizon. $r = R_S$ is a *null surface,* on which $dx^\mu dx_\mu = 0$.

vice grips of the light cone boundaries, must obey $dx^\mu dx_\mu = 0$. Now, we are free to define this same surface by an equation of the form $f(x^\mu) = 0$. Then, by definition, it must also be true that on these constant-f surfaces,

$$df \equiv dx^\mu \, \partial_\mu f = 0. \tag{446}$$

Since dx^μ is a null vector, so then must be $\partial_\mu f$, because f is a scalar function. We therefore take the definition of a null surface to be one that satisfies

$$(\partial^\mu f) \, \partial_\mu f = 0 \tag{447}$$

on $f(x^\mu) = 0$. In problems with a high degree of symmetry in which f is a function of only one coordinate, as is the case for the present problem with $f = f(r)$, the null surface condition becomes

$$g^{rr}(\partial_r f)^2 = 0. \tag{448}$$

This will in general require $g^{rr} = 0$, or equivalently for the Kerr metric, $g_{rr} \to \infty$.

We may now summarise the singular behaviour structure of the Kerr metric, shown schematically in fig. (8). The reader should be both cautious (as well as intrigued!) that a full understanding of the spacetime structure surrounding the Kerr singularity has yet to be achieved.

Moving radially inward in the equatorial plane from a region external to all singularities, we first cross the stationary limit surface $r = r'_+$ (labelled I in the diagram), a surface of infinite redshift from the point-of-view of an external observer. In the crossing, causal contact with the exterior is not lost. The next surface to be encountered, $r = r_+$, is an event horizon (II in the figure). Along the rotation axis of the black hole, the r'_+ and r_+ surfaces coincide, but in general there will be a well-defined gap between them. This gap-region is known as the *ergosphere*, because (as shown by Penrose) it is possible to extract energy and

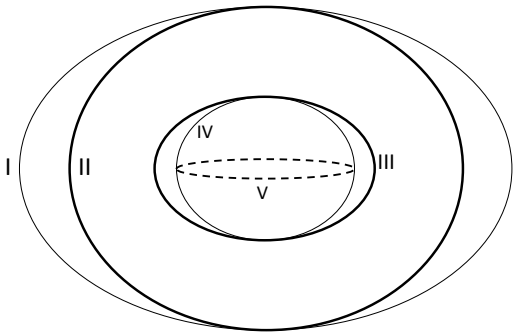


Figure 8: Schematic structure of the Kerr geometry. Surfaces II and III correspond to event horizons; I and IV to stationary surfaces. See text for further details.

angular momentum from the black hole by entering and leaving this domain. If a particle enters the ergosphere, splits in two, and only *one* of the decay products emerges, then it is possible for the emergent particle to have more energy coming out than the parent particle had going in. In other words, via this *Penrose process*, the mass M, and (it turns out), the angular momentum parameter a of the hole both decrease, with balance taken up by the particle that has reemerged from within the ergosphere.

Moving inward through surface II, we cross an event horizon, thereby entering a spacetime region that is causally cut-off from the exterior. This is similar to the $r < R_S$ region of a standard Schwarzschild hole, but here with the addition of frame-dragging effects. When a second event horizon is crossed at surface III, however, g_{00} becomes once again negative and timelike. This is, in principle, another ergosphere condition for the metric, but without the means to extract any energy from the hole: the event horizon passages have cut that off! Finally, passing through surface IV, the metric returns to the $g_{00} < 0$, $g_{rr} > 0$ sign convention it has in the asymptotically flat region lying exterior to the outer stationary r'_+ surface. Watch how you go, however: the equator of surface IV is a true spacetime singularity! Recall that this corresponds to the coordinate conditions $r = 0$ or $x^2 + y^2 = a^2$. Now, $x^2 + y^2 < a^2$ is certainly an allowed region of spacetime. But we have "used up" all of our positive r allocation in reaching the singularity! Remember, however, that r is simply a mathematical coordinate defined by equation (413); it is only asmptotically that we identify r with a spherical radius. Indeed, it may be shown that allowing r to become *negative* analytically continues the spacetime structure right up through $r \to -\infty$. Analytic continuity is not a guarantee of causal structure, and here we pay the price: when r is small and negative, $g_{\phi\phi} < 0$ (timelike!) and $g_{00} > 0$ (spacelike). With the spacelike variables $dt = dr = d\theta = 0$, $\theta = \pi/2$, and r small and negative,

$$c^2 d\tau^2 = \left(\frac{2GMa^2}{|r|c^2} \right) d\phi^2 \tag{449}$$

In other words, the proper time interval consists of *closed* timelike curves as ϕ advances from 0 to 2π, two otherwise identical points in the spacetime manifold. Is this a time machine? Our intuition recoils, but our intuitive concept of time has been honed in the provincial, asymptotically flat regions of spacetime. Might nature really allow this? All one can say is that at present, we have no basis for excluding this type of mathematical structure. What

is almost certainly true is that whether we accept closed timelike curves or prove them impossible to sustain, we may expect to encounter new physics.

7.3.4 The angular velocity of the outer event horizon

As we have noted, the azimuthal null geodesics in §(7.3.2), bracket the allowed range of angular velocities for any material particle in the ergosphere. We have,

$$\omega_d = \frac{c|g_{0\phi}|}{g_{\phi\phi}} = \frac{2r_g acr}{(r^2 + a^2\cos^2\theta)(r^2 + a^2) + 2r_g r a^2 \sin^2\theta}, \tag{450}$$

$$\frac{g_{00}}{g_{\phi\phi}} = \frac{2r_g r - r^2 - a^2\cos^2\theta}{[(r^2 + a^2\cos^2\theta)(r^2 + a^2) + 2r_g r a^2 \sin^2\theta]\sin^2\theta} \tag{451}$$

At the horizon location $r = r_+$, we may substitute $2r_g r_+$ for $r_+^2 + a^2$, with the result that

$$\omega_d(r_+) = \frac{ac}{2r_+ r_g}, \quad \frac{g_{00}}{g_{\phi\phi}} = \frac{a^2}{4r_+^2 r_g}. \tag{452}$$

Note that both of these quantities are independent of θ. Moreover, according to equations (443) and (444) the + and − solutions for $d\phi/dt$ are one and the same: the horizon angular velocity is unique, and given by

$$\omega_H \equiv \omega_d(r_+) = \frac{ac}{2r_+ r_g}. \tag{453}$$

This is the maximum allowed angular velocity in the ergosphere. (Since it occurs on an otherwise inaccessible boundary, ω_H is more accurately a least upper bound.)

7.4 The Penrose Process

7.4.1 Energy and angular momentum transferral

Kerr black holes differ from their Schwarzschild counterparts by more than mathematical complexity. In 1969, Penrose showed that Kerr holes are "live" in a way that Schwarzschild holes are not: it is possible to extract useable energy from a rotating black hole in the process of accreting mass through the outer event horizon.

The argument is simple, but subtle. Consider a body approaching a Kerr black hole from an asymptotically flat spacetime, a large distance away, region A. The 4-momentum of the body here is $p_\mu^{(A)}$, using the covariant form to take advantage of the fact that p_0 is conserved along its geodesic orbit, since the Kerr metric is time-independent. Far from the hole, in a nearly Minkowskian spacetime, p_0 may be identified with $-E/c$, where E is the energy of the body. Whether or not the body actually passes through the ergosphere, when it completes its fly-by of the hole and approaches another asymptotically flat region C at large distances, $p_0^{(A)} = p_0^{(C)}$. (The same equality is true is for the component p_ϕ, since the Kerr metric is axisymmetric.)

In the Penrose process, the body in question splits into two pieces while it is within the ergosphere, at a point denoted B. One of the daughter bodies ($B1$) is accreted by the hole,

the other ($B2$) continues on to region C. The schematic 4-momentum conservation equations for this problem may be written:

$$p_0^{(A)} = p_0^{(B)} = p_0^{(B1)} + p_0^{(B2)} = p_0^{(B1)} + p_0^{(C)}, \tag{454}$$

whence

$$E^{(C)} = E^{(A)} - E^{(B1)}, \tag{455}$$

from $E = -p_0 c$. At face value, this appears to show the very reasonable result that the energy of the body emerging in region C is less than the parent body of region A, due to the loss of the energy of the $B1$ mass which has been accreted by the hole. But this interpretation ignores the fact that in the ergosphere, g_{00} has flipped sign from negative to positive! This means that $-p_0^{(B1)} = E^{(B1)}$ is no longer a timelike energy component, but a spacelike momentum component, and as such its sign is not fixed. In particular, for some orbits $E^{(B1)}$ will be negative. When $E^{(B1)} < 0$, we find $E^{(C)} > E^{(A)}$: there is more energy in the final body than there was at the start of the orbit in the parent body! The orbits for which $E^{(B1)}$ is negative never allow the $B1$ mass to leave the ergosphere; in these cases it is always accreted by the hole. It is not difficult to demonstrate this behaviour by direct numerical solution of the Kerr geodesic equations.

Now, while $E^{(B1)}$ may be a negative quantity, the physical energy that an *in situ,* locally comoving inertial observer would measure must still be positive, by the laws of special relativity! In Boyer-Lindquist coordinates, the 4-velocity of such an observer, whom we shall locate for simplicity at fixed r and θ, would be

$$U^\mu = (cdt/d\tau, 0, 0, d\phi/d\tau) = U^0(c, 0, 0, \Omega) \tag{456}$$

where the U^r and U^θ components vanish, and $\Omega = d\phi/dt > 0$. The positive-energy requirement for this local observer is found by contracting U^μ with $p_\mu^{(B1)}$ and demanding that

$$U^\mu p_\mu^{(B1)} = U^0(cp_0^{(B1)} + \Omega p_\phi^{(B1)}) < 0, \tag{457}$$

(note the sign!), or

$$p_\phi^{(B1)} < -\frac{cp_0^{(B1)}}{\Omega} < -\frac{cp_0^{(B1)}}{\omega_H} \tag{458}$$

where in the last inequality we make use of the maximum property of ω_H. Since $cp_0^{(B1)} = -E^{(B1)} > 0$ for the Penrose process, the angular momentum $p_\phi^{(B1)}$ must be negative. The black hole loses both mass *and* angular momentum, which are then taken up by the escaping body which emerges in region C. Identifying $p_\phi^{(B1)}$ with incremental change δJ of the hole's angular momentum, and $-cp_0^{(B1)}$ with energy change $c^2\delta M$ of the hole, we may rewrite equation (458) as

$$\delta J < c^2 \frac{\delta M}{\omega_H}. \tag{459}$$

Note that both sides of the inequality are negative: the magnitude of δJ *exceeds* $|c^2\delta M/\omega_H|$.

7.4.2 Relation to the Hawking Area Theorem

In 1969, the same year that Penrose discovered his energy extraction process, Hawking showed that the area of a black hole event horizon can only increase, never decrease. (This

remarkable theorem is discussed in more detail in §11.x.x.) How does this relate to equation (459)? Might Hawking's theorem further restrict the domain of validity for angular momentum transfer given by equation (459)? In fact, as we now demonstrate, the requirement that the (outer) event horizon area increase in the Penrose Process is precisely equivalent to (459).

The event horizon area is defined with respect with constant t- and r-hypersurfaces, so that $dt = dr = 0$ in the Kerr metric area element. The event horizon area at $r = r_+$ is then given by

$$A_+ = \int_0^{2\pi} \int_0^{\pi} \sqrt{|g_{\phi\phi} g_{\theta\theta}|} d\theta \, d\phi, \tag{460}$$

where $r = r_+ = r_g + \sqrt{r_g^2 - a^2}$ in the metric coefficients. The integral is not difficult, and leads to

$$A_+ = \frac{8\pi G M}{c^2} r_+ = \frac{8\pi G^2}{c^4} \left(M^2 + \sqrt{M^4 - c^2 J^2 / G^2} \right). \tag{461}$$

Then, if M and J change by δM and δJ respectively, a direct calculation shows that $\delta A_+ > 0$ if and only if

$$\delta M - \frac{c^2 J \delta J}{2G^2 M \ (M^2 + \sqrt{M^2 - c^2 J^2 / G^2})} = \delta M - \frac{a \delta J}{2 c r_g r_+} > 0 \tag{462}$$

This is identical to the inequality (459).

7.5 Orbital equations

The orbital equations for Kerr spacetime may be derived analogously to their Schwarzschild counterparts. We shall treat only the case of orbits that lie entirely in the $\theta = \pi/2$ plane. Even with this restriction, it is important to emphasise that *unlike* the Schwarzschild metric, the Kerr metric is *not,* in general, a valid solution for the exterior geometry of an actual rotating star.

No additional techniques are needed for the general problem, but the resulting equations are considerably more complex. The $\theta = \pi/2$ Kerr metric coefficients we shall need are:

$$g_{00} = -(1 - 2r_g/r), \quad g_{0\phi} = g_{\phi 0} = -2r_g a / r, \quad g_{rr} = r^2/\Delta, \quad g_{\phi\phi} = r^2 + a^2 + 2r_g a^2 / r \tag{463}$$

with Δ unchanged from its (421) value, and ρ is now identical to r. For later reference, the corresponding inverse metric coefficients are

$$g^{00} = -\frac{1}{\Delta} \left(r^2 + a^2 + \frac{2r_g a^2}{r} \right), \ g^{0\phi} = g^{\phi 0} = -\frac{2r_g a}{r\Delta}, \ g^{rr} = \frac{\Delta}{r^2}, \ g^{\phi\phi} = \frac{1}{\Delta}(1 - \frac{2r_g}{r}). \tag{464}$$

Begin with the Lagrangian

$$\mathcal{L} = g_{00} c^2 \dot{t}^2 + 2 g_{0\phi} c \dot{t} \dot{\phi} + g_{rr} \dot{r}^2 + g_{\phi\phi} \dot{\phi}^2 \tag{465}$$

where the dot signifies $d/d\tau$. Since $\mathcal{L}$ is t-independent, the associated Lagrangian equation of motion is simply $\partial_i \mathcal{L} =$ constant, or

$$g_{00} c \dot{t} + g_{0\phi} \dot{\phi} = -\gamma_\infty c \tag{466}$$

where the right side is a constant chosen to remind us (much as in the Schwarzschild calculation) that γ_∞ is the special relativistic γ-factor if the orbit extends to $r \to \infty$. The ϕ equation is similar:

$$g_{0\phi} c\dot{t} + g_{\phi\phi}\dot{\phi} = \gamma_\infty J \tag{467}$$

where now J is an angular momentum constant. Notice that the left side of of equations (466) and (467) are identical to the covariant 4-vectors U_0 and U_ϕ respectively. That they are indeed constant also follows directly from the covariant form of the geodesic equation and the symmetry of $g_{\mu\nu}$. The solution to equations (466) and (467) for $\dot{t}$ and $\dot{\phi}$ is straightforward, and may be found most simply by implementing the inverse 2×2 matrix calculated in equation (430):

$$c\dot{t} = -\frac{\gamma_\infty}{D}(g_{\phi\phi} c + g_{0\phi} J) \tag{468}$$

$$\dot{\phi} = \frac{\gamma_\infty}{D}(g_{0\phi} c + g_{00} J) \tag{469}$$

where the determinant D is given by (431) and (433). Using the metric coefficients of equation (420) with $\theta = \pi/2$, this becomes

$$\dot{t} = \frac{\gamma_\infty}{\Delta}\left(r^2 + a^2 + \frac{a^2 R_S}{r} - \frac{aJR_S}{rc}\right) \tag{470}$$

$$\dot{\phi} = \frac{\gamma_\infty}{\Delta}\left[\frac{acR_S}{r} + J\left(1 - \frac{R_S}{r}\right)\right] \tag{471}$$

with

$$\Delta = r^2 + a^2 - rR_S. \tag{472}$$

The radial equation of motion may be found by setting $\mathcal{L} = -c^2$ in equation (465) (Why?). The result may be written in the following suggestive manner:

$$-c^2 = c\dot{t}(g_{00} c\dot{t} + g_{0\phi}\dot{\phi}) + \dot{\phi}(g_{0\phi} c\dot{t} + g_{\phi\phi}\dot{\phi}) + g_{rr}\dot{r}^2 \tag{473}$$

The first two groupings of terms are just constants of motion, equations (466) and (467). Hence:

$$-c^2 = -\gamma_\infty c^2 \dot{t} + \gamma_\infty J\dot{\phi} + \dot{r}^2 g_{rr}. \tag{474}$$

Using (470) and (471), and solving for $\dot{r}$ leads us to:

$$\left(\frac{dr}{d\tau}\right)^2 = c^2\left[(\gamma_\infty^2 - 1)\left(1 + \frac{a^2}{r^2}\right) + \frac{R_S}{r}\right] - \frac{\gamma_\infty^2}{r^2}\left[J^2 - \frac{R_S}{r}(J - ac)^2\right] \tag{475}$$

Further use of (470) and (471) allows us to solve for dr/dt and $dr/d\phi$, parameterising the orbit in terms of the time and azimuthal variables as measured by an observer at infinity:

$$\left(\frac{dr}{dt}\right)^2 = \frac{\Delta^2\{\ \ldots\ \}}{\gamma_\infty^2[r^2 + a^2 + (aR_S/cr)(ac - J)]^2}, \tag{476}$$

$$\left(\frac{dr}{d\phi}\right)^2 = \frac{\Delta^2\{\ \ldots\ \}}{\gamma_\infty^2[J + (R_S/r)(ac - J)]^2}, \tag{477}$$

where the notation $\{\ \ldots\ \}$ denotes the right hand side ofequation (475). The equations for photon orbits are obtained by taking the limit $\gamma_\infty \to \infty$.

7.6 Circular and quasi-circular orbits

The midplane Kerr orbital equations are more complex than their Schwarzschild counterparts, and even a comparatively simple problems like radial infall require a numerical solution. An interesting exception to this rule is the case of infall from circular orbits, trajectories we shall henceforth refer to as *quasi-circular orbits.* Such orbits, in both Schwarzschild and Kerr spacetimes, exhibit a behaviour very different from anything seen in Newtonian dynamics: close enough to the black hole, circular orbits become unstable. These orbits are of great astrophysical interest, as they are important for an understanding of black hole accretion. The resulting inspiralling orbit may be solved for in closed form, as we shall now show.

In our discussion of Schwarzschild orbits in §6.5.2, we noted that the circular 4-velocity components U^ϕ and U^0 generally satisfy the coupled equations

$$g_{\mu\nu}U^\mu U^\nu = -c^2, \quad \partial_r g_{\mu\nu}U^\mu U^\nu = 0. \tag{478}$$

The first of these follows from the definition of $d\tau^2$, the second is a direct consequence of the circular orbit geodesic equation $dU_r/d\tau = 0$. Using the $\theta = \pi/2$ Kerr metric $g_{\mu\nu}$ components, and assuming that only U^ϕ and U^0 are present, the solution emerging from these two equations is

$$U^\phi = \left(\frac{GM/r^3}{1 - 3r_g/r + 2ar_g^{1/2}/r^{3/2}}\right)^{1/2}, \quad U^0 = (a + r^{3/2}/r_g^{1/2})U^\phi, \tag{479}$$

where as usual r_g is the gravitational radius GM/c^2. The sign convention is that U^ϕ is always positive for an undisturbed circular orbit; if the hole is rotating in the opposite sense, then $a < 0$. The Newtonian angular velocity limit of U^ϕ is clear. The covariant vectors U_0 and U_ϕ, which correspond to the conserved energy (minus thereof) and angular momentum respectively, may be calculated as usual by contracting $g_{\mu\nu}$ with their contravariant counterparts. We then find

$$U_\phi = \frac{c(r_g/r)^{1/2}(r^2 + a^2 - 2a(r_g r)^{1/2})}{(r^2 - 3r_g r + 2a(r_g r)^{1/2})^{1/2}}, \quad U_0 = -\frac{c(r^2 - 2r_g r + a(r_g r)^{1/2})}{r(r^2 - 3r_g r + 2a(r_g r)^{1/2})^{1/2}} \tag{480}$$

The Schwarzschild results are recovered as $a \to 0$.

A key feature of axisymmetric rotational motion of any type—not just orbital dynamics, but laboratory fluids as well—is that if the specific angular momentum decreases radially outward, the motion is unstable: small disturbances from circularity grow. First shown by Lord Rayleigh, to prove this rigorously would take us a bit far afield, so instead we offer the following intuitive argument. If the angular momentum of circular orbits is decreasing outward, then a small outward radial displacement of a test mass would place it in an orbit for which its (conserved) angular momentum is too high. The excess angular momentum centrifugally forces the test mass to a yet more distant orbit, where the mismatch is yet worse, and the motion is unstable. The same reasoning applies equally well for an inward displacement. By contrast, an increasing-outward angular momentum profile will always return a displaced test mass toward the direction of the orginal orbit, and is therefore stable. We seek, therefore, the point at which $\partial_r U_\phi = 0$. After a straightforward but lengthy calculation, it may be shown that the vanishing angular momentum gradient condition is

$$0 = \frac{\partial U_\phi}{\partial r} = \frac{cr_g^{1/2}}{2r^{1/2}\mathcal{D}^3}\left(1 + \frac{ar_g^{1/2}}{r^{3/2}}\right)\left[1 - \frac{6r_g}{r} - \frac{3a^2}{r^2} + \frac{8ar_g^{1/2}}{r^{3/2}}\right], \tag{481}$$

where $\mathcal{D}^2 = 1 - 3r_g/r + 2ar_g^{1/2}/r^{3/2}$. The U_ϕ gradient will vanish only when the final factor is square brackets is zero. This radial location, denoted by r_I, corresponds to the innermost stable circular orbit, or ISCO. It satisfies the equation:

$$\begin{aligned} r_I^2 &= 6r_g r_I + 3a^2 - 8a(r_g r_I)^{1/2} \\ &= \frac{2r_g r_I}{3} + \frac{16 r_g r_I}{3}\left(1 - \frac{3a}{4(r_g r_I)^{1/2}}\right)^2 \end{aligned} \tag{482}$$

The second form of this equation will be especially convenient for the development to follow. For $-1 \le a/r_g \le 1$, the solution for the ISCO location ranges from $9 \ge r_I/r_g \ge 1$, decreasing monotonically with a. (The Schwarzschild $a = 0$, $r_I/r_g = 6$ solution is easily recovered.)

We next evaluate U_ϕ and U_0 at $r = r_I$, hereafter $U_\phi^{(I)}$ and $U_0^{(I)}$. Substituting for the resulting r_I^2 terms wherever they appear in (480) by using equation (482), we find

$$\begin{aligned} U_\phi^{(I)} &= \frac{c(r_g/r_I)^{1/2}(4a^2 + 6r_g r_I - 10a(r_g r_I)^{1/2})}{(3a^2 + 3r_g r - 6a(r_g r)^{1/2})^{1/2}} \\ &= 2\sqrt{3} c r_g \left(1 - \frac{2a}{3(r_g r_I)^{1/2}}\right), \end{aligned} \tag{483}$$

and

$$\begin{aligned} U_0^{(I)} &= -\frac{c(3a^2 + 4r_g r_I - 7a(r_g r_I)^{1/2})}{r_I(3a^2 + 3r_g r_I - 6a(r_g r)^{1/2})^{1/2}}, \\ &= -\frac{4c}{3}\left(\frac{3r_g}{r_I}\right)^{1/2}\left(1 - \frac{3a}{4(r_g r_I)^{1/2}}\right), \\ &= -c\left(1 - \frac{2r_g}{3r_I}\right)^{1/2} \end{aligned} \tag{484}$$

In each of the two preceding cases, the second line follows from the first after factoring the a^2 polynomials in the numerator and denominator and cancelling common terms. The final equality of (484) follows directly from the second equality of equation (482).

The task now before us is to find orbits that have a radial component U^r of the 4-velocity, but retain the conserved angular momentum and energy constants $U_\phi^{(I)}$ and $U_0^{(I)}$ of the ISCO. The condition $U^\mu U_\mu = -c^2$ may be written

$$g_{rr}(U^r)^2 + U_\phi^{(I)}(g^{\phi\phi}U_\phi^{(I)} + g^{\phi 0}U_0^{(I)}) + U_0^{(I)}(g^{00}U_0^{(I)} + g^{0\phi}U_\phi^{(I)}) = -c^2 \tag{485}$$

Remember, the $U_\mu^{(I)}$ covariant 4-velocities are constants of the motion; the r-dependence of this equation arises entirely from the metric coefficients. Substituting explicitly for the latter, our equation takes the form:

$$\begin{aligned} (U^r)^2 &+ \frac{U_\phi^{(I)}}{r^2}\left[\left(1 - \frac{2r_g}{r}\right)U_\phi^{(I)} - \frac{2r_g a U_0^{(I)}}{r}\right] \\ &- \frac{U_0^{(I)}}{r^2}\left[U_0^{(I)}\left(r^2 + a^2 + \frac{2r_g a^2}{r}\right) + \frac{2r_g a U_\phi^{(I)}}{r}\right] = -c^2\left(1 - \frac{2r_g}{r} + \frac{a^2}{r^2}\right). \end{aligned} \tag{486}$$

Don't be put off by the unwieldy appearance of this equation! The development is about to take a dramatic simplification.

Equation (486) is of the form

$$(U^r)^2 = -2V_{eff}(r) \tag{487}$$

where $V_{eff}(r)$ is an effective potential function. Consider first the simple case of Newtonian dynamics. Then for a hypothetical orbit that had the energy and angular momentum of some circular orbit of radius r_c, the function $V_{eff}(r)$ would be a simple quadratic polynomial in $1/r$. In particular, with specific angular momentum $J = \sqrt{GMr_c}$ and total energy $E = -GM/2r_c$, we have

$$V_{eff}(r) = -E + \frac{J^2}{2r^2} - \frac{GM}{r} = \frac{GM}{2r_c}\left(1 - \frac{r_c}{r}\right)^2, \tag{488}$$

a perfect square. This is very artificial because it has only the one real solution, $r = r_c$ and $U^r = 0$, but it illustrates the following point. An actual noncircular bound orbit would have two V_{eff} roots, corresponding to the radii of most distant and closest approach. In the limit of a circular orbit, these roots coalesce, and only one radial location is possible for the relevant circular energy and angular momentum. In other words, both the first and second radial derivatives of V_{eff} vanish for a circular orbit at $r = r_c$.

Moving to the relativistic problem, V_{eff} is cubic in $1/r$, but there is an additional simplifying constraint: the ISCO radius r_I is by definition a *marginally stable* circular orbit. That means the *second* radial derivative of V_{eff} must also vanish: its constituent polynomial is a perfect cube in $1/r$. Therefore, once the relevant circular orbit energy and angular momentum are selected, V_{eff} must have the form

$$(U^r)^2 = -2V_{eff}(r) = -2A\left(1 - \frac{r_I}{r}\right)^3, \tag{489}$$

where A is a constant to be determined. Taking the limit $r \to \infty$ in both equations (486) and (489), and using (484) we obtain:

$$(U^r)^2 \to -2A = (U_0^{(I)})^2 - c^2 = -\frac{2c^2 r_g}{3r_I}. \tag{490}$$

We finally arrive at our desired result:

$$(U^r)^2 = \frac{2c^2}{3}\frac{r_g}{r_I}\left(\frac{r_I}{r} - 1\right)^3. \tag{491}$$

In addition to its simplicity, note the generality of equation (491): it holds for any a-value in Kerr geometry. Moreover, radial motion is allowed even with circular orbit values for the energy and angular momentum! We have found the radial 4-velocity of any orbit spiralling in from the ISCO radius for any Kerr black hole.

The infalling branch of equation (491) for $r < r_I$ takes the form:

$$\frac{dr}{d\tau} = -c\left(\frac{2r_g}{3r_I}\right)^{1/2}\left(\frac{r_I}{r} - 1\right)^{3/2}. \tag{492}$$

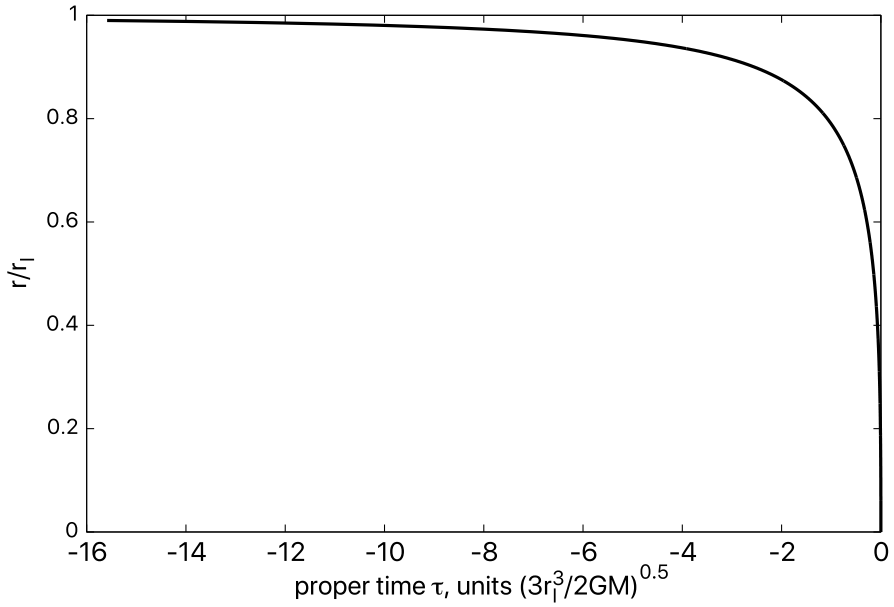


Figure 9: Radial coordinate r as a function of proper time τ for a test particle falling in from the marginally stable circular orbit at $r = r_I$. Normalised to r_I, the solution is valid for any Kerr geometry. The circular energy and angular momentum constants of the circular orbit are conserved in the radial infall.

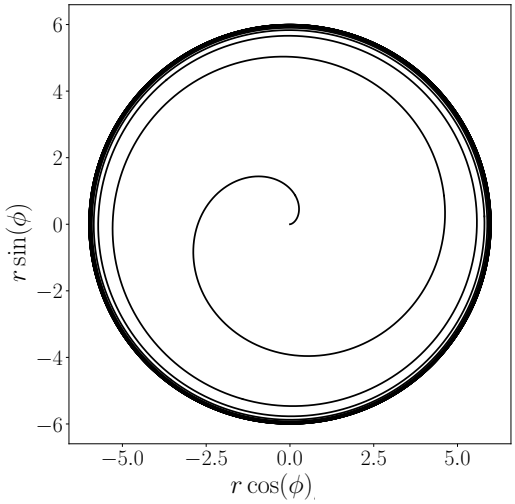


Figure 10: The infalling spiral orbit of equation (496).

With $x = r/r_I$, $x_g = r_g/r_I$, this may be written

$$\left(\frac{3r_I^3}{2r_g}\right)^{1/2} \int_0^x \left(\frac{x'}{1-x'}\right)^{3/2} dx' = -c\tau. \tag{493}$$

Notice that τ runs from $-\infty$ at $x = 1$, to $\tau = 0$ at $x = 0$. The integral may be done straightforwardly via the substitution $x' = \sin^2\theta'$, a problem we shall leave for the reader. The result is:

$$c\tau = \sqrt{\frac{3r_I^3}{2r_g}} \left(3\sin^{-1}\sqrt{x} - (x-3)\sqrt{\frac{x}{1-x}}\right) \tag{494}$$

The function $r(\tau)$ is shown in figure (9).

In the Schwarzschild limit $a = 0, r_I = 6r_g$, the orbit $r(\phi)$ is particular simple. Then, with $\dot{\phi} = 2\sqrt{3}cr_g/r^2$,

$$\frac{dr}{d\phi} = \frac{\dot{r}}{\dot{\phi}} = -\frac{1}{3\sqrt{2r_g r_I}} r^2 \left(\frac{r_I}{r} - 1\right)^{3/2} \tag{495}$$

Once again, this is an equation that is easily integrated (the substitution $u = 1/r$ is a good place to begin), with the result

$$r = \frac{6r_g}{1 + 12/\phi^2}, \tag{496}$$

where $-\infty < \phi \leq 0$. We emphasise once again that, despite its simplicity of form, this is an *exact*, fully relativistic solution of the Schwarzschild orbit equations. (See figure [10].)

Because of the fact that general relativity allows for noncircular orbits for values of the energy and angular momentum corresponding to a circular orbit, there is a symmetry that may be exploited which allows an entire class of explicit analytic solutions to the Kerr orbital equations to be constructed. These involve functions no more complicated than hyperbolic trigonometric, and exhibit some remarkable properties. Further exploration of these interesting orbits is left to the problems.

[T]he probable effect of deviations from the Fermi equation of state suggests that actual stellar matter after the exhaustion of thermonuclear energy sources will, if massive enough, contract indefinitely, although more and more slowly, never reaching true equilibrium.

— J. R. Oppenheimer & G. Volkoff

8 Self-Gravitating Relativistic Hydrostatic Equilibrium

8.1 Historical Introduction

"'[T]he glorious sun stays in his course and plays the alchemist..." wrote Shakespeare[14] more than three centuries in advance of our understanding of thermonuclear fusion. The alchemical course of the lives of the vast majority of stars, from nuclear buring to a remnant core known as "white dwarf," runs to completion without any practical influence of general relativity. However, as we shall see, white dwarfs obey a strict upper mass limit of some $1.4M_\odot$, the exact number depending weakly upon the star's composition. While fewer in number than their low mass counterparts, more massive stars have a very different life history. For members of this stellar population, fate has decreed a much larger, and much more quite dramatic, role for general relativity.

A natural point with which to begin our discussion of relativistic stars is the 1932 discovery of the second major denizen of the atomic nucleus by James Chadwick: the neutron, aptly named for its electrical charge neutrality, with a mass nearly identical to that of the proton. By the 1930's, it was well-understood that white dwarf stars, with a mass $\sim 1M_\odot$ but a radius no larger than that of the Earth, were able to support themselves against gravitational collapse by the pressure of degenerate electrons. What might happen to a star once its nuclear fuel had been spent, if its core mass exceeded $1.4M_\odot$, was completely unknown. Shortly after of the discovery of the neutron, the observational astronomers Baade and Zwicky proposed that the degeneracy pressure of a much more massive particle, a pressure much larger than that provided by electrons, would be available to support a hypothetical star composed entirely of neutrons. They went on to propose, with astonishing prescience, that the transition of a normal star into a "neutron star" was the physical process responsible for the production of supernovae, and that these unusual stars could form as the remnant of a supernova event. An order-of-magnitude estimate obtained by equating the standard Fermi momentum with $m_n c$, the mass of a neutron times the speed of light, allowed one to deduce that a neutron star of a few solar masses would have a radius of order 10 km, far smaller than the planetary scale of 10^4 km typical of a white dwarf. (The difference in radial stems directly from the ratio of the electron to neutron mass.) The density of such an extraordinary star would be comparable to that of an atomic nucleus: some 10^{17} kg per cubic metre. The bold and brilliant speculations of Baade and Zwicky did not go unnoticed —both were very high profile researchers—but they were generally met with skepticism in

[14] King John, Act 1, Scene 3.

the astronomical community. That skepticism persisted into the late 1960's when it at last crumbled, faced with the stunning discovery of what are now known as "pulsars," which were convincingly identified with rotating neutron stars.

Theoretical physicists were more willing to engage with speculative ideas early on. Even before Baade and Zwicky, the Soviet theorist Lev Landau had boldly proposed that *all* ordinary stars contain a neutron core in the form of a macroscopic nucleus as their energy source, with energy liberated via the break up of atomic nuclei as they accrete onto this degenerate core! Even though a hydrogen fusion source for stellar energy had been (correctly) proposed by Eddington as early as 1920, this was considered to be far from certain in the early 1930's, and physicists were interested in Landau's work. Do the ideas hold together? Are the permitted masses of neutron cores in the right range for this to be viable? Just how small could a self-consistent neutron core be? Conversely, is there an upper limit to its mass, analogous to the case of a white dwarf star? To answer these questions properly, however, would require the full strong gravity machinery of general relativity.

In 1939, J. Robert Oppenheimer and his student G. Volkoff developed the underlying theoretical techniques necessary to address this problem. Their paper contains for the first time, the relativistic equation for the hydorstatic balance of stellar equilibrium. This work established the basis of modern neutron star theory in particular, and of general relativistic internal stellar dynamics more broadly. Because it built on preliminary mathematical groundwork laid down by R. C. Tolman, the equation is known today as the Tolman-Oppenheimer-Volkoff (TOV) equation.

Alas, Oppenheimer and Volkoff found that a fully degenerate, ideal gas of neutrons could not form a stable self-gravitating body with a mass in excess of 0.7 $M_\odot$, less even than the modest $1.4M_\odot$ upper mass limit of a white dwarf. It became quickly apparent, however, that in contrast to the degenerate electrons of a white dwarf, the strongly interacting nucleons of a neutron star were not well-modelled by an ideal Fermi gas. Using a more realistic (but still quite uncertain) equation of state thought to be appropriate at these extraordinarily high neutron densities, later research found that neutron star masses could exist with values closer to $2M_\odot$, or perhaps even slightly beyond. We should therefore expect to find neutron stars in the very narrow mass range between 1.4 $M_\odot$, the well-established white dwarf upper mass limit, and $\sim 2\ M_\odot$, our best guess for the neutron star maximum mass. In fact, modern astrophysics has brilliantly confirmed this very tight prediction.

This is not much of a reprieve for the high mass population of stars. What of stellar cores with masses in excess of $2M_\odot$? We know today that there essentially no hope: nothing can prevent their collapse to a black hole. Fittingly, it was Oppenheimer once again, also in 1939, in collaboration with Hartland Snyder, who published the first general relativistic calculation of the spherical gravitational collapse of ordinary matter to a pointlike singularity we would now call a black hole. Many regard Oppenheimer, who successfuly directed the Atomic Bomb Manhatten Project of WWII, as the true founding father of neutron star and black hole theory. Oppenheimer died in 1967 at the relatively young age of 62, just before the spectacular 1968 discovery of the Crab Nebula Pulsar, which provided the "smoking gun" evidence for the existence of neutron stars.

The physical radius of a neutron star is only just slightly larger than its Schwarzschild radius, so if neutron stars had now become real, black holes represented at most a modest leap of faith. Throughout Oppenheimer's lifetime, neutron stars and black holes had been mere theoretical constructs for most researchers —some would have said toys—and were largely ignored among astronomers. Starting in the 1950's, with a then small cadre of physicists led by John Wheeler and supported by Oppenheimer himself, there developed a growing imperative that gravitational collapse must be taken very seriously. This concerted effort les to a golden era of productivity in general relativity research not rivaled since the days of Einstein's early papers. On the observational front, we now know that neutron stars and

black holes are present in vast numbers throughout the Universe, and the research focus of much of present day high energy astrophysics. We have even recently witnessed, via the emergent electromagnetic and gravitational radiation, black holes merging and neutron stars colliding with another! This is truly a propitious time to be studying the general relativistic theory of self-gravitating stars.

8.2 Fundamentals

If you have not done the exercise on page (76), now is the moment to do so. In fact, we shall start our discussion with the results derived there. We write the field equations in "source function form"

$$R_{\mu\nu} = -\frac{8\pi G}{c^4} S_{\mu\nu} \equiv -\frac{8\pi G}{c^4}\left(T_{\mu\nu} - \frac{g_{\mu\nu}T}{2}\right), \tag{497}$$

with the stress tensor $T_{\mu\nu}$ that of an ideal static fluid:

$$T_{\mu\nu} = Pg_{\mu\nu} + (\rho + P/c^2)U_\mu U_\nu, \quad U_i = 0, \quad (i = 1, 2, 3).$$

The metric is spherically symmetric and static, which we represent by the usual $A(r), B(r)$ functions:

$$-c^2 d\tau^2 = -Bc^2 dt^2 + A dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2. \tag{498}$$

The sole component of a static 4-velocity is $U^t = c/\sqrt{B}$, and $U_t = -c\sqrt{B}$. The components of the Ricci tensor have been worked out in §5.4. The fundamental field equations take the source-function form:

$$R_{tt} = -\frac{B''}{2A} + \frac{B'}{4A}\left(\frac{B'}{B} + \frac{A'}{A}\right) - \frac{B'}{rA} = -\frac{8\pi G}{c^4} S_{tt} = -\frac{4\pi G B}{c^4}(3P + \rho c^2), \tag{499}$$

$$R_{rr} = \frac{B''}{2B} - \frac{B'}{4B}\left(\frac{B'}{B} + \frac{A'}{A}\right) - \frac{A'}{rA} = -\frac{8\pi G}{c^4} S_{rr} = \frac{4\pi G A}{c^4}(P - \rho c^2), \tag{500}$$

$$R_{\theta\theta} = -1 + \frac{1}{A} + \frac{r}{2A}\left(-\frac{A'}{A} + \frac{B'}{B}\right) = -\frac{8\pi G}{c^4} S_{\theta\theta} = \frac{4\pi G r^2}{c^4}(P - \rho c^2). \tag{501}$$

The prime $'$ stands for d/dr. We may eliminate B completely from our equations by forming the following combination:

$$\frac{R_{tt}}{B} + \frac{R_{rr}}{A} + \frac{2R_{\theta\theta}}{r^2} = -\frac{2A'}{rA^2} - \frac{2}{r^2} + \frac{2}{Ar^2} = -\frac{16\pi G\rho}{c^2}. \tag{502}$$

This may be simplified:

$$\left(\frac{r}{A}\right)' = 1 - \frac{8\pi G\rho r^2}{c^2}. \tag{503}$$

Integrating this equation from $r = 0$, the centre of our star, to some interior radius r,

$$\frac{1}{A} = 1 - \frac{2G\mathcal{M}(r)}{rc^2} + \left(\frac{C}{r}\right), \tag{504}$$

where C is an integration constant and $\mathcal{M}(r)$ is the mass-like quantity

$$\mathcal{M}(r) = 4\pi \int_0^r \rho r'^2 \, dr'. \tag{505}$$

(The r' notation is of course just a dummy integration variable, not a derivative!) Were C not to vanish, then A would have to vanish at $r = 0$, and the metric would be singular. In the absence of a black hole at the origin (which is what C allows for), this is not physical: with all the mass exterior to the origin and a finite pressure at the centre, the metric must be regular. Hence $C = 0$ and

$$A = \left(1 - \frac{2G\mathcal{M}(r)}{rc^2}\right)^{-1}. \tag{506}$$

Equation (505) certainly looks at first sight like a good definition for "mass within r", but remember that the space inside the star has curvature! Where, for example, is $A(r)$, the metric coefficient needed for proper radial intervals: $A^{1/2}dr$? When the r upper limit of the integral exceeds the outer radius R of the star, then $\mathcal{M}$ must be the constant M in the Schwarzschild solution. This "M" includes what we would think of as the Newtonian mass *plus* the effective contribution from the energy from all other sources, including the gravitational field itself! Interestingly, this naive Newtonian mass expression, with its compounded errors of neglecting curved space and the varied relativistic contributions, is correct. The errors, evidently, are offsetting.

Our next step is to return to equation (501), using the hydrostatic equilibrium equation for $B' = dB/dr$:

$$\frac{B'}{B} = -\frac{2P'}{\rho c^2 + P}. \tag{507}$$

Upon substitution for A and B in (501), this leads to:

$$\boxed{\frac{dP}{dr} = -\left(\frac{G\mathcal{M}(r)}{r^2} + \frac{4\pi GPr}{c^2}\right)\left(\rho + \frac{P}{c^2}\right)\left(1 - \frac{2G\mathcal{M}(r)}{rc^2}\right)^{-1}}. \tag{508}$$

This is known as the *Tolman-Oppenheimer-Volkoff,* or TOV, equation. We shall use the TOV equation, together with equation (505) in differential form,

$$\frac{d\mathcal{M}}{dr} = 4\pi\rho r^2 \tag{509}$$

as the governing coupled pair of equations for relativistic hydrostatic equilibrium.

Students of classical stellar structure should recognise the TOV equation as the equation of Newtonian stellar equilibrium,

$$\frac{dP}{dr} = -\frac{G\mathcal{M}(r)\rho}{r^2} \tag{510}$$

with some notable relativistic additions. We replace ρ by $\rho + P/c^2$, a combination that includes the pressure contribution to the inertial energy density, and occurs repeatedly in our study of relativity. Perhaps somewhat less intuitively, the gravitating mass within r, $\mathcal{M}$ has acquired an additional contribution of $4\pi Pr^3/c^2$. But the pressure is a full-fledged component-member of the source function $S_{\mu\nu}$, and we should not be surprised to see it "pull its own weight" (so to speak) as a material source for the gravitational field. Indeed, we shall see this same combination of terms put in an appearance in the cosmological equations for a homogeneous universe. As to the final relativistic factor, which is just $A(r)$, *that* we

can make "disappear" with some contravariant sleight-of-hand. Watch carefully now, and write the TOV equation in the following simpler form:

$$g^{rr}\frac{dP}{dr} = -\frac{4\pi G r}{3}\left(\bar{\rho}(r) + \frac{3P}{c^2}\right)\left(\rho + \frac{P}{c^2}\right) \tag{511}$$

where g^{rr} is the rr component of the $g^{\mu\nu}$ metric tensor (*i.e.* $1/A$), and $\bar{\rho}(r)$ is the average density within radius r. The gradient ∂_r operator acting on the scalar P is the radial component of a *covariant* tensor; the factor g^{rr} changes it to a standard *contravariant* tensor radial component. The right side of the equation must then itself be proportional to the radial component of a standard gravitational acceleration contravariant vector, let us say a^r. While this g^{rr} coefficient is manifestly geometrical and ensures that a contravariant form of the gradient appears, it can also be thought of more physically as embodying the contribution to gravity arising from curvature itself: "the gravity of gravity." In general relativity, the effects of gravity act geometrically, showing up in the metric terms or their derivatives. By contrast, the density and pressure modifications act dynamically: they come directly from the stress tensor source.

We are not quite finished, because we need to determine $B(r)$ to complete the solution of our metric. This is a simple matter of combining equations (507) and (508):

$$\frac{B'}{B} = \left(\frac{2G\mathcal{M}(r)}{r^2c^2} + \frac{8\pi GPr}{c^4}\right)\left(1 - \frac{2G\mathcal{M}(r)}{rc^2}\right)^{-1} \tag{512}$$

We use the boundary condition that $B = 1$ at $r = \infty$, the Minkowski space limit far from the star. Then our solution may be written

$$\ln B = -\int_r^\infty \left(\frac{2G\mathcal{M}(r')}{r'^2c^2} + \frac{8\pi GPr'}{c^4}\right)\left(1 - \frac{2G\mathcal{M}(r')}{r'c^2}\right)^{-1} dr' \tag{513}$$

The integral takes on a more familiar form when $r > R$, i.e., when we are outside the star. Then $\mathcal{M}(r')$ is just the constant M, the total mass of the star, and the pressure $P = 0$.

$$\ln B = -\int_r^\infty \frac{2GM}{r'^2c^2}\left(1 - \frac{2GM}{r'c^2}\right)^{-1} dr' = \ln\left(1 - \frac{2GM}{rc^2}\right) \tag{514}$$

In other words,

$$B(r > R) = 1 - \frac{2GM}{rc^2}. \tag{515}$$

We recover the Schwarzschild solution in the vacuum outside the spherical star, as we must.

8.3 Constant density stars

An exact, but relatively simple, mathematical solution of the TOV equation is that for a star of constant energy density ρc^2. The Newtonian case is particularly simple. With $\mathcal{M}(r) = 4\pi G\rho r^3/3$, the solution to (510) is:

$$\frac{P}{\rho} = \frac{2\pi G\rho}{3}(R^2 - r^2) \tag{516}$$

where R is the outer radius, the location at which the pressure P vanishes. The central pressure P_c is then just

$$P_c = \frac{2\pi G\rho^2}{3}R^2 \tag{517}$$

In the TOV case, the differential equation is more complicated, but it does separate nicely:

$$\frac{dP}{(\rho + 3P/c^2)(\rho + P/c^2)} = -\frac{4\pi Gr}{3}\left(1 - \frac{8\pi G\rho r^2}{3c^2}\right)^{-1} dr.v \tag{518}$$

The solution with $P(R) = 0$ is obtained via a partial fraction expansion on the left side of this equation, then standard logarithmic integrals over P. On the right side, we have already, up to a constant factor, an exact derivative. Carrying out these straightforward integrations, we find:

$$\frac{\rho + 3P/c^2}{\rho + P/c^2} = \left(\frac{1 - 8\pi G\rho r^2/3c^2}{1 - 8\pi G\rho R^2/3c^2}\right)^{1/2}. \tag{519}$$

Now this is a *very* interesting solution. With mass $M = 4\pi\rho R^3/3$, by demanding that the solution make physical sense at $r = 0$ we find the restriction

$$\left(\frac{1}{1 - 2GM/Rc^2}\right)^{1/2} = \frac{\rho + 3P/c^2}{\rho + P/c^2} \equiv 3 - \frac{2\rho}{\rho + P/c^2} < 3. \tag{520}$$

This condition can be fulfilled only if

$$R > \frac{9GM}{4c^2} = \frac{9R_S}{8}, \tag{521}$$

which states that the actual radius of a constant density star must always be *larger* than $9R_S/8$, not R_S itself as you might have expected. What if the density is not constant? If the matter is compressible, it doesn't help at all. The constant density case is actually an absolute upper limit to the M/R ratio in any star. If the stellar material can be compressed, then for a given mass M, you will have to distribute the material over a *larger* radius R to prevent inward gravitational collapse. (Collapse is clearly easier to trigger when the self-gravitating matter is softer!) We therefore have a very powerful result:

All stable stars must be larger in radius than $9GM/4c^2$.

Note as well that equation (520) states that the maximum observed-to-emitted frequency ratio possible (the largest redshift) for radiation emerging from the surface of any kind of star is 3.

Exercise. Show that the inequality (521) is actually an *upper* bound on R for a given density, and therefore, the maximum mass of a constant density star is

$$M_{\max} = \frac{4}{27}\sqrt{\frac{3}{\pi}}\frac{c^3}{G^{3/2}\rho^{1/2}},$$

corresponding to infinite central density. For a density of $\rho \simeq 2 \times 10^{15}$ g cm^{-3}, what is $M_{\max}$ in solar masses? (Answer: $2.54M_\odot$.) The assumption of a constant density is much too crude a model to provide an accurate physical upper limit for real stars, but the calculation illustrates an important point of principle: an upper mass limit exists.

8.4 White Dwarfs

The existence of an upper bound to the mass of a white dwarf star may be understood by simple dimensional scaling arguments.

The potential energy of a (nonrelativistic) star is always a number that is of order $-GM^2/R$, where M is the mass and R the radius of the star. For a fixed mass, this behaves as $1/R$, so shrinking the star always lowers the gravitational energy (more negative). In the case of a pressure-free body ("dust"), we could in principle shrink the mass down to zero radius and attain a state of arbitrarily low energy. Stars exist in nature because this does not happen, pressure fights back. The thermal energy of the star generally grows rapidly as it shrinks, and the star effectively resists the gravitational collapse. Too small a radius would provide too much of a push-back, and the star would reexpand. But too large a radius would make pressure inefficient, and the star could once again lower its gravitational energy by contraction, so there is a happy medium that minimises the total energy, the sum of the potential plus thermal energies. This minimum energy state is where the star lives. This is stellar equilibrium.

The thermal energy of the star is of order PR^3, some sort of average pressure times the stellar volume. The argument we outined above works both for ordinary matter, as well as for degenerate matter, which obeys an equation of state of the form $P \propto \rho^{5/3}$: the pressure is proportional to the 5/3 power of the density and nothing else. This is the equation of state for a *nonrelativistic* white dwarf supported by electron degeneracy pressure. Then

$$PR^3 \sim \rho^{5/3} R^3 \sim M^{5/3}/R^2 \tag{522}$$

At fixed mass but variable R, we see that the thermal energy dominates at small radius (pushing back) and gravitational potential energy wins at large radius (pulling in). As outlined above, there is always a stable equilbrium somewhere in between. *This holds for any mass at all,* so if this were always the correct equation of state, a white dwarf star could have any mass.

An upper mass limit arises, in fact, because of the following interesting behaviour. We have seen that stellar equilibrium is achieved when the potential and thermal energies are comparable. For a nonrelativistic white dwarf, this means that the balance is

$$M^{5/3}/R^2 \sim M^2/R, \tag{523}$$

which leads immediately to the scaling relation $R \sim M^{-1/3}$: the more massive a degenerate star is, the *smaller* it is! This is in stark contrast to ordinary stars (and ordinary people), whose radius increases with mass. For nondegenerate stars, R grows close to linearly with M. (This is related to a more-or-less constant internal temperature regulated by a nuclear burning thermostat. Why does this lead to $R \sim M$?) The point here is that at larger and larger mass, the density M/R^3 rises as M^2, the electrons become ever more tightly packed, and their kinetic energy (of order their Fermi energy) must enter the relativistic regime. Then the equation of state is no longer $P \propto \rho^{5/3}$, but $P \propto \rho^{4/3}$. Under these circumstances, the thermal energy scales as

$$PR^3 \sim \rho^{4/3} R^3 \sim M^{4/3}/R. \tag{524}$$

In other words, both the thermal and potential energies scale (at fixed mass) in the same way with radius, as $1/R$. Moreover, the potential energy is proportional to M^2, versus $M^{4/3}$ for the thermal energy, so that at large enough mass the star can always lower its total energy by collapsing, like dust, down to $R = 0$. *Thus, the properties of relativistic electron degeneracy lead directly to the existence of a maximum mass, beyond which degenerate electrons cannot prevent gravitational collapse.* This is known as the *Chandrasekhar Mass*, after the

astrophysicist who first calculated its value accurately. (He did this as a student!) Denoted M_{Ch}, its value is close to $1.42 M_\odot$ for a typical abundance of elements. For astronomers, this is really not a very big mass at all.

It is most instructive to work out the precise value of M_{Ch}. Begin with the Newtonian equation of stellar equilibrium, in which the gravitational force is written as the gradient of a potential function Φ:

$$\frac{1}{\rho}\frac{dP}{dr} = -\frac{d\Phi}{dr}, \quad \Phi = -\int_r^\infty \frac{G\mathcal{M}(r')}{r'^2}\,dr' \tag{525}$$

Then, taking the divergence in spherical coordinates gives

$$\frac{1}{r^2}\frac{d}{dr}r^2\left[\frac{1}{\rho}\frac{dP}{dr}\right] = -\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -4\pi G\rho, \tag{526}$$

where the final equality uses the spherical Poisson equation $\nabla^2\Phi = 4\pi G\rho$. Given an equation of state in form of the function $P = P(\rho)$, this is a self-contained second-order differential equation, which may be solved (at least by numerical methods).

Recall that a fully degenerate gas has one fermion per available energy state, up to the Fermi energy, ϵ_F. With a density of states per unit volume of $8\pi p^2 dp/h^3$ (this includes two spin states at each level; h is Planck's constant), the total number of electrons per unit volume is

$$n_e = \int_0^{p_F} 8\pi p^2 \frac{dp}{h^3} = \left(\frac{8\pi}{3}\right)\left(\frac{p_F^3}{h^3}\right) \tag{527}$$

where p_F is the Fermi momentum associated with ϵ_F. The total pressure, calculated most easily as the momentum per unit time per unit volume imparted to an imaginary wall in xy plane is (note that we now explicity integrate over θ in spherical momentum coordinates):

$$P = 2 \times \int_0^{p_F}\int_0^{\pi} v_z(2p_z) \times (1/2)\, 2\pi p^2 \sin\theta\, d\theta \frac{dp}{h^3} = \frac{8\pi}{3h^3}\int_0^{p_F} v\, p^3 dp \tag{528}$$

where we have used $v_z = v\cos\theta$ and similarly for p_z. (The various factors of 2 are present for spin, for twice p_z imparted with each elastic collsion to the wall, with a 1/2 because only half the particles are moving toward the wall. The 2π is from the ϕ integration.) The θ integral is 2/3. This expression, as written, is valid both relativistically and nonrelativistically. For the nonrelativistic case, $v = p/m_e$ (m_e is the electron mass), and

$$P = \frac{8\pi p_F^5}{15 m_e h^3} \quad \text{(nonrelativistic)} \tag{529}$$

With m_p the proton mass and $\mu_e m_p$ the total mass per electron in the gas (μ_e would be very close to 2 for equal proton and neutron numbers), we then have

$$P = \left[\frac{h^2}{20 m_e (m_p\mu_e)^{5/3}}\right]\left(\frac{3}{\pi}\right)^{2/3} \rho^{5/3} \equiv K_1\,\rho^{5/3} \quad \text{(nonrelativistic)} \tag{530}$$

Numerically, $K_1 = 3.123 \times 10^6$ in SI units. An important characteristic density at which p_F is given by $m_e c$ is

$$\rho_{char} = \frac{8\pi}{3}\frac{m_e^3 c^3}{h^3}\mu_e m_p = 0.97\mu_e \times 10^9 \text{ kg m}^{-3}. \tag{531}$$

When the mass density is comparable to or exceeds ρ_{char}, the gas enters the relativistic regime.

For the extreme relativistic case, $v = c$ in the rightmost integral of equation (528) and

$$P = \frac{2\pi c p_F^4}{3h^3} \quad \text{(extreme relativistic)} \tag{532}$$

or

$$P = \left[\frac{hc}{8(m_p\mu_e)^{4/3}}\right]\left(\frac{3}{\pi}\right)^{1/3}\rho^{4/3} \equiv K_2\,\rho^{4/3} \quad \text{(extreme relativistic)} \tag{533}$$

Numerically, $K_2 = 4.887 \times 10^9$ in SI units.

Let us next consider equation (526) for the general so-called polytropic equation of state, $P = K\rho^\gamma$. Substituting this expression for P leads to

$$\frac{1}{r^2}\frac{d}{dr}r^2\left[\frac{d\rho^{\gamma-1}}{dr}\right] = -\frac{4\pi G(\gamma-1)}{\gamma K}\rho \tag{534}$$

Let ρ_0 be the central density of the star. If we now make the substitutions,

$$n = \frac{1}{\gamma-1}, \quad \rho = \rho_0\theta^n, \quad r = \left[\frac{K(n+1)\rho_0^{(1/n)-1}}{4\pi G}\right]^{1/2}\xi \equiv A\xi \tag{535}$$

the equation takes the form of what is known as the *Lane-Emden* equation,

$$\frac{1}{\xi^2}\frac{d}{d\xi}\xi^2\left[\frac{d\theta}{d\xi}\right] = -\theta^n. \tag{536}$$

The original Newtonian equation of hydrostatic equilibrium may be written in an inverted form as

$$\mathcal{M}(r) = -\frac{r^2}{\rho}\frac{dP}{dr} = -\left[\frac{K^3(n+1)^3}{4\pi G^3}\right]^{1/2}\rho_0^{(3-n)/2n}\,\xi^2\frac{d\theta}{d\xi}. \tag{537}$$

Thus, we may read off the mass of the star as given by the above formula, with the final Lane-Emden terms evaluated at the location for which $\theta = 0$, the dimensionless stellar radius.

The boundary conditions to be applied for the solution of (536), integrating outward from the centre $r = 0$, are

$$\theta(0) = 1, \quad \frac{d\theta}{d\xi} \equiv \theta'(0) = 0. \tag{538}$$

The vanishing derivative condition follows from (537) and the fact that the mass contained within a very small ξ neighbourhood of $\xi = 0$ goes like $\mathcal{M}(r) \sim \xi^3$. For example, an exactly soluble solution to the Lane-Emden equation is available for the case $n = 1$,

$$\theta = \frac{\sin\xi}{\xi}, \quad n = 1 \tag{539}$$

The surface of the star for this solution corresponds to $\xi = \pi$, and at this point $\theta' = -1/\pi$. Equation (537) then gives for the mass of the star

$$M = \rho_0\left[\frac{2\pi K^3}{G^3}\right]^{1/2}. \tag{540}$$

More generally, the equation must be solved by numerical methods, which is a straightforward task.

The cases $n = 1.5$ and $n = 3$, corresponding to an ordinary white dwarf and a relativistic degenerate white dwarf, are the most interesting for astrophysics. With $n = 1.5$,

$$M = -\left[\frac{125K_1^3}{32\pi G^3}\right]^{1/2} \rho_0^{1/2}\, \xi^2 \frac{d\theta}{d\xi} \tag{541}$$

For the given numerical value of K_1, $\mu_e = 2$, a central density of 10^9 kg per cubic metre and the numerical result $-\xi^2\theta' = 2.7$, this gives a mass of $0.48\mathrm{M}_\odot$, a typical low mass white dwarf.

Yet more interesting is the relativistic equation of state, $n = 3$. Then,

$$M = -\left[\frac{16K_2^3}{\pi G^3}\right]^{1/2} \xi^2 \frac{d\theta}{d\xi}, \tag{542}$$

which is *independent* of the central density ρ_0, and depends only upon fundamental constants of physics. From the solution of the Lane-Emden equation, $-\xi^2\theta' = 2.02$, conveniently very close to 2. For $\mu_e = 2$, this gives a numerical value of $M = 1.43\mathrm{M}_\odot$. It is known as the Chandrasekhar mass, M_{Ch}, and it is the fundamental upper mass limit to a white dwarf star. It is a sufficiently important quantity to be written down on its own, expressed in terms of fundamental constants of nature. Replacing, for the sake of elegance, 2.02 with 2, we have

$$M_{Ch} = 8\left[\frac{K_2^3}{\pi G^3}\right]^{1/2} = \left(\frac{3}{8\pi^2}\right)^{1/2}\left(\frac{1}{\mu_e m_p}\right)^2\left(\frac{hc}{G}\right)^{3/2} \tag{543}$$

(Purists may wish to multiply this by a factor of $2.018/2 = 1.009$ for accuracy.)

With this explicit solution in hand, let us recall why this mass is a true upper mass limit. Imagine a sequence of white dwarf models with ever increasing total mass. In the non-relativistic regime, the central density ρ_0 increases as M^2. The effective adiabatic index, at first fixed at $n = 1.5$, starts to rise toward $n = 3$ as the growing central density compels the electrons into the relativistic regime of their equation of state. As M continues to rise, n closes in on 3, and the central density now scales as the mass M to the power $6/(3-n)$, becoming infinite at $M = M_{Ch}$. At this point, there is nothing to prevent gravitational collapse to either a neutron star or, failing that, a black hole.

Exercise. What is the actual physical radius of an $n = 1.5$ white dwarf? Use $\rho_0 = 10^9$ kg per m^3, and the fact that θ vanishes at $\xi = 3.65$.

8.5 Neutron Stars

We have seen that a massive stellar core or a white dwarf that has accreted enough matter to bring its total mass above M_{Ch} is in deep trouble. Such a body will rapidly collapse. As the core density shoots up, the inverse of the classical beta-decay reaction

$$p + e^- \rightarrow n + \nu_e$$

induces electrons (e^-) to recombine with protons (p) to form a sea of neutrons (n). In the process, electron neutrinos (ν_e) escape from the initial *implosion,* which is thought then

to immediately rebound as a supernova *explosion.* This is a fascinating topic unto itself, controversial and far from well-understood, but our interest here is in the neutrons, the dominant majority species left behind. The normal beta-decay of free neutrons (typically 10 minutes) is halted by a small residual population (typically a few percent) of protons and electrons, which fill their respective "Fermi seas." This keeps the forward and backward beta reaction rates in balance. The neutrons accordingly have one last chance to avoid the fate of complete collapse into a black hole: their degeneracy pressure can resist gravity at these ultra high densities, just as electrons do in a white dwarf. For such a neutron star, the characteristic density equation (531) becomes

$$\rho_{char} = \frac{8\pi}{3}\frac{m_n^4 c^3}{h^3} = 6.11 \times 10^{18} \text{ kg m}^{-3} = 6.11 \times 10^{15} \text{ g cm}^{-3}. \tag{544}$$

We have replaced the typical Fermi momentum condition of a white dwarf, $p_F \sim m_e c$, with $p_F \sim m_n c$. (m_n is the neutron mass, differing very little from m_p.) Neutron stars have, not by coincidence, densities characteristic of atomic nuclei.

The question immediately arises as to whether there is an upper limit to the mass of a neutron star analogous to that of M_{Ch} for a white dwarf. That this is to be expected follows from the fact that both the pressure and the gravitational field itself act as sources of self-gravity: a bigger self-attraction for the same amount of mass/energy. If anything, it should be even more difficult to stave off collapse in the case of a neutron star as compared with a white dwarf. This is indeed the case, as found in a classical calculation by Oppenheimer and Volkoff, at least for an ideal Fermi gas of neutrons.

Exercise. Show that relativistic parameter GM/Rc^2 is of order m_e/m_p for a white dwarf, and of order unity for a neutron star. Use the appropriate ρ_{char} in each case.

In the case of a white dwarf, the energy density ρc^2 is completely dominated by the nonrelativistic nucleon rest mass. This is no longer so for a neutron star, and the equation of state must take this into account. In particular, for a fully degenerate relativistic gas of neutrons, ρc^2 is given by

$$\rho c^2 = \int_0^{p_F} \frac{8\pi}{h^3} p^2 (p^2c^2 + m_n^2 c^4)^{1/2} dp = 3\rho_{char} c^2 \int_0^{p_F/m_n c} u^2 (1+u^2)^{1/2} du \tag{545}$$

The pressure is given now by equation (528), with $v \equiv pc^2/E$ (E here is the particle energy), as

$$P = \frac{8c^2\pi}{3h^3} \int_0^{p_F} \frac{p^4\, dp}{(p^2c^2 + m_n^2 c^4)^{1/2}} = \rho_{char} c^2 \int_0^{p_F/m_n c} \frac{u^4\, du}{(1+u^2)^{1/2}} \tag{546}$$

The integrals may be evaluated explicitly (with $x = p_F/m_n c$):

$$\int_0^x u^2 (1+u^2)^{1/2} du = \frac{1}{8}\left(x(1+2x^2)\sqrt{1+x^2} - \sinh^{-1} x \right), \tag{547}$$

$$\int_0^x \frac{u^4}{(1+u^2)^{1/2}} du = \frac{1}{8}\left(x(2x^2-3)\sqrt{1+x^2} + 3\sinh^{-1} x \right), \tag{548}$$

Equations (547) and (548) may be written in much more elegant forms, which are also convenient for numerical computation. If we substitute in (547) $\alpha = \sinh^{-1} x$, the right hand side (RHS) becomes

$$\text{RHS} = \frac{1}{32}(\sinh 4\alpha - 4\alpha), \tag{549}$$

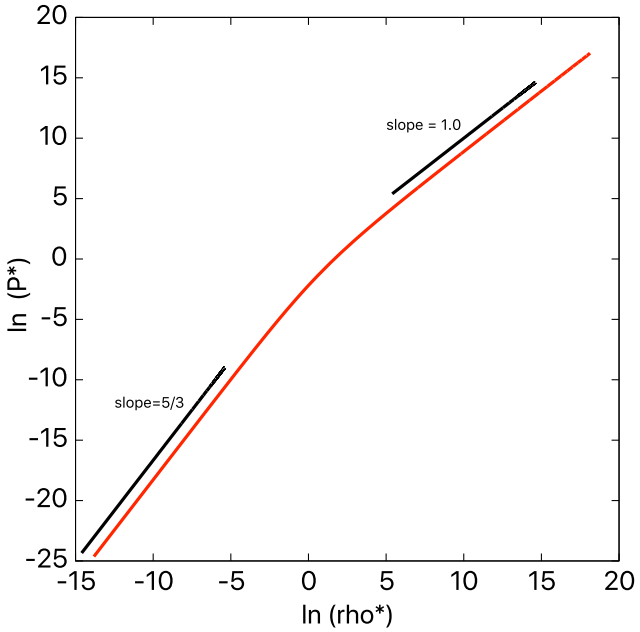


Figure 11: The equation of state for a relativistic neutron gas, via the parametric solution of equations (547) and (548). At small p_F/m_nc, the nonrelativistic $P = K_1\rho^{5/3}$ relation is recovered. At large p_F/m_nc, the extreme relativistic equation of state $P = \rho c^2/3$, analogous to radiation, is recovered. (Notation: rho$^* = \rho/\rho_{char}$, $P^* = P/\rho_{char}c^2$.)

where we have made use of the hyperbolic trig identities (prove, if they are not familiar!)

$$1 + \sinh^2\alpha = \cosh^2\alpha, \quad 2\cosh\alpha\sinh\alpha = \sinh 2\alpha, \quad 1 + 2\sinh^2\alpha = \cosh 2\alpha.$$

With $t \equiv 4\alpha$ (using the t-notation of Oppenheimer and Volkoff), we may parameterise ρ by the compact equation:

$$\rho(t) = \frac{3\rho_{char}}{32}(\sinh t - t), \quad t = 4\sinh^{-1}(p_F/m_nc). \tag{550}$$

Using exactly the same tricks for the pressure integral produces the slick result

$$P(t) = \frac{\rho_{char}c^2}{32}[\sinh t - 8\sinh(t/2) + 3t] = \frac{c^2}{3}[\rho(t) - 8\rho(t/2)]. \tag{551}$$

With explicit forms for $\rho(t)$ and $P(t)$ at hand, the differential equations to be solved consist of the mass equation for $d\mathcal{M}/dr$,

$$\frac{d\mathcal{M}}{dr} = 4\pi r^2\rho, \tag{552}$$

and the TOV equation, now written as an equation for dt/dr:

$$\frac{1}{\dot{P}}\frac{dP}{dr} = \frac{dt}{dr} = -\frac{1}{\dot{P}}\left(\frac{G\mathcal{M}}{r^2} + \frac{4\pi GPr}{c^2}\right)\left(\rho + \frac{P}{c^2}\right)\left(1 - \frac{2G\mathcal{M}}{rc^2}\right)^{-1}, \tag{553}$$

where $\dot{P} = dP/dt$, and both ρ and P are explicit functions of t, equations (550) and (551). The boundary conditions at $r = 0$ are $\mathcal{M} = 0$ and $t = t_0$ where t_0 is implicitly defined for a given central density ρ_0 by the equation

$$\rho_0 = \frac{3\rho_{char}}{32}(\sinh t_0 - t_0). \tag{554}$$

An explicit equation of state of the form $P = f(\rho)$ is shown in Fig. (11). The actual $f(\rho)$ functional relation is remarkably simple in appearance, changing smoothly from one power law to another.

When the central density of the star ρ_0 is much less than ρ_{char}, the entire star behaves exactly like an $n = 1.5$ white dwarf, but with m_n substituting for $\mu_e m_p$ in the formulae. The mass of this type of low mass neutron star can be written down directly from the relevant solution of the Lane-Emden equation, something we now know how to do. The answer turns out to be

$$M = 2.71(\rho_0/\rho_{char})^{1/2} M_\odot \tag{555}$$

If $\rho_0/\rho_{char} = 0.03$, say, this is $0.47\mathrm{M}_\odot$.

When $\rho_0 \gg \rho_{char}$, the equation of state becomes $P = \rho c^2/3$, very much like a gas of photons. Then, the TOV equation (508) is

$$\frac{d\rho}{dr} = -\frac{4G\mathcal{M}(r)\rho}{r^2c^2}\left(1 + \frac{4\pi\rho r^3}{3\mathcal{M}(r)}\right)\left(1 - \frac{2G\mathcal{M}(r)}{rc^2}\right)^{-1} \tag{556}$$

Surely it is *obvious* that an exact solution is

$$\rho = \frac{3c^2}{56\pi G r^2}! \tag{557}$$

(Try it out. Start with $\mathcal{M}(r) = 3rc^2/14G$ and go from there.) This simple inverse square law is in fact the correct asymptotic local behaviour near the origin of the true physical solution in the limit $\rho_0 \to \infty$. Note that the integrated mass does not diverge, but remains finite near $r = 0$. An equilibrium solution that vanishes at some finite r and has a well-defined finite mass may be found numerically. This was first attempted by Oppenheimer and Volkoff, but the details of their reported large density behaviour are not quite right. Given the cumbersome technology of 1930's hand calculators, we should not judge them too harshly! By the early 1960's, matters had been resolved.

A plot[15] of the stellar mass as a function of central density ρ_0 is given in figure (12). The limiting mass as $\rho_0 \to \infty$ is $0.42\,M_\odot$. (The oscillatory manner in which this value is approached is surprising, and can be related to underlying stability properties of the star.) But equation (555) shows that neutron stars in the Lane-Emden low ρ_0 regime, in which M is rising with increasing ρ_0, clearly includes stellar masses exceeding this value. There must therefore be a maximum mass somewhere between the high and low density regimes, and figure (12) bears this out. There is a maximum mass of $0.71\mathrm{M}_\odot$ and a radius of 9.6km (correctly calculated by Oppenheimer and Volkoff), the classic upper limit for a gas of degenerate neutrons This occurs for t_0 very close to 3, $\rho_0 = .66\rho_{char} = 4 \times 10^{15}$ g cm^{-3}.

If the ideal Fermi neutron gas were the whole story, there probably would not be very many neutron stars in the Universe. Stars would become white dwarfs with masses up to

[15] I am grateful to L. Fraser-Taliente, R. McDonald and A. Morris, all of whom flawlessly carried out this calculation as an optional exercise for this course in 2019. Prior to their work, I was unaware that the original Oppenheimer-Volkoff mass calculation was inaccurate at large densities.

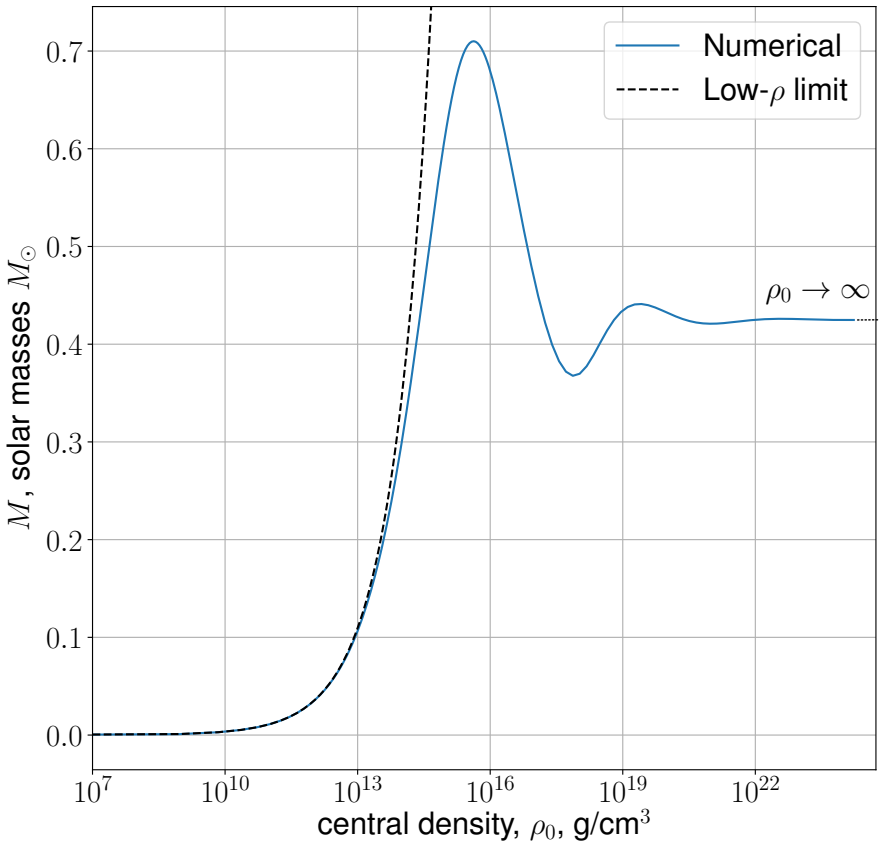


Figure 12: Computed mass M of a model neutron star for an ideal Fermi gas equation of state. The x-axis is central density ρ_0, in units of g cm^{-3}, on a logarithmic scale. At low ρ_0, the mass rises as the square root of the density, like a white dwarf star (dotted line). At higher densities, M reaches a maximum of $0.71\mathrm{M}_\odot$. M then decreases, oscillating with decreasing amplitude, approaching $0.42\mathrm{M}_\odot$ as $\rho_0 \to \infty$. (Figure credit: L. Fraser-Taliente.)

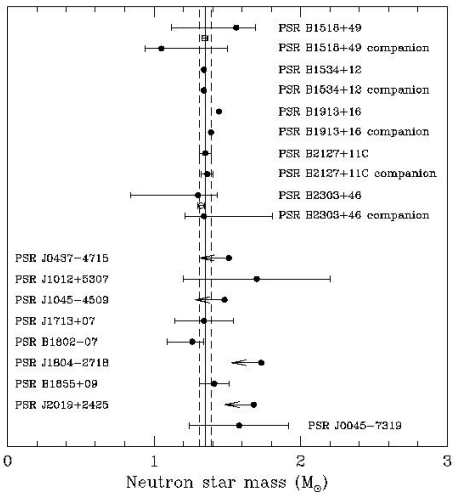


Figure 13: Neutron star (NS) masses from binary systems, determined by precise orbital timing measurments. The dotted lines show a spread of white dwarf upper mass limits, allowing for variation in composition. Left pointing arrows are observational upper limits for the NS mass . NS masses cluster near the white dwarf upper limit.

about $1.4M_\odot$. Beyond this value, only black holes would be possible. Perhaps an imploding white dwarf might throw off enough mass to get below the Oppenheimer-Volkoff limit and become a neutron star, but this is hardly certain. In fact, at ultrahigh densities, neutrons interact via the strong nuclear force, and the ideal gas equation of state used by Oppenheimer and Volkoff is not a good approximation. (White dwarf electrons experience no such complication.) There has long been great uncertainty regarding the proper equation of state to use for calculating the maximum mass for a neutron star. Recent (2017) data from the merger of two actual neutron stars, in the form of what the post-collision debris did and did not do, has been used to constrain the parameter space allowed for a proper equation of state. These data suggest an upper mass limit of 2.17 $M_\odot$, more than a factor of 3 larger than the original Oppenheimer-Volkoff value.

Our modern view of neutron stars is therefore somewhat different. There is a range of masses between $1.4M_\odot$ and $2.17M_\odot$ where a star appears to be able to reside happily as a neutron star. A white dwarf in a binary system that has accreted enough matter to bring its mass just above the Chandrasekhar mass can form a neutron star after implosion, without losing a significant fraction of its mass in the process. We would expect, therefore, to find many neutron stars with masses near the Chandrasekhar value of $1.4M_\odot$. That is indeed what observations show (Fig. 13).

8.6 Pumping iron: the physics of neutron star matter

To understand the behaviour of neutron stars in their proper context, we follow the physicist John Wheeler (mentor of, among other notables, Richard Feynman and Kip Thorne) who in the 1950s computed the equation of state for matter in a massive star at the end of all nuclear processing. This corresponds to pure iron nuclei, since iron has the greatest (negative) binding energy per nucleon of any atomic nucleus. It is not possible to extract

energy from an iron nucleus either by fusing it with other particles or by breaking it apart. Wheeler worked with B. Harrison on this project, and the result is known as the Harrison-Wheeler (HW) equation of state, the first of its kind for calculating neutron star interiors. In this section, we will give a qualitative description of how iron, upon steadily increasing compression, makes the transition from ordinary matter to the stuff of neutron stars. Our discussion in this section will be largely descriptive.

In our description we shall refer both to the quantity $\gamma = d\ln P/d\ln\rho$, the adiabatic index of the matter, as well as to the physical density. For ordinary iron ($\rho = 7.6$g per cubic cm), γ is huge. This is why such a solid feels hard: to achieve a tiny density compression requires a vast exertion of pressure. As we keep squeezing, γ actually plummets, because $d\ln P = dP/P$ drops very rapidly with the sharply rising P. The internal electrons, for their part, congregate more and more tightly around their parent nuclei, resisting the compression by standard electron degeneracy pressure. The applied external pressure changes the spacing of the internal energy levels in the atoms, but not the number of electrons per level. That is fixed by degeneracy, and it is what makes this squeezing process an adiabatic one. This state of affairs persists for compression factors of up to 10^4, i.e. up to densities of $\sim 10^5$ g cm^{-3}. That may seem impressively high, but we are still very far from a neutron star.

Next, we enter the white dwarf regime. At densities in excess of 10^5 g cm^{-3}, the forces of degeneracy pressure start to dominate over the electrical attraction of the iron nuclei. As a consequence, the confining atomic orbitals cease to exist, and something much more like a metal appears, in which the electrons are free to roam between the nuclei. This is true white dwarf matter. The adiabatic index is now easy to understand: $\gamma = 5/3$. As we continue to increase the pressure and density, the confined electron degeneracy motions approach the speed of light. At a density of $\rho = 10^7$ g cm^{-3}, γ has dropped to 4/3. This is the relativistic white dwarf regime of the Chandrasekhar limit.

We remain in the $\gamma = 4/3$ phase as the density compression rises through a factor 40,000. Then, at a density of 4×10^{11} g cm^{-3}, the next transition occurs: the electrons no longer have any space between the nuclei! The minimum energy state is one in which they are absorbed by the nuclear protons. This enriches the neutron content of the matter by inverse beta decay (proton plus electron goes to neutron plus electron neutrino):

$$p + e^- \rightarrow n + \nu_e$$

The loss of electron degeneracy pressure at these densities causes γ to drop well below unity. This is a global minimum for the adiabatic index.

As the density continues to rise, it become energetically more favourable for the neutron rich nuclei to "drip neutrons" back out into the ambient surroundings, where densities (and thus the degeneracy pressure) are lower. As the compression continues and the overall density climbs, the dripped-out neutrons start to resist back with their own degeneracy pressure. Finally, by 4×10^{12} g cm^{-3}, the nuclei have broken apart, and we have a gas of free, degenerate neutrons. This is the stuff from which an Oppenheimer-Volkoff neutron star is made.

This was as far as Harrison and Wheeler could go in the 1950's. This was enough, however, to show that there were no stable stars in the density gap between white dwarfs and neutron stars. In the following decades, work by many individuals on the equation of state for ultrahigh densities showed a greater resistance and thus a stiffer equation of state than a simple extrapolation of the ideal gas equation of state used by Oppenheimer and Volkoff. It is this added stiffness that is the basis for our understanding of how neutron stars can exist with masses in excess of the Oppneheimer-Volkoff upper limit of $0.7M_\odot$.

Figure (14), taken from the semi-popular text of Kip Thorne, *Black Holes & Time Warps: Einstein's Outrageous Legacy* (1994), shows how the scenario we have just outlined translates more generally into equilibrium masses for actual gravitating bodies. (I apologise for the

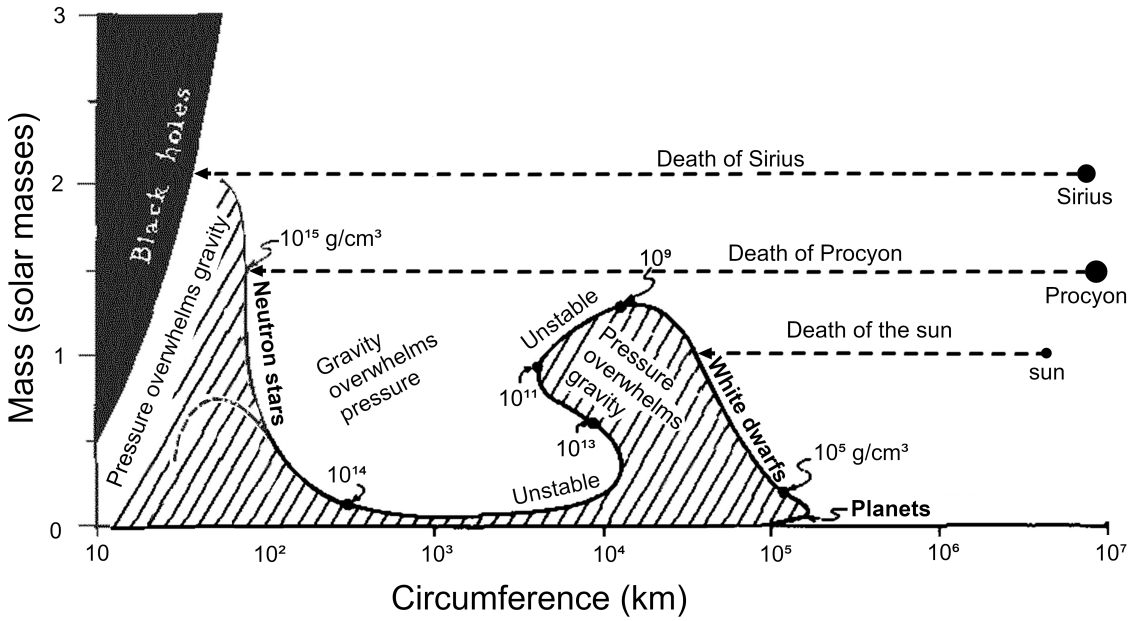


Figure 14: Equilibrium mass as a function of circumference for self-gravitating iron. (Circumference is easier to work with radius since it is not directly affected by curvature.) Only the branches labelled 'White dwarfs' and 'Neutron stars' are stable. See text for a detailed description. (Figure from *Thorne,* modified for clarity by L. Fraser-Taliente.)

difficulty in reading some of the numbers.) The figure is a plot of Stellar Mass (in units of the Sun's mass) as a function of circumference in kilometres (a quantity less distorted by curvature than radius itself) for pure iron. The km-circumference tick points are $10^1, 10^2, ...10^7$. The solid line plots the equilibrium mass value (or possible multiple values!) for a particular circumference. In the region that is shaded by heavy lines, pressure is larger than gravity; in the white region gravity is larger than pressure forces. The solid line marks the equilibrium boundary. The powers of 10 in the figure that are shown pointing with wiggly arrows to the equilibrium mass curve indicate the corresponding central density in units of grams per cubic cm.

It is very important to note that not every value of stellar mass along the equilibrium curve is actually realisable in nature. Only the portions of the mass curve which are monotonically decreasing with circumference are physically stable! Consider, for example, the curve portion labelled "White Dwarfs." Move slightly to the right, off the curve into the white region. The mass boundary is extended, and the pressure has dropped while the gravity has gone up: the mass feels a net force that returns it back toward equilibrium. Similarly, move leftward into the shaded zone from the same White Dwarf line. Now contraction drives up the pressure relative to the gravity, and once again the body experiences a return stabilising force.

Contrast this behaviour with the "unstable" branch, say the uppermost between densities of 10^9 and 10^{11}. On this side of the equilibrium curve, a move to right puffing up star leads to pressure dominating over gravity and further puffing! Similarly, a move to the left—a contraction—leads to gravity dominance, and further collapse.

The two stable portions of the equilibrium curve, at circumferences of roughly 60,000 and 60 km, correspond to white dwarf and neutron stars respectively. The dotted line within the shaded portion of the neutron star section corresponds to the classical Oppenheimer-Volkoff calculation with an upper mass limit of $0.7M_\odot$. The solid curve is based on a much better

calculation, taking account of a modern nuclear equation of state. The current most massive neutron star actually measured (as of September 2019) weighs in at $2.14M_\odot$, and there is good reason to believe that this is very close to the maximum mass a neutron star can have before imploding to a black hole. (The black hole region of this diagram is clearly visible on the left.)

Three real stars are shown, contracting leftward toward their fates: the Sun, Procyon, and Sirius. (The latter two are actually binary systems with white dwarf companions. We refer in our diagram to the bright "normal star.") Normal stars are not on the equilibrium curve because they are not cold iron supported by degeneracy presure! But at the end of their nuclear burning they will eventually contract and move to the left in the diagram. The Sun will find solace on the white dwarf branch. Procyon is too massive to end up as a white dwarf, but will catch the stable neutron star branch. For poor Sirius, unless it manages to shed some mass, its fate will be complete gravitational collapse to a black hole.

To go deeper into the topic of the equation of state at ultrahigh densities would take us too far afield from our primary interest in general relativity. The reader who would like to pursue this topic in more detail, starting with the basic physics, should begin by consulting Chapter 2 of the text by Shapiro and Teukolsky for a clear exposition and a number of excellent exercises.

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* 1×10^{-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*

9 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, the language of Riemannian geometry that we have been studying. Finally, there is the language of quantum field theory: gravity arises from the exchange of spin 2 particles, gravitons, much as electromagnetism arises from the exchange of spin 1 photons. Just as the starting point of quantum electrodynamics is the classical radiation theory of Maxwell, the starting point of quantum gravity must be a classical radiation theory of gravity. However, 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 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 now good reasons for the practical astrophysicist to get involved. In Februrary 2016, the first detection of gravitational waves was announced. The event signal had been received and recorded earlier, 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.

In the limit when $g_{\mu\nu}$ is very close to $\eta_{\mu\nu}$, the theory of general relativity is a classical theory of gravitational radiation (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 $h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu}$ become, in this weak field limit, a set of rather ordinary looking linear wave equations with source terms, completely analogous to Maxwell's Equations. The principal difference is that electrodynamics is sourced by a vector current J^{μ}, 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 gravitational field be weak: the gravitational radiation makes a contribution to its own source, something electromagnetic radiation cannot do. 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 theorists 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 to guide 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 well and 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 analysed by Hulse and Taylor. So much orbital information could be extracted from this remarkable system that it was possible to predict, and 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 waves. 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 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[16]. 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 will be more to come. In the near future, it is anticipated by many that extremely delicate pulsar timing experiments, in which arrival times of pulses are measured to fantastic precision, will detect distortions in interstellar 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

[16]Update: This was recognised with the 2017 Nobel Prize in Physics, awarded to Barish, Thorne, and Weiss for the detection of gravitational radiation by the LIGO facility.

rewarded.

9.1 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 begin with the assumption that the spacetime metric is close to Minkowski. 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:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{558}$$

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

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} \tag{559}$$

to ensure $g_{\mu\nu}g^{\nu\kappa} = \delta^{\kappa}_{\mu}$. (You can raise the index of g with η *only* when approximating $g^{\mu\nu}$ as its leading order value, which is $\eta^{\mu\nu}$.) Note that

$$\eta^{\mu\nu}h_{\nu\kappa} = h^{\mu}{}_{\kappa}, \quad \eta^{\mu\nu}\frac{\partial}{\partial x^{\nu}} = \frac{\partial}{\partial x_{\mu}}, \tag{560}$$

so that we can slide dummy indices "up-down" as follows:

$$\frac{\partial h_{\mu\nu}}{\partial x_{\mu}} = \eta_{\mu\rho}\frac{\partial h^{\rho}{}_{\nu}}{\partial x_{\mu}} = \frac{\partial h^{\rho}{}_{\nu}}{\partial x^{\rho}} \equiv \frac{\partial h^{\mu}{}_{\nu}}{\partial x^{\mu}}. \tag{561}$$

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

$$R_{\mu\nu} = R^{(1)}_{\mu\nu} + R^{(2)}_{\mu\nu} + ...\text{etc.} \tag{562}$$

and

$$G^{(1)}_{\mu\nu} = R^{(1)}_{\mu\nu} - \eta_{\mu\nu}\frac{R^{(1)}}{2} \tag{563}$$

where $R^{(1)}_{\mu\nu}$ is all the Ricci tensor terms linear in $h_{\mu\nu}$, $R^{(2)}_{\mu\nu}$ all terms quadratic in $h_{\mu\nu}$, and so forth. The linearised affine connection is

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}\eta^{\lambda\rho}\left(\frac{\partial h_{\rho\nu}}{\partial x^{\mu}} + \frac{\partial h_{\rho\mu}}{\partial x^{\nu}} - \frac{\partial h_{\mu\nu}}{\partial x^{\rho}}\right) = \frac{1}{2}\left(\frac{\partial h^{\lambda}{}_{\nu}}{\partial x^{\mu}} + \frac{\partial h^{\lambda}{}_{\mu}}{\partial x^{\nu}} - \frac{\partial h_{\mu\nu}}{\partial x_{\lambda}}\right). \tag{564}$$

In terms of $h_{\mu\nu}$ and $h = h^{\mu}{}_{\mu}$, from contracting the linear-in-g terms of equation (258) on page 63, we explicitly find

$$R^{(1)}_{\mu\nu} = \frac{1}{2}\left(\frac{\partial^2 h}{\partial x^{\mu}\partial x^{\nu}} - \frac{\partial^2 h^{\lambda}{}_{\mu}}{\partial x^{\nu}\partial x^{\lambda}} - \frac{\partial^2 h^{\lambda}{}_{\nu}}{\partial x^{\mu}\partial x^{\lambda}} + \Box h_{\mu\nu}\right) \tag{565}$$

where

$$\Box \equiv \frac{\partial^2}{\partial x^\lambda \partial x_\lambda} = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \tag{566}$$

is the d'Alembertian (clearly a Lorentz invariant), making a most welcome appearance into the proceedings. Contracting μ with ν, we find that

$$R^{(1)} = \Box h - \frac{\partial^2 h^{\mu\nu}}{\partial x^\mu \partial x^\nu} \tag{567}$$

where we have made use of

$$\frac{\partial h^\lambda_\mu}{\partial x_\mu} = \frac{\partial h^{\lambda\mu}}{\partial x^\mu}.$$

Assembling $G^{(1)}_{\mu\nu}$, we find

$$G^{(1)}_{\mu\nu} = \frac{1}{2}\left[\frac{\partial^2 h}{\partial x^\mu \partial x^\nu} - \frac{\partial^2 h^\lambda{}_\mu}{\partial x^\lambda \partial x^\nu} - \frac{\partial^2 h^\lambda{}_\nu}{\partial x^\lambda \partial x^\mu} + \Box h_{\mu\nu} - \eta_{\mu\nu}\left(\Box h - \frac{\partial^2 h^{\lambda\rho}}{\partial x^\lambda \partial x^\rho}\right)\right]. \tag{568}$$

The full, nonlinear field equations may then formally be written

$$G^{(1)}_{\mu\nu} = -\left(\frac{8\pi G T_{\mu\nu}}{c^4} + G_{\mu\nu} - G^{(1)}_{\mu\nu}\right) \equiv -\frac{8\pi G(T_{\mu\nu} + \tau_{\mu\nu})}{c^4}, \tag{569}$$

where

$$\tau_{\mu\nu} = \frac{c^4}{8\pi G}(G_{\mu\nu} - G^{(1)}_{\mu\nu}) \simeq \frac{c^4}{8\pi G}\left(R^{(2)}_{\mu\nu} - \eta_{\mu\nu}\frac{R^{(2)}}{2}\right) \tag{570}$$

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 §9.4. In linear theory, $\tau_{\mu\nu}$ is neglected in comparison with the ordinary matter $T_{\mu\nu}$, and the equation is

$$\frac{\partial^2 h}{\partial x^\mu \partial x^\nu} - \frac{\partial^2 h^\lambda{}_\mu}{\partial x^\lambda \partial x^\nu} - \frac{\partial^2 h^\lambda{}_\nu}{\partial x^\lambda \partial x^\mu} + \Box h_{\mu\nu} - \eta_{\mu\nu}\left(\Box h - \frac{\partial^2 h^{\lambda\rho}}{\partial x^\lambda \partial x^\rho}\right) = -\frac{16\pi G T_{\mu\nu}}{c^4} \tag{571}$$

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

$$\frac{\partial^2 h}{\partial x^\mu \partial x^\nu} - \frac{\partial^2 h^\lambda{}_\mu}{\partial x^\lambda \partial x^\nu} - \frac{\partial^2 h^\lambda{}_\nu}{\partial x^\lambda \partial x^\mu} + \Box h_{\mu\nu} = -\frac{16\pi G}{c^4} S_{\mu\nu}$$

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 Φ and vector potential $\boldsymbol{A}$:

$$\nabla^2\Phi + \frac{1}{c}\frac{\partial}{\partial t}(\boldsymbol{\nabla}\cdot\boldsymbol{A}) = -4\pi\rho \tag{572}$$

$$\nabla^2\boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial \Phi}{\partial t}\right) = -\frac{4\pi}{c}\boldsymbol{J} \tag{573}$$

(Note: I have used esu units, which are much more natural for relativity. Here ρ is the electric charge density.) Do the following exercise!

Exercise. In covariant notation, with $A^\alpha = (\Phi, \boldsymbol{A})$ and $J^\alpha = (\rho, \boldsymbol{J}/c)$ 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^α are $\Box A^\alpha$ and $\partial^\alpha \partial_\beta A^\beta$. Show that if $\partial_\alpha J^\alpha = 0$ identically, then our equation relating A^α to J^α must be of the form

$$\Box A^\alpha - \partial^\alpha \partial_\beta A^\beta = C J^\alpha$$

where C is a constant to be determined, and that this equation remains unchanged when the transformation $A^\alpha \to A^\alpha + \partial^\alpha \Lambda$ in 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:

$$\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial \Phi}{\partial t} = 0 \tag{574}$$

In covariant 4-vector language, this is simply a vanishing divergence condition, $\partial_\alpha A^\alpha = 0$. Then, the dynamical equations simplify:

$$\nabla^2 \Phi - \frac{1}{c^2}\frac{\partial \Phi}{\partial t^2} = \Box \Phi = -4\pi\rho \tag{575}$$

$$\nabla^2 \boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = \Box \boldsymbol{A} = -\frac{4\pi}{c}\boldsymbol{J} \tag{576}$$

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 $G^{(1)}_{\mu\nu}$ can be written entirely in terms of the Bianchi-like quantity

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{\eta_{\mu\nu} h}{2}, \quad \text{or} \quad \bar{h}^\mu_\nu = h^\mu_\nu - \frac{\delta^\mu_\nu h}{2}. \tag{577}$$

Using this in (568), the *linearised* field equation becomes

$$2G^{(1)}_{\mu\nu} \equiv \Box \bar{h}_{\mu\nu} - \frac{\partial^2 \bar{h}^\lambda{}_\mu}{\partial x^\nu \partial x^\lambda} - \frac{\partial^2 \bar{h}^\lambda{}_\nu}{\partial x^\mu \partial x^\lambda} + \eta_{\mu\nu}\frac{\partial^2 \bar{h}^{\lambda\rho}}{\partial x^\lambda \partial x^\rho} = -\frac{16\pi G T_{\mu\nu}}{c^4}. \qquad \text{(linearised)} \tag{578}$$

(It is easiest to verify this by starting with (578), substituting with (577), and showing that this leads to (568).)

Interesting. Except for $\Box \bar{h}_{\mu\nu}$, 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 (578) 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 (568) and on the right side, change $h_{\mu\nu}$ to $h'_{\mu\nu}$, where

$$h'_{\mu\nu} = h_{\mu\nu} - \frac{\partial \xi_\nu}{\partial x^\mu} - \frac{\partial \xi_\mu}{\partial x^\nu}, \quad h' = h - 2\frac{\partial \xi^\mu}{\partial x^\mu} \tag{579}$$

and the ξ_μ represent any vector function. You will find that the form of the equation is completely unchanged, i.e., the ξ_μ terms cancel out identically! This is a true gauge invariance.

Exercise. Show that under a gauge transformation,

$$\bar{h}'_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{\partial \xi_\nu}{\partial x^\mu} - \frac{\partial \xi_\mu}{\partial x^\nu} + \eta_{\mu\nu}\frac{\partial \xi^\rho}{\partial x^\rho}, \quad \frac{\partial \bar{h}'^\mu{}_\nu}{\partial x^\mu} = \frac{\partial \bar{h}^\mu{}_\nu}{\partial x^\mu} - \Box \xi_\nu.$$

Show as well that equation (578) may be written

$$\frac{\partial \Theta^{\lambda\mu\nu}}{\partial x^\lambda} = -\frac{16\pi G}{c^4} T^{\mu\nu},$$

where

$$\Theta^{\lambda\mu\nu} = \partial^\lambda \bar{h}^{\mu\nu} - \partial^\mu \bar{h}^{\lambda\nu} + \eta^{\mu\nu}\partial_\rho \bar{h}^{\lambda\rho} - \eta^{\lambda\nu}\partial_\rho \bar{h}^{\mu\rho}$$

and that $\Theta^{\lambda\mu\nu} = -\Theta^{\mu\lambda\nu}$. This is the gravitational analogue to the Maxwell equation (85). Is $\Theta^{\lambda\mu\nu}$ gauge invariant on its own?

Note that if we replace $T^{\mu\nu}$ on the right with $T^{\mu\nu} + \tau^{\mu\nu}$, following equations (569) and (570), the Θ-equation becomes an exact statement of energy-momentum conservation for the combined fields arising from material and gravitational sources.

What is manifesting as gauge invariance is actually invariance to coordinate transformations, which acts *to leading linear order* as a gauge transformation. To see this, consider the infinitesimal coordinate transformation,

$$x'^\mu = x^\mu + \xi^\mu(x), \quad \text{or} \quad x^\mu = x'^\mu - \xi^\mu(x') \qquad \text{(to leading order)}. \tag{580}$$

Then

$$g'_{\mu\nu} = \eta'_{\mu\nu} + h'_{\mu\nu} = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma} = \left(\delta^\rho_\mu - \frac{\partial \xi^\rho}{\partial x^\mu}\right)\left(\delta^\sigma_\nu - \frac{\partial \xi^\sigma}{\partial x^\nu}\right)(\eta_{\rho\sigma} + h_{\rho\sigma}) \tag{581}$$

Notice that for the derivatives of ξ^ρ, we may use x instead of x', since the difference is quadratic in ξ. With η' identical to η, to leading order in ξ we must have

$$h'_{\mu\nu} = h_{\mu\nu} - \frac{\partial \xi_\nu}{\partial x^\mu} - \frac{\partial \xi_\mu}{\partial x^\nu}, \tag{582}$$

as we found earlier. 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 $\boldsymbol{\nabla}\Lambda$ to $\boldsymbol{A}$ in Maxwell's theory. We find that the equations remain identical, just as we would find if we took $\boldsymbol{\nabla}\times(\boldsymbol{A} + \boldsymbol{\nabla}\Lambda)$ 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 $\partial_\mu \xi_\nu$) 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 here are of quadratic order, product terms like $\partial_\lambda \xi_\sigma$ contracted with $\partial^\lambda \partial^\sigma \bar{h}_{\mu\nu}$. The terms that we retain are of *linear* order, $\partial_\mu \xi_\nu$ alone. This is the right and proper thing to do. We are, after all, only working to linear order in our overall 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.' "

[Note that "longitudinal" refers to the direction parallel to the propagation of the wave; "transverse" to the axes at right angles to this direction.]

The final line of the 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 type [1] or type [2] waves, which involve at least one longitudinal component in an h-index) and those which are gauge-independent (type [3], which is completely transverse). Physical waves 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 sinuosities in the co-ordinate system," and they are in fact 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*:

$$\frac{\partial \bar{h}^\mu_\nu}{\partial x^\mu} = \frac{\partial h^\mu_\nu}{\partial x^\mu} - \frac{1}{2}\frac{\partial h}{\partial x^\nu} = 0 \tag{583}$$

In this gauge, the Field Equations (578) take the "wave-equation" form

$$\boxed{\Box \bar{h}_{\mu\nu} = -\frac{16\pi G T_{\mu\nu}}{c^4}} \tag{584}$$

How we can be sure that, even with our gauge freedom, we can find the right ξ^μ to get into a harmonic gauge and ensure the emergence of (584)? Well, if we have been unfortunate enough to be working in a gauge in which equation (583) is not satisfied, then follow equation (582) and form a new improved $h'_{\mu\nu}$, with the stipulation that our new gauge is harmonic: $\partial h'^\mu_\nu / \partial x^\mu = (1/2)\partial h'/\partial x^\nu$. We find that this implies the ξ_ν must satisfy

$$\Box \xi_\nu = \frac{\partial \bar{h}^\mu_\nu}{\partial x^\mu}, \tag{585}$$

a wave equation for ξ_ν identical in form to (584). For this equation, a solution certainly exists. Indeed, our experience with electrodynamics has taught us that the solution to the

fundamental radiation equation (584) takes the form

$$\bar{h}_{\mu\nu}(\boldsymbol{r},t) = \frac{4G}{c^4}\int \frac{T_{\mu\nu}(\boldsymbol{r}',t-R/c)}{R}\,d^3r', \qquad R \equiv |\boldsymbol{r}-\boldsymbol{r}'| \tag{586}$$

and hence a similar solution exisits for (585). The $\bar{h}_{\mu\nu}$, like their electrodynamic counterparts, are determined at time t and location $\boldsymbol{r}$ by a source intergration over $\boldsymbol{r'}$ 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. If you want to know how big $h_{\mu\nu}$ is going to be in your detector when black holes merge, simply count the number of expected Schwarzschild radii to the source and take its 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.

9.1.1 Come to think of it...

You may not have actually seen the solution (586) before, or maybe, you know, you just need a little reminding. It is important. Let's derive it.

Consider the equation

$$-\frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} + \nabla^2\Psi = -4\pi f(\boldsymbol{r},t) \tag{587}$$

We specialise to the Green's function solution

$$-\frac{1}{c^2}\frac{\partial^2 G}{\partial t^2} + \nabla^2 G = -4\pi\delta(\boldsymbol{r})\delta(t) \tag{588}$$

Of course, our particular choice of origin is immaterial, as is our zero of time, so that we shall replace $\boldsymbol{r}$ by $\boldsymbol{R} \equiv \boldsymbol{r} - \boldsymbol{r'}$ ($R \equiv |\boldsymbol{R}|$), 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 (588) by integrating over $\int e^{i\omega t}dt$ and denote the Fourier transform of G by $\tilde{G}$:

$$k^2\tilde{G} + \nabla^2\tilde{G} = -4\pi\delta(\boldsymbol{r}) \tag{589}$$

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 $k \to 0$, in which case we recover the correctly normalised $1/r$ potential of a point charge at the origin. The back transform gives

$$G = \frac{1}{2\pi r}\int_{-\infty}^{\infty} e^{\pm ikr - i\omega t}\,d\omega = \frac{1}{2\pi r}\int_{-\infty}^{\infty} e^{-i\omega(t\mp r/c)}\,d\omega \tag{590}$$

which we recognise as a Dirac delta function (remember $\omega/k = c$):

$$G = \frac{\delta(t \mp r/c)}{r} \rightarrow \frac{\delta(t - r/c)}{r} \rightarrow \frac{\delta(\tau - R/c)}{R}. \tag{591}$$

We have selected the *retarded* time solution $t - r/c$ as a nod to causality, and moved thence to $(\tau,\ R)$ variables for an arbitary time and space origin. The solution shows that a flash occurring at $t = t'$ and $\boldsymbol{r} = \boldsymbol{r'}$, produces an effect at a distance R from the flash at a time R/c later. The general solution constructed from a superposition of our Green's function is

$$\Psi(\boldsymbol{r}, t) = \int \frac{f(\boldsymbol{r'}, t')}{R} \delta(t - t' - R/c) dt' \boldsymbol{dr'} = \int \frac{f(\boldsymbol{r'}, t - R/c)}{R} \boldsymbol{dr'}. \tag{592}$$

Remember that R depends on both $\boldsymbol{r}$ and $\boldsymbol{r'}$.

9.2 Plane waves

Summary: Linear gravitational radiation comes in only two independent 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 $h_{xx} = -h_{yy}$ 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 coordindates are the most widely used for studying linear gravitational radiation.

To understand more fully the solution (586), 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,

$$R \simeq r - \boldsymbol{e_r} \cdot \boldsymbol{r'} \tag{593}$$

where $\boldsymbol{e_r}$ is a unit vector in the $\boldsymbol{r}$ direction, and

$$\bar{h}_{\mu\nu}(\boldsymbol{r}, t) \simeq \frac{e^{i(kr - \omega t)}}{r} \frac{4G}{c^4} \int T_{\mu\nu}(\boldsymbol{r'}) \exp(-i\boldsymbol{k} \cdot \boldsymbol{r'})\, d^3 r', \tag{594}$$

with $\boldsymbol{k} = (\omega/c)\boldsymbol{e_r}$, a wavenumber in the radial direction. Since r is large, this has the asymptotic form of a plane wave. Hence, both $\bar{h}_{\mu\nu}$ and $h_{\mu\nu}$ take on the form of simple plane waves, travelling in the radial outward direction, at great distances from the generating source. These plane waves have simple and important polarisation properties, which we now discuss.

9.2.1 The transverse-traceless (TT) gauge

Consider a traveling plane wave for $h_{\mu\nu}$, orienting our z axis along $\boldsymbol{k}$, so that

$$k^0 = \omega/c,\ k^1 = 0,\ k^2 = 0,\ k^3 = \omega/c \qquad \text{and} \qquad k_0 = -\omega/c,\ k_i = k^i \tag{595}$$

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

$$h_{\mu\nu} = e_{\mu\nu} a \exp(i k_\rho x^\rho) \tag{596}$$

where a is an amplitude and $e_{\mu\nu} = e_{\nu\mu}$ a symmetric polarisation tensor, again with the η's raising and lowering subscripts. Thus

$$e_{ij} = e^i_j = e^{ij} \tag{597}$$

$$e^{0i} = -e_0{}^i = e^0{}_i = -e_{0i} \tag{598}$$

$$e^{00} = e_{00} = -e^0{}_0 \tag{599}$$

The harmonic constraint

$$\frac{\partial h^\mu_\nu}{\partial x^\mu} = \frac{1}{2}\frac{\partial h^\mu_\mu}{\partial x^\nu} \tag{600}$$

implies

$$k_\mu e^\mu_\nu = k_\nu e^\mu_\mu/2. \tag{601}$$

This leads to several linear dependencies between the $e_{\mu\nu}$. For example, when $\nu = 0$ this means

$$k_0 e^0_0 + k_3 e^3_0 = k_0(e^i_i + e^0_0)/2, \tag{602}$$

or, with $k_3 = -k_0$,

$$-(e_{00} + e_{30}) = (e_{ii} - e_{00})/2. \tag{603}$$

When $\nu = j$ (a spatial index),

$$k_0 e^0_j + k_3 e^3_j = k_j(e_{ii} - e_{00})/2. \tag{604}$$

The $j = 1$ and $j = 2$ cases reduce to

$$e_{01} + e_{31} = e_{02} + e_{32} = 0, \tag{605}$$

while $j = 3$ yields

$$e_{03} + e_{33} = (e_{ii} - e_{00})/2 = -(e_{00} + e_{03}). \tag{606}$$

Equations (605) and the first=last equality of (606) yield

$$e_{01} = -e_{31},\ e_{02} = -e_{32},\ e_{03} = -(e_{00} + e_{33})/2. \tag{607}$$

Using the above expression for e_{03} in the first=second equality of (606) then gives

$$e_{22} = -e_{11}. \tag{608}$$

Of the 10 independent components of the symmetric $e_{\mu\nu}$ the harmonic condition (600) 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 (582) *within* the harmonic constraint. We can still continue to eliminate components of $e_{\mu\nu}$. In particular, let us choose

$$\xi_\mu(x) = i\epsilon_\mu \exp(ik_\rho x^\rho) \tag{609}$$

where the ϵ_μ are four constants to be chosen, and we will do so to make the calculation as simple as possible. This satisfies $\Box\xi_\mu$=0, and therefore keeps the harmonic coordinate condition, $\partial_\mu \bar{h}^\mu_\nu = 0$, intact. Following the prescription of (582), we now generate a new, but physically equivalent polarisation tensor,

$$e'_{\mu\nu} = e_{\mu\nu} + k_\mu\epsilon_\nu + k_\nu\epsilon_\mu. \tag{610}$$

By choosing the ϵ_μ 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 (610),

$$e'_{11} = e_{11}, \qquad e'_{12} = e_{12} \tag{611}$$

both remain *unchanged.* But, with $k = \omega/c$,

$$e'_{13} = e_{13} + k\epsilon_1, \quad e'_{23} = e_{23} + k\epsilon_2, \quad e'_{33} = e_{33} + 2k\epsilon_3, \quad e'_{00} = e_{00} - 2k\epsilon_0, \tag{612}$$

so that these four components may be set to zero by a simple choice of the ϵ_μ. 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 in time.

We thus see that there are only two independent modes of polarisation for a gravitational wave. The first, with only e_{11} and $e_{22} = -e_{11}$ present, is known as the "+" mode. The second, with only $e_{12} = e_{21}$ present, is the cross or "×" mode. All linear gravitational radiation is a superposition of these two modes.

9.3 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 (586) is:

$$\bar{h}^{\mu\nu}(\boldsymbol{r}, t) = \frac{4G}{rc^4} \int T^{\mu\nu}(\boldsymbol{r}', t') d^3r', \tag{613}$$

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. Start with the identity

$$0 = \int \frac{\partial(x'^j T^{ik})}{\partial x'^k} \, d^3r' = \int \left(\frac{\partial T^{ik}}{\partial x'^k}\right) x'^j \, d^3r' + \int T^{ij} \, d^3r', \tag{614}$$

where the first equality follows because the first integral reduces to a surface integration of $x'^j T^{ik}$ at infinity, where it is presumed to vanish. Thus

$$\int T^{ij} \, d^3r' = -\int \left(\frac{\partial T^{ik}}{\partial x'^k}\right) x'^j \, d^3r' = \int \left(\frac{\partial T^{i0}}{\partial x'^0}\right) x'^j \, d^3r' = \frac{1}{c}\frac{d}{dt'} \int T^{i0} x'^j \, d^3r' \tag{615}$$

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,

$$\frac{d}{dt'} \int T^{i0} x'^j \, d^3r' = \frac{d}{dt'} \int T^{j0} x'^i \, d^3r' \tag{616}$$

Continuing in this same spirit,

$$0 = \int \frac{\partial(T^{0k}x'^i x'^j)}{\partial x'^k}\, d^3r' = \int \left(\frac{\partial T^{0k}}{\partial x'^k}\right) x'^i x'^j\, d^3r' + \int (T^{0i}x'^j + T^{0j}x'^i)\, d^3r' \tag{617}$$

Using exactly the same reasoning as before,

$$\int (T^{0i}x'^j + T^{0j}x'^i)\, d^3r' = \frac{1}{c}\frac{d}{dt'}\int T^{00}x'^i x'^j\, d^3r' \tag{618}$$

Differentiating with respect to t', and using (616) and (615) gives an elegant result:

$$\int T^{ij}\, d^3r' = \frac{1}{2c^2}\frac{d^2}{dt'^2}\int T^{00}x'^i x'^j\, d^3r'. \tag{619}$$

Inserting this in (613), we obtain the *quadrupole formula* for gravitational radiation:

$$\boxed{\bar{h}^{ij} = \frac{2G}{c^6 r}\frac{d^2 I^{ij}}{dt'^2}} \tag{620}$$

where I^{ij} is the *quadrupole-moment tensor* of the energy density:

$$I^{ij} = \int T^{00}x'^i x'^j\, d^3r' \tag{621}$$

To estimate this numerically, we write

$$\frac{d^2 I^{ij}}{dt'^2} \sim Ma^2c^2\omega^2 \tag{622}$$

where M is the characteristic mass of the rotating system, a an internal separation, and ω a characteristic frequency, an orbital frequency for a binary say. Then

$$\bar{h}^{ij} \sim \frac{2GMa^2\omega^2}{c^4 r} \simeq 7\times 10^{-22}(M/M_\odot)(a_{11}^2\omega_7^2/r_{100}) \tag{623}$$

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), ω_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 3×10^{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 (623) as

$$\bar{h}^{ij} \sim \frac{2GMa^2\omega^2}{c^4 r} = \left(\frac{2.9532}{r_{km}}\right)\left(\frac{M}{M_\odot}\right)\left(\frac{a\omega}{c}\right)^2 \simeq 1\times 10^{-22}\frac{M/M_\odot}{r_{Gpc}}\left(\frac{a\omega}{c}\right)^2, \tag{624}$$

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 (620) is an *exact* solution of $\Box\bar{h}^{ij} = 0$, for any r, even if r is not large.

9.4 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.

9.4.1 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 gravitational radiation looked like this:

$$-\frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2}+\nabla^2\Phi=4\pi G\rho. \tag{625}$$

This is what a relativistic theory would look like if the source ρ were just a simple scalar quantity, rather than a component of a stress tensor. To turn this into an energy equation, 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

$$-\frac{1}{8\pi G}\frac{\partial}{\partial t}\left[\frac{1}{c^2}\left(\frac{\partial\Phi}{\partial t}\right)^2+|\boldsymbol{\nabla}\Phi|^2\right]+\boldsymbol{\nabla}\boldsymbol{\cdot}\left(\frac{1}{4\pi G}\frac{\partial\Phi}{\partial t}\boldsymbol{\nabla}\Phi\right)=\rho\frac{\partial\Phi}{\partial t}. \tag{626}$$

But

$$\rho\frac{\partial\Phi}{\partial t}=\frac{\partial(\rho\Phi)}{\partial t}-\Phi\frac{\partial\rho}{\partial t}=\frac{\partial(\rho\Phi)}{\partial t}+\Phi\boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{v})=\frac{\partial(\rho\Phi)}{\partial t}+\boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{v}\Phi)-\rho\boldsymbol{v}\boldsymbol{\cdot}\boldsymbol{\nabla}\Phi, \tag{627}$$

where $\boldsymbol{v}$ is the velocity, and the mass conservation equation

$$\frac{\partial\rho}{\partial t}+\boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{v})=0$$

has been used in the second "=" sign from the left. Combining (626) and (627), and rearranging the terms a bit leads to

$$\frac{\partial}{\partial t}\left[\rho\Phi+\frac{1}{8\pi G}\left(\frac{1}{c^2}\left(\frac{\partial\Phi}{\partial t}\right)^2+|\boldsymbol{\nabla}\Phi|^2\right)\right]+\boldsymbol{\nabla}\boldsymbol{\cdot}\left(\rho\boldsymbol{v}\Phi-\frac{1}{4\pi G}\frac{\partial\Phi}{\partial t}\boldsymbol{\nabla}\Phi\right)=\rho\boldsymbol{v}\boldsymbol{\cdot}\boldsymbol{\nabla}\Phi. \tag{628}$$

The right side is just minus the rate at which work is being done on the sources per unit volume. (Recall that the force per unit volume is $-\rho\boldsymbol{\nabla}\Phi$.) For those cases of interest when the source ρ vanishes outside a certain radius, the left side of the equation may then be interpreted as a far-field wave energy density of

$$\mathcal{E}=[(\partial_t\Phi/c)^2+|\boldsymbol{\nabla}\Phi|^2]/8\pi G,\quad\text{(Scalar gravity energy density)} \tag{629}$$

and a wave energy flux of

$$\mathcal{F}=-(\partial_t\Phi)\boldsymbol{\nabla}\Phi/4\pi G.\quad\text{(Scalar gravity energy flux)} \tag{630}$$

(Is the sign of the flux sensible for outgoing waves?) Does an analogous procedure 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.

9.4.2 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 (578):

$$\Box\bar{h}_{\mu\nu} - \frac{\partial^2\bar{h}^\lambda_\mu}{\partial x^\nu\partial x^\lambda} - \frac{\partial^2\bar{h}^\lambda_\nu}{\partial x^\mu\partial x^\lambda} + \eta_{\mu\nu}\frac{\partial^2\bar{h}^{\lambda\rho}}{\partial x^\lambda\partial x^\rho} = -\frac{16\pi G T_{\mu\nu}}{c^4}. \tag{631}$$

To facilitate the construction of a conserved flux, we need to rewrite the final term on the left side of this equation and the source term in the right. Begin by contracting on $\mu\nu$: the first term on the left becomes $\Box\bar{h}$, the second and third each become $-\partial_\lambda\partial_\rho\bar{h}^{\lambda\rho}$, while the final contraction turns $\eta_{\mu\nu}$ into a factor of 4. (Why?) This leads us to

$$\Box\bar{h} + 2\frac{\partial^2\bar{h}^{\lambda\rho}}{\partial x^\lambda\partial x^\rho} = -\kappa T, \tag{632}$$

where we have written $\kappa = 16\pi G/c^4$. Using this relation to substitute for $\partial_\lambda\partial_\rho\bar{h}^{\lambda\rho}$, we may recast our original equation as

$$\Box\bar{h}_{\mu\nu} - \frac{\partial^2\bar{h}^\lambda_\mu}{\partial x^\nu\partial x^\lambda} - \frac{\partial^2\bar{h}^\lambda_\nu}{\partial x^\mu\partial x^\lambda} - \frac{\eta_{\mu\nu}}{2}\Box\bar{h} = -\kappa S_{\mu\nu}, \tag{633}$$

where we have once again made use of the source function

$$S_{\mu\nu} = T_{\mu\nu} - \frac{\eta_{\mu\nu}T}{2}. \tag{634}$$

Equation (633) is in fact the fourth form of the linearised field equation we have encountered (one of which is in the *Exercises*), depending upon whether the fields are expressed in terms of $h_{\mu\nu}$ or $\bar{h}_{\mu\nu}$, and the source in terms of $T_{\mu\nu}$ or $S_{\mu\nu}$. For ease of reference, these have been collected and shown in figure (15).

Now, and only now, may we proceed with the same technique we used successfully for our toy problem. Contract (633) with $\partial\bar{h}^{\mu\nu}/\partial x^\sigma$, summing over μ and ν as usual but keeping σ free. We want to fold everything on the left side into a divergence. The first term on the left becomes

$$\frac{\partial\bar{h}^{\mu\nu}}{\partial x^\sigma}\Box\bar{h}_{\mu\nu} = \frac{\partial\bar{h}^{\mu\nu}}{\partial x^\sigma}\frac{\partial^2\bar{h}_{\mu\nu}}{\partial x_\rho\partial x^\rho} = \frac{\partial}{\partial x_\rho}\left(\frac{\partial\bar{h}_{\mu\nu}}{\partial x^\rho}\frac{\partial\bar{h}^{\mu\nu}}{\partial x^\sigma}\right) - \frac{\partial\bar{h}_{\mu\nu}}{\partial x^\rho}\frac{\partial^2\bar{h}^{\mu\nu}}{\partial x_\rho\partial x^\sigma}$$

$$= \frac{\partial}{\partial x_\rho}\left(\frac{\partial\bar{h}_{\mu\nu}}{\partial x^\rho}\frac{\partial\bar{h}^{\mu\nu}}{\partial x^\sigma}\right) - \frac{\partial\bar{h}_{\mu\nu}}{\partial x^\rho}\frac{\partial}{\partial x^\sigma}\frac{\partial\bar{h}^{\mu\nu}}{\partial x_\rho} = \frac{\partial}{\partial x_\rho}\left(\frac{\partial\bar{h}_{\mu\nu}}{\partial x^\rho}\frac{\partial\bar{h}^{\mu\nu}}{\partial x^\sigma}\right) - \frac{\partial}{\partial x^\sigma}\left(\frac{1}{2}\frac{\partial\bar{h}_{\mu\nu}}{\partial x^\rho}\frac{\partial\bar{h}^{\mu\nu}}{\partial x_\rho}\right) \tag{635}$$

Do you understand why the final equality is valid for the $\partial/\partial x^\sigma$ exact derivative? Here, 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 group and which is the down group.

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 side of the equation as an exact divergence. The second term on the left of (633) is

$$-\frac{\partial^2\bar{h}^\lambda_\mu}{\partial x^\nu\partial x^\lambda}\frac{\partial\bar{h}^{\mu\nu}}{\partial x^\sigma} = -\frac{\partial^2\bar{h}^{\lambda\mu}}{\partial x_\nu\partial x^\lambda}\frac{\partial\bar{h}_{\mu\nu}}{\partial x^\sigma} = -\frac{\partial}{\partial x_\nu}\left(\frac{\partial\bar{h}^{\lambda\mu}}{\partial x^\lambda}\frac{\partial\bar{h}_{\mu\nu}}{\partial x^\sigma}\right) + \frac{\partial\bar{h}^{\lambda\mu}}{\partial x^\lambda}\frac{\partial^2\bar{h}_{\mu\nu}}{\partial x^\sigma\partial x_\nu}.$$

$$\frac{\partial^2 h}{\partial x^\mu \partial x^\nu} - \frac{\partial^2 h^\lambda{}_\mu}{\partial x^\lambda \partial x^\nu} - \frac{\partial^2 h^\lambda{}_\nu}{\partial x^\lambda \partial x^\mu} + \Box h_{\mu\nu} - \eta_{\mu\nu}\left(\Box h - \frac{\partial^2 h^{\lambda\rho}}{\partial x^\lambda \partial x^\rho}\right) = -\frac{16\pi G T_{\mu\nu}}{c^4}$$

$$\Box h_{\mu\nu} - \frac{\partial^2 h^\lambda{}_\mu}{\partial x^\lambda \partial x^\nu} - \frac{\partial^2 h^\lambda{}_\nu}{\partial x^\lambda \partial x^\mu} + \frac{\partial^2 h}{\partial x^\mu \partial x^\nu} = -\frac{16\pi G}{c^4} S_{\mu\nu}$$

$$\Box \bar{h}_{\mu\nu} - \frac{\partial^2 \bar{h}^\lambda{}_\mu}{\partial x^\nu \partial x^\lambda} - \frac{\partial^2 \bar{h}^\lambda{}_\nu}{\partial x^\mu \partial x^\lambda} + \eta_{\mu\nu}\frac{\partial^2 \bar{h}^{\lambda\rho}}{\partial x^\lambda \partial x^\rho} = -\frac{16\pi G T_{\mu\nu}}{c^4}$$

$$\Box \bar{h}_{\mu\nu} - \frac{\partial^2 \bar{h}^\lambda_\mu}{\partial x^\nu \partial x^\lambda} - \frac{\partial^2 \bar{h}^\lambda_\nu}{\partial x^\mu \partial x^\lambda} - \frac{\eta_{\mu\nu}}{2}\Box \bar{h} = -\frac{16\pi G}{c^4} S_{\mu\nu}$$

Figure 15: Four equivalent forms for the linearised field equation. Recall that $\bar{h}_{\mu\nu} = h_{\mu\nu} - \eta_{\mu\nu}h/2$, $S_{\mu\nu} = T_{\mu\nu} - \eta_{\mu\nu}T/2$.

Replacing dummy-index ν with dummy-index ρ in the first group on the right side, and recognising an exact derivative in the second group on the right leads to:

$$-\frac{\partial^2 \bar{h}^\lambda_\mu}{\partial x^\nu \partial x^\lambda}\frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma} = -\frac{\partial}{\partial x_\rho}\left(\frac{\partial \bar{h}^{\lambda\mu}}{\partial x^\lambda}\frac{\partial \bar{h}_{\mu\rho}}{\partial x^\sigma}\right) + \frac{1}{2}\frac{\partial}{\partial x^\sigma}\left(\frac{\partial \bar{h}^{\lambda\mu}}{\partial x^\lambda}\frac{\partial \bar{h}_{\mu\nu}}{\partial x_\nu}\right). \tag{636}$$

The third term of equation (633) is

$$-\frac{\partial^2 \bar{h}^\lambda_\nu}{\partial x^\mu \partial x^\lambda}\frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma}.$$

But this is exactly the same case as we have just done: simply interchange the dummy indices μ and ν 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 (633) is

$$-\frac{1}{2}\frac{\partial \bar{h}}{\partial x^\sigma}\frac{\partial^2 \bar{h}}{\partial x^\rho \partial x_\rho} = -\frac{1}{2}\frac{\partial}{\partial x_\rho}\left(\frac{\partial \bar{h}}{\partial x^\rho}\frac{\partial \bar{h}}{\partial x^\sigma}\right) + \frac{1}{4}\frac{\partial}{\partial x^\sigma}\left(\frac{\partial \bar{h}}{\partial x^\rho}\frac{\partial \bar{h}}{\partial x_\rho}\right). \tag{637}$$

Thus, after dividing our fundamental equation by 2κ, the left side of equation (633) takes on a more compact form, and we obtain,

$$\frac{\partial \overline{\mathcal{T}}_{\rho\sigma}}{\partial x_\rho} = -\frac{1}{2}S_{\mu\nu}\frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma}, \tag{638}$$

where $\overline{\mathcal{T}}_{\rho\sigma}$ is a flux tensor whose general form is by now familiar, a source function for the radiation:

$$\overline{\mathcal{T}}_{\rho\sigma} = \mathcal{T}_{\rho\sigma} - \frac{1}{2}\eta_{\rho\sigma}\mathcal{T}, \tag{639}$$

and $\mathcal{T}_{\rho\sigma}$ is given by

$$\mathcal{T}_{\rho\sigma} = \frac{1}{2\kappa}\left(\frac{\partial \bar{h}_{\mu\nu}}{\partial x^\rho}\frac{\partial \bar{h}^{\mu\nu}}{\partial x^\sigma}\right) - \frac{1}{\kappa}\left(\frac{\partial \bar{h}^{\lambda\mu}}{\partial x^\lambda}\frac{\partial \bar{h}_{\mu\rho}}{\partial x^\sigma}\right) - \frac{1}{4\kappa}\left(\frac{\partial \bar{h}}{\partial x^\rho}\frac{\partial \bar{h}}{\partial x^\sigma}\right), \tag{640}$$

with $\mathcal{T} \equiv \mathcal{T}^{\rho}_{\rho}$.

We consider exclusively vacuum plane waves, which in fact form a Fourier basis for any disturbance. If we choose an harmonic gauge to work in, then from the linearised form of $R_{\mu\nu} = 0$ it follows that we must have $k^{\mu}k_{\mu} = 0$. The scalar $\mathcal{T}$ manifestly vanishes in an harmonic gauge. It can be shown, though we will not do so here, that for plane wave solutions $\mathcal{T}_{\rho\sigma}$, and therefore its scalar trace $\mathcal{T}$ are both gauge invariant. Hence, $\mathcal{T}$ *must vanish in general.* Moreover, the tensor $\mathcal{T}_{\rho\sigma} - \mathcal{T}_{\sigma\rho}$ must also be gauge invariant, vanishes for an harmonic gauage, and so must vanish in any gauge. Hence, despite appearances, $\overline{\mathcal{T}}_{\rho\sigma} = \mathcal{T}_{\rho\sigma}$ for plane wave solutions, and may be replaced by its symmetrised form:

$$\mathcal{T}_{\rho\sigma} = \frac{1}{2\kappa}\left(\frac{\partial\bar{h}_{\mu\nu}}{\partial x^{\rho}}\frac{\partial\bar{h}^{\mu\nu}}{\partial x^{\sigma}}\right) - \frac{1}{2\kappa}\left(\frac{\partial\bar{h}^{\lambda\mu}}{\partial x^{\lambda}}\frac{\partial\bar{h}_{\mu\rho}}{\partial x^{\sigma}} + \frac{\partial\bar{h}^{\lambda\mu}}{\partial x^{\lambda}}\frac{\partial\bar{h}_{\mu\sigma}}{\partial x^{\rho}}\right) - \frac{1}{4\kappa}\left(\frac{\partial\bar{h}}{\partial x^{\rho}}\frac{\partial\bar{h}}{\partial x^{\sigma}}\right). \tag{641}$$

Finally, for the TT gauge of choice, matters simplify yet more dramatically, and we find the elegant result

$$\mathcal{T}_{\rho\sigma} = \frac{c^4}{32\pi G}\left(\frac{\partial\bar{h}_{\mu\nu}}{\partial x^{\rho}}\frac{\partial\bar{h}^{\mu\nu}}{\partial x^{\sigma}}\right) \quad \text{(TT gauge)}. \tag{642}$$

Why did we choose to divide by 2κ for our overall normalisation constant? Why not just κ, or for that matter, 4κ? It is the right side of our energy equation that answers this. This is

$$-\frac{1}{2}S_{\mu\nu}\frac{\partial\bar{h}^{\mu\nu}}{\partial x^{\sigma}} = -\frac{1}{2}\left(T_{\mu\nu} - \frac{\eta_{\mu\nu}}{2}T\right)\left(\frac{\partial h^{\mu\nu}}{\partial x^{\sigma}} - \frac{\eta^{\mu\nu}}{2}\frac{\partial h}{\partial x^{\sigma}}\right) = -\frac{1}{2}T_{\mu\nu}\frac{\partial h^{\mu\nu}}{\partial x^{\sigma}}. \tag{643}$$

Choose $\sigma = 0$, the time component. We work in the Newtonian limit $h^{00} \simeq -2\Phi/c^2$, where Φ 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 $T_{\mu\nu}$. Next, we do two integrations by parts, in which we shall ignore the exact derivative. Why? Because the exact derivative of a bounded periodic function (and everything here *is* periodic) must oscillate away to zero on average. It is not a sustained energy source for the waves. By contrast, the product of two periodic functions need not oscillate to zero; for example the average of $\cos^2(\omega t) = 1/2$. This is a sustained source. (If $\cos^2(\omega t)$ is expressed as the derivative of a function, that function contains a decidedly nonperiodic growing term proportional to t and is not bounded.) We therefore keep product terms only when they are not exact derivatives. Then, using the right arrow $\rightarrow$ notation to mean "integrate by parts and ignore the pure derivatives," we perform the following manipulations on the right side of equation (643):

$$-\frac{1}{2}T_{00}\frac{\partial h^{00}}{\partial x^{0}} \rightarrow \frac{1}{2}\frac{\partial T_{00}}{\partial x^{0}}h^{00} = -\frac{1}{2}\frac{\partial T^{0i}}{\partial x^{i}}h^{00} \rightarrow \frac{1}{2}T^{0i}\frac{\partial h^{00}}{\partial x^{i}} \simeq -\rho\frac{\boldsymbol{v}}{c}\cdot\boldsymbol{\nabla}\Phi, \tag{644}$$

where the first equality follows from $\partial_{\nu}T^{0\nu} = 0$. We have therefore arrived at an expression for the rate at which the effective Newtonian potential does net work on the matter. The $1/c$ factor cancels out with the same factor on the left (flux) side of the original equation. The expression is overall *negative* if the force $-\rho\boldsymbol{\nabla}\Phi$ is oppositely directed to the velocity, so that the source is losing energy by generating outgoing waves. Our harmonic gauge expression (642) for $\mathcal{T}_{0i}$ 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, $\mathcal{T}^{0i}$ would be positive, as befits an outward moving wave energy flux.

The fact that division by 2κ 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 (642) is the true TT gauge energy flux of gravitational radiation in the weak field limit:

$$\mathcal{F}^i = \mathcal{F}_i = c\,\mathcal{T}^{i0} = -c\,\mathcal{T}_{i0} = -\frac{c^4}{32\pi G}\left(\frac{\partial \bar{h}_{\mu\nu}}{\partial x^i}\frac{\partial \bar{h}^{\mu\nu}}{\partial t}\right) \quad \text{(TT gauge)}. \tag{645}$$

9.5 The energy loss formula for gravitational waves

Summary: Einstein's classic formula for the total gravitational energy radiated by a source is $G/(5c^9)\dddot{J}_{ij}\dddot{J}_{ij}$, 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}:

$$J_{ij} = I_{ij} - \frac{\delta_{ij}}{3}I \tag{646}$$

where I is the trace of I_{ij}. Next, we address the transverse property. The projection of a vector $\boldsymbol{v}$ onto a plane perpendicular to a unit direction vector $\boldsymbol{n}$ is accomplished simply by removing the component of $\boldsymbol{v}$ along $\boldsymbol{n}$. Denoting the resulting projected vector as $\boldsymbol{w}$,

$$\boldsymbol{w} = \boldsymbol{v} - (\boldsymbol{n}\cdot\boldsymbol{v})\boldsymbol{n} \tag{647}$$

or

$$w_j = (\delta_{ij} - n_i n_j)v_i \equiv P_{ij}v_i \tag{648}$$

where we have introduced the projection tensor

$$P_{ij} = \delta_{ij} - n_i n_j,$$

which has the easily shown properties

$$n_i P_{ij} = n_j P_{ij} = 0, \qquad P_{ij}P_{jk} = P_{ik}, \qquad P_{ii} = 2. \tag{649}$$

Notice that $P_{ij}P_{jk} = P_{ik}$ simply states that if you project something that is already projected, nothing is changed! Projection operators are *idempotent*: applying them twice is the same as applying them once.

Projecting tensor components presents no difficulties,

$$w_{ij} = P_{ik}P_{jl}v_{kl} \to n_i w_{ij} = n_j w_{ij} = 0, \tag{650}$$

nor does the extraction of a projected tensor that is both traceless and transverse:

$$w_{ij}^{TT} = \left(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl}\right)v_{kl} \quad \to w_{ii}^{TT} = (P_{ik}P_{il} - P_{kl})v_{kl} = (P_{kl} - P_{kl})v_{kl} = 0. \tag{651}$$

Using the properties of P_{ij} given in (649), it is a straightforward calculation to show that the analogue of $P_{ij}P_{jk} = P_{ik}$ for the traceless transverse operator is:

$$(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl})(P_{im}P_{jn} - \frac{1}{2}P_{ij}P_{mn}) = (P_{mk}P_{nl} - \frac{1}{2}P_{kl}P_{mn}). \tag{652}$$

Again, 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. The transverse-traceless projection operator, like all projection operators, is idempotent. Recognising this simple mathematical fact in advance will save a great deal of calculational effort.

Let us define

$$J_{ij}^{TT} = \left(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl} \right) J_{kl}. \tag{653}$$

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:

$$(J_{ij} - J_{ij}^{TT})J_{ij}^{TT} = J_{ij}J_{ij}^{TT} - J_{kl}(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl})(P_{im}P_{jn} - \frac{1}{2}P_{ij}P_{mn})J_{mn} \tag{654}$$

Following (652), the right side of this is just

$$J_{ij}J_{ij}^{TT} - J_{kl}(P_{mk}P_{nl} - \frac{1}{2}P_{kl}P_{mn})J_{mn} = J_{ij}J_{ij}^{TT} - J_{kl}J_{kl}^{TT} = 0 \tag{655}$$

This will come in handy in a moment.

Next, we write down the traceless-transverse part of the quadrupole formula:

$$h_{ij}^{TT} = \frac{2G}{c^6 r}\frac{d^2 J_{ij}^{TT}}{dt'^2}. \tag{656}$$

Recalling that $t' = t - r/c$ and the J^{TT}'s are functions of t' (not t!),

$$\frac{\partial h_{ij}^{TT}}{\partial t} = \frac{2G}{c^6 r}\frac{d^3 J_{ij}^{TT}}{dt'^3}, \quad \frac{\partial h_{ij}^{TT}}{\partial r} = -\frac{2G}{c^7 r}\frac{d^3 J_{ij}^{TT}}{dt'^3} \tag{657}$$

where, in the second expression we retain only the dominant term in $1/r$. The radial flux of gravitational waves is then given by (645):

$$\mathcal{F}_r = \frac{G}{8\pi r^2 c^9}\frac{d^3 J_{ij}^{TT}}{dt'^3}\frac{d^3 J_{ij}^{TT}}{dt'^3} \tag{658}$$

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.

$$\left(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl} \right) = (\delta_{ik} - n_i n_k)(\delta_{jl} - n_j n_l) - \frac{1}{2}(\delta_{ij} - n_i n_j)(\delta_{kl} - n_k n_l) \tag{659}$$

With J_{kl} traceless, we may work this through to find

$$J_{ij}^{TT} = \left(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl}\right) J_{kl} = J_{ij} + \frac{1}{2}(\delta_{ij} + n_i n_j) n_k n_l J_{kl} - n_i n_k J_{jk} - n_j n_k J_{ik} \tag{660}$$

Now,

$$\dddot{J}_{ij}^{TT} \dddot{J}_{ij}^{TT} = [\dddot{J}_{ij} + (\dddot{J}_{ij}^{TT} - \dddot{J}_{ij})] \dddot{J}_{ij}^{TT}. \tag{661}$$

But we've seen in (655) that

$$(\dddot{J}_{ij}^{TT} - \dddot{J}_{ij}) \dddot{J}_{ij}^{TT} = 0,$$

so we are left with

$$\dddot{J}_{ij}^{TT} \dddot{J}_{ij}^{TT} = \dddot{J}_{ij} \dddot{J}_{ij}^{TT} \equiv \dddot{J}_{ij}(\dddot{J}_{ij} + \frac{1}{2}(\delta_{ij} + n_i n_j) n_k n_l \dddot{J}_{kl} - n_i n_k \dddot{J}_{jk} - n_j n_k \dddot{J}_{ik}) =$$

$$\dddot{J}_{ij} \dddot{J}_{ij} - 2 n_j n_k \dddot{J}_{ij} \dddot{J}_{ik} + \frac{1}{2} n_i n_j n_k n_l \dddot{J}_{ij} \dddot{J}_{kl} \tag{662}$$

Remember, the vector $\boldsymbol{n}$ is just the unit vector pointing in the direction of spherical angles θ and ϕ. In Cartesian coordinates:

$$\boldsymbol{n} = (n_x, n_y, n_z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).$$

The total gravitational wave luminosity is an integration of the distribution (662) over all solid angles,

$$L_{GW} = \int r^2 \mathcal{F}_r \, d\Omega = \frac{G}{8\pi c^9} \int \dddot{J}_{ij}^{TT} \dddot{J}_{ij}^{TT} \, d\Omega. \tag{663}$$

To evaluate this, you will need

$$\int n_i n_j \, d\Omega = \frac{4\pi}{3} \delta_{ij}. \tag{664}$$

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:

$$\int n_i n_j n_k n_l \, d\Omega = \frac{4\pi}{15}(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{kj}). \tag{665}$$

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 (658), (662), (664) and

(665), remembering $J_{ii} = 0$, and carrying out the angular integral, the total gravitational luminosity is thus given by

$$L_{GW} = \frac{G}{8\pi c^9} \times \left[4\pi - 2 \times \frac{4\pi}{3} + \frac{1}{2} \times \frac{4\pi}{15} \times (0+1+1) \right] \dddot{J}_{ij} \dddot{J}_{ij},$$

which amounts to:

$$\boxed{L_{GW} = \frac{G}{5c^9} \dddot{J}_{ij} \dddot{J}_{ij} = \frac{G}{5c^9} \left(\dddot{I}_{ij} \dddot{I}_{ij} - \frac{1}{3} \dddot{I}_{ii} \dddot{I}_{jj} \right).} \tag{666}$$

The calculation has not been easy. But the endpoint is a beautifully simple formula, a classical result first derived by Albert Einstein[17] 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.

9.6 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 681 below), but you will be provided with the formula.

Despite an initial burst of optimism surrounding an erroneous detection in 1968 (see §9.7), until relatively recently the direct detection of gravitational radiation seemed to most physicists a very distant possibility, at best. The coverage of the topic in Weinberg's classic text emphasises the possibility that rapidly rotating neutron stars might be a viable source. Alas, for this to occur the neutron star would have to possess a sizeable and rapidly varying quadrupole moment, and this is not an observed property of neutron stars. Even when rotating rapidly as pulsars, neutron stars would not develop a quadrupole moment that changes with time; rotation maintains strict axisymmetry.

By contrast, the possibility that *Keplerian orbits* might be interesting for measuring gravitational radiation is examined in some detail in MTW. Ordinary orbits involving ordinary stars are not a very promising source of observable gravitational radiation. But a close binary with two *compact objects* was discovered within two years of the textbook's publication, a system presciently considered by these authors. The binary, PSR 1913+16, was an unusual double neutron star system, with an orbital period of 7.75 hours. Compact object binaries have become the central focus of modern-day gravitational wave research. As noted in the Introduction, the first indirect confirmation of the existence of gravitational radiation came from radio observations of PSR 1913+16, in which a change in orbital period (from the loss of mechanical orbital energy into gravitational radiation energy) was inferred via the diminishing interval between arrival times of the pulsar's signal. The amplitude of the gravitational waves was itself well-below the threshold of any direct detection. In the range of emitted frequencies, it remains so even today. Over long enough time scales, however, a tight compact object binary will lose enough energy through gravitational radiation that the resulting inspiral goes all the way to completion, and the system either coalesces (e.g. two black holes) or explodes (e.g., GW170817: two neutron stars). This end state scenario *is* observable. Predictions suggest that there are enough merging binaries in the universe to produce a rather high detection rate: several per year at a minimum. The final frenzied

[17] Actually, Einstein found a coefficient of 1/10, not 1/5. Eddington put matters right a few years later. Tricky business, this gravitational radiation.

seconds before coalescence produce detectable gravitational wave signatures rich in physical content, at frequencies for which LIGO has been specifically designed. Through the end of 2020, there have been 20 rigorously confirmed binary merger gravitational wave events, and dozens of candidate signals under active consideration. Black hole merger waveforms can now be determined by high precision numerical calculation (but only since relatively recently: see, e.g., F. Pretorius 2005, Phys. Rev. Lett. 95, 121101).

Let us apply equation (666) 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 leading order 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 92. The separation r of the two bodies is given as a function of azimuth ϕ as

$$r = \frac{L}{1 + \epsilon \cos\phi}, \tag{667}$$

where L is the semilatus rectum and ϵ 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 the symbol J for angular momentum to avoid confusion with J_{ij}), and a the semi-major axis, we have

$$r^2 \frac{d\phi}{dt} = l, \qquad L = \frac{l^2}{GM} = a(1 - \epsilon^2), \tag{668}$$

and thus

$$\frac{d\phi}{dt} = \left(\frac{GM}{a^3(1-\epsilon^2)^3}\right)^{1/2} (1 + \epsilon\cos\phi)^2, \qquad \frac{dr}{dt} = \left(\frac{GM}{a(1-\epsilon^2)}\right)^{1/2} \epsilon \sin\phi \tag{669}$$

The distance of each body from the center-of-mass is denoted r_1 and r_2 respectively. Writing these as vector quantities,

$$\boldsymbol{r_1} = \frac{m_2 \boldsymbol{r}}{M}, \qquad \boldsymbol{r_2} = -\frac{m_1 \boldsymbol{r}}{M}. \tag{670}$$

Thus, the xy coordinates in the orbital plane are

$$\boldsymbol{r_1} = \frac{m_2 r}{M}(\cos\phi, \sin\phi), \qquad \boldsymbol{r_2} = \frac{m_1 r}{M}(-\cos\phi, -\sin\phi). \tag{671}$$

The nonvanishing moment tensors I_{ij} are then

$$I_{xx} = \frac{m_1 m_2^2 + m_1^2 m_2}{M^2} r^2 \cos^2\phi = \mu r^2 \cos^2\phi \tag{672}$$

$$I_{yy} = \mu r^2 \sin^2\phi \tag{673}$$

$$I_{xy} = I_{yx} = \mu r^2 \sin\phi\cos\phi \tag{674}$$

$$I_{ii} = I_{xx} + I_{yy} = \mu r^2 \tag{675}$$

where μ is the reduced mass $m_1 m_2/M$. (Notice that if m_1 and m_2 are very different, it is the *smaller* of the two masses that determines μ.) It is a now lengthy, but utterly straightforward task to differentiate each of these moments three times. If you choose to verify these results, you may wish to begin with the relatively easy $\epsilon = 0$ case in the course of reproducing the formulae below:

$$\frac{d^3 I_{xx}}{dt^3} = \alpha(1+\epsilon\cos\phi)^2(2\sin 2\phi + 3\epsilon\sin\phi\cos^2\phi), \tag{676}$$

$$\frac{d^3 I_{yy}}{dt^3} = -\alpha(1+\epsilon\cos\phi)^2[2\sin 2\phi + \epsilon\sin\phi(1+3\cos^2\phi)], \tag{677}$$

$$\frac{d^3 I_{xy}}{dt^3} = \frac{d^3 I_{yx}}{dt^3} = -\alpha(1+\epsilon\cos\phi)^2[2\cos 2\phi - \epsilon\cos\phi(1-3\cos^2\phi)], \tag{678}$$

where

$$\alpha^2 \equiv \frac{4G^3 m_1^2 m_2^2 M}{a^5(1-\epsilon^2)^5} \tag{679}$$

Equation (666) yields, after some assembling:

$$L_{GW} = \frac{32}{5}\frac{G^4}{c^5}\frac{m_1^2 m_2^2 M}{a^5(1-\epsilon^2)^5}(1+\epsilon\cos\phi)^4\left[(1+\epsilon\cos\phi)^2 + \frac{\epsilon^2}{12}\sin^2\phi\right] \tag{680}$$

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 $2\pi\sqrt{a^3/GM}$ to produce a *time* average. The answer is (Peters and Mathews 1963):

$$\langle L_{GW}\rangle = \frac{32}{5}\frac{G^4}{c^5}\frac{m_1^2 m_2^2 M}{a^5} f(\epsilon) = 1.00\times 10^{25}\ m_{\odot 1}^2 m_{\odot 2}^2 M_\odot (a_\odot)^{-5}\ f(\epsilon)\ \text{Watts}, \tag{681}$$

where

$$f(\epsilon) = \frac{1 + (73/24)\epsilon^2 + (37/96)\epsilon^4}{(1-\epsilon^2)^{7/2}}, \tag{682}$$

and $\odot$ indicates solar units of mass (1.99×10^{30} kg) and length (one solar radius is 6.955×10^8 m). 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 (681) tells us that L_{GW} is a number of order $E\Omega(v/c)^5$, where E is the Keplerian orbital energy, Ω 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 2π angles in ϕ. 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 (681).

Exercise. By way of comparison, we calculate the average electromagnetic power when two equal and opposite charges, e and $-e$, are in an elliptical orbit about one another. The lowest order electromagnetic *dipole* radiation is given by Larmor's formula

$$L_{EM} = \frac{2}{3c^3}\ddot{d}_i\ddot{d}_i$$

where d_i is the ith (spatial) component of the dipole moment of the charges. In vector notation, $\boldsymbol{d} = e(\boldsymbol{r_1} - \boldsymbol{r_2})$ where 1 and 2 identify each particle. (We use CGS units. For SI units, replace e here and below by $e/\sqrt{4\pi\epsilon_0}$.) Note that only the second time derivative is required for dipole EM radiation. Following the procedure we used for the gravitational wave calculation, show that the time averaged luminosity is:

$$\langle L_{EM} \rangle = \frac{2}{3} \frac{e^6}{c^3 \mu^2 a^4} f(\epsilon), \quad f(\epsilon) = \frac{1 + \epsilon^2/2}{(1-\epsilon^2)^{5/2}},$$

where μ is the reduced mass $m_1 m_2/M$. (Notation is the same as used in our *GW* problem.)

Equations (681) and (682) 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

$$\dot{P} = -\frac{192\pi}{5} \left(\frac{m_1 m_2}{M^2}\right) \left(\frac{GM}{ac^2}\right)^{5/2} f(\epsilon)$$

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 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.

9.7 Detection of gravitational radiation

9.7.1 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 $\Gamma^\mu_{\nu\lambda}$ and apply the geodesic equation,

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0 \tag{683}$$

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

$$\frac{d^2x^i}{d\tau^2} + \Gamma^i_{00}c^2 = 0 \tag{684}$$

But equation (564) clearly shows that $\Gamma^i_{00} = 0$ 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 $\delta l = h_{ik}dx^k/2$. 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 × modes, "plus" and "cross." They behave similarly, just rotated by 45°. 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 vertical prolate. The × mode ($e_{21} = e_{12}$) shows the same oscillation pattern along a rotation pattern rotated by 45°. (An excellent animation is shown in the Wikipedia article Gravitational Wave.) 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 Milky 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 really been taken very seriously up to this point. Like Adams's false detection

of a white dwarf gravitational redshift forty years before, an incorrect measurement caught the attention of the community. Post Weber, theorists and observers began to consider very seriously how one might go about detecting these signals properly, and what physics might be learnt from an unambiguous detection. Gravitational radiation detection became part of the physics mainstream, with leading figures in relativity becoming 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 LIGO design role.) Almost fifty years after Weber, the LIGO facility 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 in the tidal breeze. This truly borders on magic: using an effective length of LIGO's interferometer arm of $l = 10$ km, then δl 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 and constrained to move 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},

$$\ddot{I}_{xx} = \frac{1}{2}mc^2\omega^3 l^2 A_{xx}\sin\omega t,$$

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 (645),

$$\mathcal{F} = \frac{c^3\omega^2 A_{xx}^2}{32\pi G}.$$

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 $m = 10$kg, $l = 10$m, $\omega = 20$ rad s^{-1} (motivated by GW150914). Compare this with a typical weak interaction cross section of 10^{-48}m^2. Just how weak is gravitational scattering?

9.7.2 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 measured 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 companion 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 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 often used as standard clocks[18]. Most obligingly, Nature had placed her most accurate clock in the middle of a neutron star binary, just the system for which such fantastically precise timing is an absolute requirement. This is 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 at some unknown angle i against the plane of the sky $a \sin i$, the eccentricity e, the orbital period P, and two technical parameters related to the periastron (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° 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 $1.42\,M_\odot$. (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 will change 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 (681). 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 [16] 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 [16] is *not* a fit to the data. Rather, this 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, (681). This beautiful precision fit leaves no doubt whatsover that the quadrupole radiation formula of

[18]Since 2011, a bank of six pulsars, observed from Gdansk Poland, has been monitored continuously as a timekeeping device.

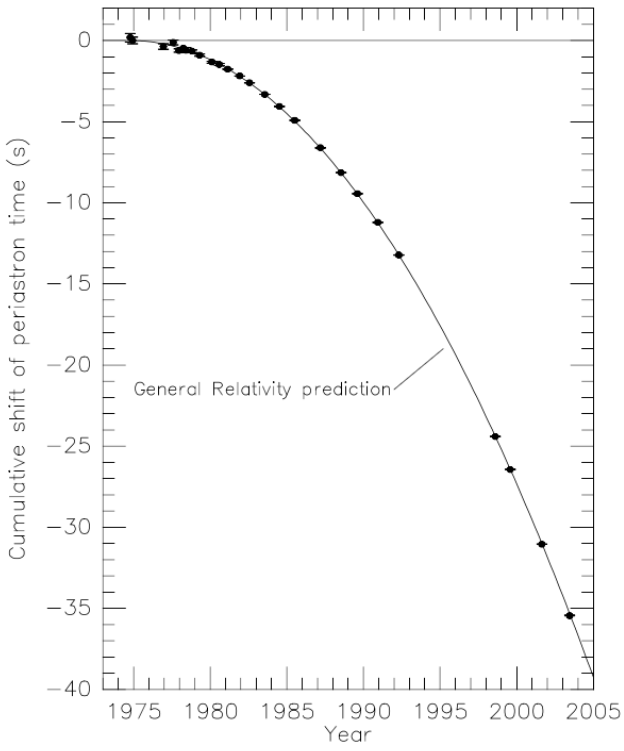


Figure 16: The cumulative change in the periastron event ("epoch") caused by the inspiral of the pulsar PSR1913+16. The dots are the data, the curve is the *prediction,* not the best fit! This prediction is confirmed to better than a fraction of a percent.

Einstein is correct. For this achievement, Hulse and Taylor won a well-deserved Nobel Prize in 1993. (It must be just a coincidence that 1993 is about the time that the observational data points become noticeably 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 interferometers are designed for wave frequencies from $\sim$ 10 Hz to 1000's of Hz. These have been detetecting sources since 2015. By contrast, the pulsar measurements are sensitive to frequencies of tens to hundreds of *micro* Hz—a very different range, produced by systems with much larger distance scales. This technique has yet to be demonstrated. The high frequency interferometers measure the gravitational radiation from stellar-mass black holes 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.

9.7.3 Direct methods: LIGO

LIGO, or Laser Interferometer Gravitational-Wave Observatory, detects gravitational waves as described in figure [17]. 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. There are two such interferometers, one in Livingston, Louisiana, the other in Hanford, Washington, a separation of 3000 km. For a signal to be verified, both detectors must show a simultaneous wave passage ("simultaneous" after taking into account an offset of 10 milliseconds for the speed of light travel time!)

This is a highly simplified description, of course. A vast amount of ingenious amplification, noise suppression, and (especially) laser frequency stabilisation techniques[19] have gone into this project, which is designed to measure induced strains at the astronishing level of one part in 10^{23}. Wave detection is possible at all only because gravitational attennae measure, not the time-averaged flux of radiation, which would have a $1/r^2$ dependence with distance r to the source, but the time-explicit amplitude $h_{ij}(t)$, which has a $1/r$ dependence.

Exercise. Following the above description, show that if one arm of the inteferometer has its phase increased by $\phi_1(t)$ due to the passage of a gravitational wave, and the other has a phase shift of $\phi_2(t)$, then the lowest order signal detected will be

$$(\phi_2 - \phi_1)\sin(\omega_c t + \delta)$$

where ω_c is the "carrier frequency" of the laser and δ is an arbitrary constant phase. The phases ϕ_1 and ϕ_2 will generally have opposite signs. Why?

Figure [18] 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

[19]The feedback system maintaining the utterly stable frequency requirements for the LIGO lasers were implemented by Ronald Drever, who passed away shortly before the LIGO Nobel Prize was announced.

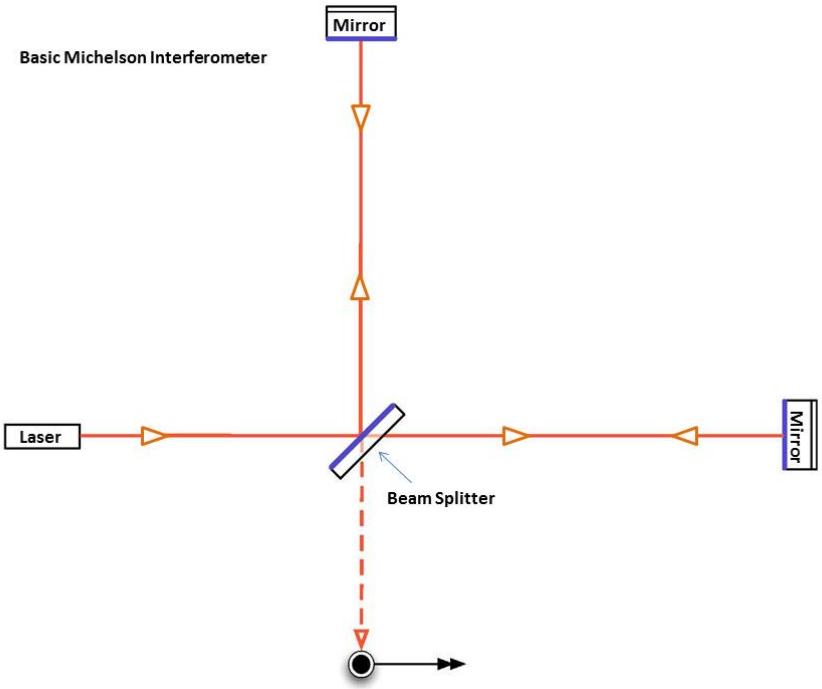


Figure 17: A schematic interferometer. Coherent light enters from the laser at the left. Half is deflected 45° upward by the beam splitter, half continues on. The two halves reflect from the mirrors. The beams re-superpose at the splitter, interfere, and are passed to a detector at the bottom. If the 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 ultra-precise destructive interference and causes laser photons to appear in the detector.

time the black holes are separated by a distance in excess of 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 §9.6, and being careful to note 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

$$\dot{P}_{GW} = -\frac{96\pi^{8/3}}{5c^5}\left(\frac{GM_c}{P_{GW}}\right)^{5/3} \tag{685}$$

where we have introduced what is known as the "chirp mass" M_c,

$$M_c = \frac{(m_1 m_2)^{3/5}}{M^{1/5}} \tag{686}$$

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

$$m_1 = m_2 \simeq 1.15 M_c.$$

Putting numbers in (685), we find

$$M_{c\odot} = 5.522 \times 10^3 P_{GW}(-\dot{P}_{GW})^{3/5}, \tag{687}$$

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

$$M_{c\odot} \simeq 30.7 \tag{688}$$

compared with "$M_c \simeq 30 M_\odot$" 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 $35 M_\odot$ 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 waves from coalescing black holes tells an interesting history, one rather well-captured even by 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 5×10^{47}J is, I believe, the largest explosion of any kind every 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 for mechanical motion ever recorded.

What a story.

Exercise. A theorem due to Stephen Hawking states that, within the realm of classical (i.e. non-quantum) general relativity, the sum of the areas of all the event horizons in any interaction involving black holes can only grow with time, never shrink. Verify that, even with a mass-energy loss of $3M_\odot c^2$ to gravitational radiation, the black hole merger discussed above is consistent with Hawking's theorem.

9.7.4 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) started to vary but the variations were correlated with one another in a mathematically predictable manner, this would be a direct indication of the deformation of space caused by the passage of a gravitational wave (fig. [19]). 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 environment is very clean. Thus, PTA has its share of skeptics. At the time of this writing (January 2022), there are only upper limits from PTA measurements.

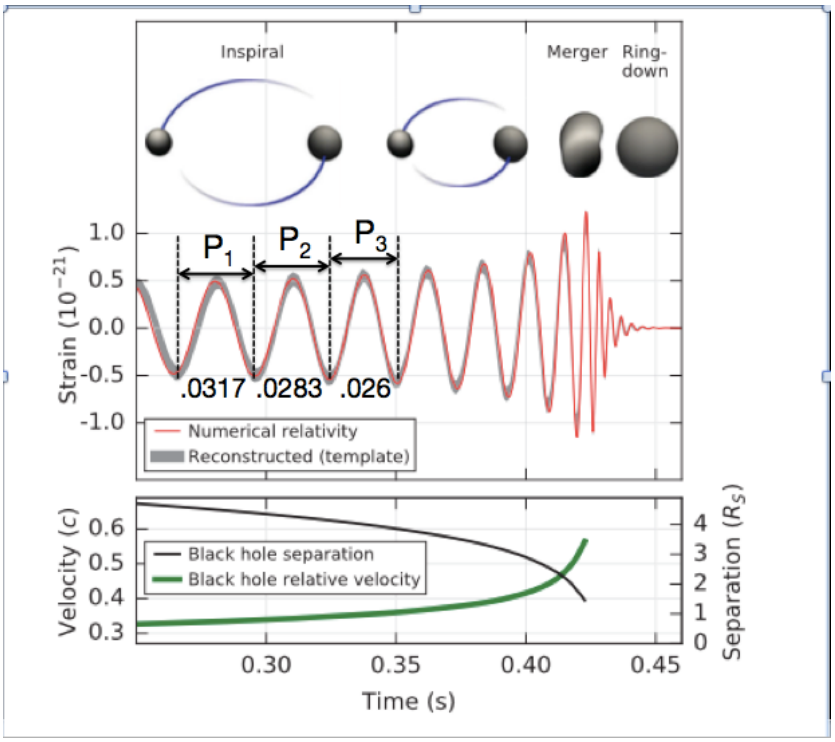


Figure 18: From Abbott et al. (2016). The upper diagram is a schematic rendering of the black hole inspiral process, from slowly evolution in a quasi-Newtonian regime, to a strongly interacting regime, followed by a coalescence and "ring-down," as the emergent single black hole settles down to its final, 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 insprial just up to the moment of coalescence.

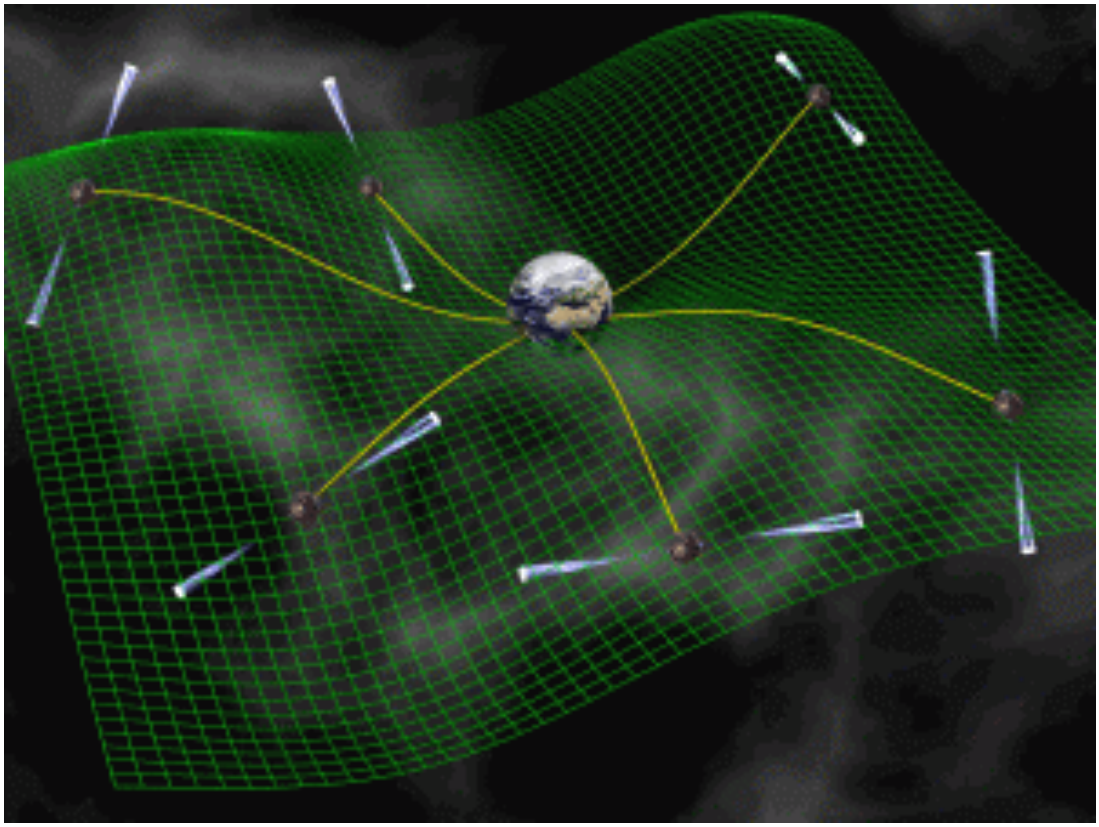


Figure 19: A schematic view of a gravitational wave passing through an array of pulsar probes, inducing regular variations in arrival times of pulses.

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

10 Cosmology

10.1 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: as measured by comoving observers, the wavelength continuously increases as the photon propagates, in direct proportion to the amount that the Universe has expanded.

10.1.1 Newtonian cosmology

The subject of the origin of the Universe is irresistible to the scientist and layperson alike. Where did the Universe come from? A Big Bang? What went bang? 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 a complete 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 place ourselves on a properly rigorous footing, we can actually get quite far using very simple ideas and hardly any relativity at all! This is an excellent way to begin. 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 might be headed, we will be in a much better position to revisit "the problem of the Universe" on a fully relativistic basis.

A plausible, if somewhat naive, first model of the Universe might be one in which space is an ordinary static Euclidian geometry, and the galaxies fill up this space uniformly (on average) everywhere out to infinity. 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 that the galaxies turned on at some finite time in the past[20]), the static Euclidian model is not even mathematically self-consistent.

Consider figure [20]. In the top diagram, 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:

[20] See W72, pp. 611-13.

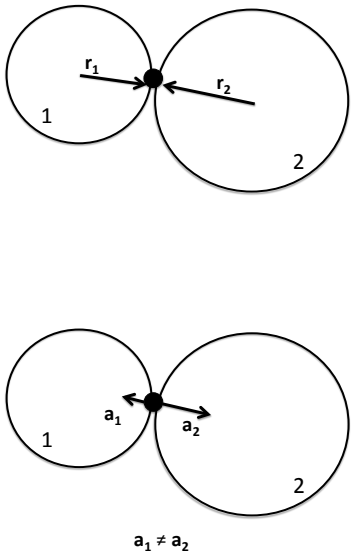


Figure 20: In a static homogeneous Euclidian universe, an observer at the center of sphere 1 would calculate a different gravitational acceleration for the dot to that of an observer at the centre of sphere 2: $\boldsymbol{a_1} \neq \boldsymbol{a_2}$. But if we take into account the actual relative acceleration 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

$$a_1 = \frac{GM(\text{within } \mathrm{r}_1)}{r_1^2} = \frac{4\pi G\rho r_1}{3}, \tag{689}$$

directed toward the centre of sphere 1. (Here ρ 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 sphere 2. This is shown in the bottom diagram of figure [20]. Both cannot be correct.

But what if the Universe is dynamically active? Then, as good Newtonian physicists, 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, provided that the Universe is allowed to expand (or contract!) and Newtonian mechanics is applied self-consistently.

Here is how it works. The observer at the centre of sphere 1 measures the acceleration of the black dot to be $-4\pi G\rho\boldsymbol{r_1}/3$ as above, with $\boldsymbol{r_1}$ pointing from the centre of sphere 1 to the surface dot. The same sphere 1 observer finds that the acceleration of the centre of sphere

2 is $-4\pi G\rho(\boldsymbol{r_1}-\boldsymbol{r_2})/3$, where $\boldsymbol{r_2}$ is the position vector oriented from the sphere 2 centre toward the big dot. Defined this way, these particular $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ 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 $-4\pi G\rho\boldsymbol{r_1}/3$ *minus* the acceleration of the observer at the centre of sphere 2:

$$-\frac{4\pi G\rho\boldsymbol{r_1}}{3}-\frac{-4\pi G\rho(\boldsymbol{r_1}-\boldsymbol{r_2})}{3}=-\frac{4\pi G\rho\boldsymbol{r_2}}{3} \tag{690}$$

Lo and behold, this is the acceleration that the observer at the centre of sphere 2 would find self-consistently in the privacy of their own study, without worrying about what anyone else thinks might be going on. A Euclidian, "linearly accelerating" universe is therefore perfectly self-consistent. The expansion process is a Newtonian requirement, not, as originally thought at the time of its discovery, a mysterious effect of general relativity. Naively, the rate of expansion ought to be slowing, since this is what gravity does: an object thrown from the surface of the Earth slows down as its upward displacement increases. This is certainly what astronomers assumed for 70 years. It turns out however, that our Universe is more devious than that: there is some real mystery here beyond the Newtonian realm.

10.1.2 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

$$r(t)=R(t)l \tag{691}$$

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. It is natural to take l to have dimensions of length, in which case $R(t)$ is a dimensionless function of time alone, a scale factor that, via this time dependence, embodies the dynamical behaviour of the Universe. The velocity $v=dr/dt\equiv\dot{r}$ of a comoving point expanding with space is then

$$v(t)=\dot{R}l=(\dot{R}/R)r. \tag{692}$$

We should emphasise the vector character of this relationship:

$$\boldsymbol{v}(t)=(\dot{R}/R)\boldsymbol{r} \tag{693}$$

where $\boldsymbol{r}$ is a vector pointing outward from our arbitrarily chosen origin. Then, the acceleration is

$$\boldsymbol{a}(t)=\frac{d\boldsymbol{v}}{dt}=(\ddot{R}/R)\boldsymbol{r}. \tag{694}$$

(Why didn't we differentiate $\boldsymbol{r}(t)/R(t)$?) This kinematic relative acceleration between two points is, by Newtonian gravitation theory, given by equation [689]):

$$\boldsymbol{a}(t)=-\frac{4\pi G\rho}{3}\boldsymbol{r}. \tag{695}$$

Hence,

$$-\frac{4\pi G\rho}{3}=\frac{\ddot{R}}{R}. \tag{696}$$

Notice how l disappears: this is an equation for R only, 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. ρR^3 is constant. We then obtain

$$\dot{R}^2 - \frac{8\pi G\rho R^2}{3} = 2E, \tag{697}$$

where E is an energy-like integration constant. In a simple, apparently naive Euclidian-Newtonian approach, this would be our fundamental dynamical cosmological equation for the evolution of the Universe. Amazingly, providing that we are prepared to allow the mass density ρ 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.

10.1.3 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 more distant observers that are moving away from the original source. If the photon had a measured wavelength λ at location r, then when the photon passes an observer a distance $dr = cdt$ away from r, who is moving at a relative velocity $\dot{R}\,dr/R$, this observer measures a Doppler change in wavelength $d\lambda$ determined by equation (693):

$$\frac{d\lambda}{\lambda} = \frac{v}{c} = \frac{\dot{R}}{cR}dr = \frac{\dot{R}}{R}dt, \tag{698}$$

or in other words

$$\frac{1}{\lambda}\frac{d\lambda}{dt} \equiv \frac{\dot{\lambda}}{\lambda} = \frac{\dot{R}}{R}. \tag{699}$$

Solving for λ, 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 to the radiation temperature $T_\gamma \propto 1/R$. The energy density in radiation (fourth power in T_γ) decreases as $1/R^4$. Evidently the total energy in a comoving volume R^3 decreases as $1/R$: energy is *not* conserved by the expansion of the Universe! By contrast, the *entropy* density in the radiation is proportional to T_γ^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 as we go back in time. If a photon is emitted with a wavelength λ_e at some time t in the past, the wavelength we would now measure at time t_0 is formally given by

$$\lambda(t_0) = \lambda_0 = \lambda_e(1+z), \tag{700}$$

where z is defined by this equation and known as the redshift parameter. Therefore,

$$\frac{\lambda_0}{\lambda_e} = 1 + z = \frac{R(t_0)}{R(t)} = \frac{1}{R(t)}. \tag{701}$$

The advantage of using z, as opposed to the more geometrical quantity R, is that z is directly observed by astronomers in their instruments. But the two quantities are mathematically equivalent via this completely general equation, $R = 1/(1+z)$. In *any* symmetric cosmological model, for example, if you measure a redshift of 2, it has come from a time when the Universe had one-third of its current size.

10.2 Cosmology models for the impatient

Summary. The concept of a spacetime metric is needed 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 emerged, given a measured redshift z. Because different cosmologies produce different functions for this $l(z)$, this quantity is a key observational tool, which in principle allows us to determine what kind of Universe we live in.

10.2.1 The large-scale spacetime metric

Although the Newtonian approach works rather well, let us remember that 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 proper speed of light is a proper universal constant c 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 everywhere. If space itself is Euclidian, a suitable spacetime metric practically leaps out of the page,

$$-c^2 d\tau^2 = -c^2 dt^2 + R^2(dx^2 + dy^2 + dz^2), \tag{702}$$

where we have used here the usual (x, y, z) Cartesian coordinates, and R satisfies equation (697). Note that x, y, z are *comoving* with the expansion, with an embedded Euclidian comoving spatial interval dl given by

$$dx^2 + dy^2 + dz^2 = dl^2.$$

In particular, a photon heading directly towards us satisfies the equation:

$$R\frac{dl}{dt} = -c \tag{703}$$

where dl is the change of radial comoving coordinate induced by the photon's passage. A useful mechanical analogue is an ant moving at a constant velocity c, crawling from the pole to the equator along the surface of an expanding globe. In this case, dl would correspond to the change in the ant's latitude.

Second, we must allow for the theoretical possibility of non-Euclidian spatial structure. Now, equation (702) 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 flat 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 on the surface exceeds 180 degrees.) The nature of a two-dimensional spherical surface is readily grasped because we can easily embed it in three dimensions to form a mental image—or just look at a globe. 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 smoothly embedded in three Euclidian dimensions (a result proved by Hilbert), so it is very hard to picture in your mind's eye! There are positively and negatively curved *three-dimensional* spaces as well, all perfectly symmetric, all logically possible alternative structures for the space of our Universe. They just happen not to fit the observations. It is fortunate for us that the real Universe seems also to be mathematically the simplest in its geometry. For now, we shall confine our attention to this expanding, familiar flat Euclidian space.

Third, we need relativity in the form of the Birkhoff theorem to properly justify neglecting exterior contributions from outside the arbitrarily chosen spheres we invoked in section 10.1. The Newtonian description cannot be strictly applied to an infinite system, whereas nothing prevents us from using Birkhoff's theorem applied to an unbounded symmetric spacetime.

Fourth, at some point we shall require a relativistically valid argument to arrive at equation (697). Nothing in the Newtonian derivation hints at this level of generality. We shall return to this calculation in section 10.3.

Finally, we need relativity theory to relate the algebraic 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.

10.2.2 The Einstein-de Sitter universe: a classic toy model

Consider equation (697) for the case $E = 0$ in the presence of ordinary matter and nothing else, which means that ρR^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 ρR^3 to be equal to $\rho(t_0) = \rho_{M0}$, the mass density at the current epoch. Equation (697) becomes

$$R^{1/2}\dot{R} = \left(\frac{8\pi G\rho_{M0}}{3}\right)^{1/2}. \tag{704}$$

Integrating,

$$\frac{2}{3}R^{3/2} = \left(\frac{8\pi G\rho_{M0}}{3}\right)^{1/2} t, \tag{705}$$

where the integration constant has been set to zero under the assumption that R was very small at early times. This leads to

$$R = \left(\frac{3H_0 t}{2}\right)^{2/3}, \tag{706}$$

where we have introduced an important structural constant of this so-called Einstein-de Sitter universe, H_0, the value of $\dot{R}$ at the current time t_0. This is known as the Hubble

constant. In the example at hand,

$$H_0 = \dot{R}(t_0) = \left(\frac{8\pi G\rho_{M0}}{3}\right)^{1/2} \rightarrow \dot{R} = H_0 R^{-1/2}. \tag{707}$$

More generally, for any model uiverse, we define the Hubble *parameter* as the logarithmic derivative of $R(t)$:

$$H(t) = \frac{\dot{R}}{R} \tag{708}$$

for arbitrary time t. This definition holds independently of the normalisation of R; with $R(t_0) = 1$ the Hubble constant $H_0 = \dot{R}_0$ as seen in equation (707).

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 by measuring the distances to nearby galaxies as well as their redshift, and then using equation (693) locally. As a problem in observational astronomy, this is hardly straighforward! 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, obtaining good distance measurments to ever more distant galaxies, and the bottom line is that *the measured value of H_0 and the measured value of the density of ordinary matter ρ_{M0} do not satisfy (707) in our Universe, even allowing room for reasonable error.* There is simply not enough ordinary matter ρ_{M0} to account for the measured H_0. On the other hand, equation (697) 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 of *ordinary* matter most certainly dominates over radiation, it is *not* the dominant contributor to the energy density of the Universe. There is now strong evidence of an additional energy density associated with the vacuum of spacetime itself! This energy density, ρ_V, is the dominant energy density of the real Universe on cosmological scales, though not at present overwhelmingly so: ρ_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 ρ_V remains constant as the Universe expands, at later times vacuum energy will completely dominate the expansion: ρ_M drops off as $1/R^3$, and the constant ρ_V takes up the slack.

What does this imply for the expansion of the Universe? With an effective vacuum Hubble parameter of

$$H_V \equiv \left(\frac{8\pi G\rho_V}{3}\right)^{1/2}, \tag{709}$$

equation (697) at later times takes the form

$$\dot{R} = H_V R \tag{710}$$

or

$$R \propto \exp(H_V t). \tag{711}$$

The Universe will expand exponentially! In other words, rather than gravity slowly decelerating the expansion by the mutual attractive force of Newtonian gravity, the vacuum energy density will actively drive an ever more vigorous expansion via a repulsive force. The converse of this is that an exponential Universe would be expanding less rapidly in the past than in the present. It is this "less rapidly than expected" earlier history of cosmic expansion that has led to our current understanding of the remarkable dynamical behaviour of our 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 happens to be taking over.

Let us return to the Einstein-de Sitter Universe. Equations (701), (703) and (707) 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 $R(t) = (t/t_0)^{2/3} = (3H_0t/2)^{2/3}$, equation (703) may be integrated over the path of the photon from its emission at l at time t:

$$\int_0^l dl' = -c\int_{t_0}^{t}\frac{dt'}{R} = c\int_R^1 \frac{dR'}{\dot{R}'R'} = \frac{2c}{H_0}\left(1-\sqrt{R}\right) = \frac{2c}{H_0}\left(1-\frac{1}{\sqrt{1+z}}\right) \equiv l(z) \qquad (712)$$

Note that as $z\to\infty$, $l\to 2c/H_0$, a constant. The most distant photons—and therefore the most distant regions of the Universe that may causally influence us today—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. On dimensional grounds it will generally be of order 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 the radial extent throughout which cosmological-scale causality may be exerted. The event horizon is a surface enveloping a black hole, within which no causal contact with the outside world is possible.

Physical distances are given not by comoving coordinate l, but by Rl. Since $R = 1$ currently, $2c/H_0$ is the physical scale of the horizon today. 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. As a practical matter, the farthest back in time that we can see is to a redshift of about $z = 1100$. At higher redshifts, the Universe was completely opaque to photons. Just as in viewing the Sun we can see only down to its photosphere and not into its interior, we can see in cosmological time only back to an era when the Universe itself was becoming similarly opaque. What would be the subtended *angular* size on the sky *today* of the horizon scale at a redshift $z = 1100$? 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 larger 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:

$$\Theta_H(z) = \frac{2c}{H_0(1+z)^{3/2}}\times\frac{1}{R(z)l(z)} = \frac{2c}{H_0(1+z)^{1/2}l(z)} = \frac{1}{\sqrt{1+z}-1} \qquad (713)$$

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 precisely the same earlier time corresponding to redshift* z (not now!). This d is the relevant proper distance to the photon source at the moment of the radiation emission when the proper size was $2c/H(t)$. The Universe was a smaller place then, and it would be inconsistent to compute the angle $\Theta_H(z)$ using the *current* proper distance $R_0l(z) = l(z)$, with the *old* proper horizon size. If you plug in $z = 1100$ into (713) and convert Θ_H to degrees, you'll find $\Theta_H = 1.8°$, a bit more than 3 times the diameter of the full moon. But the actual Universe looks very regular on much larger angular scales. Indeed, we know that the radiation temperature over the entire

sky is unchanging to one part in 10^5. Even if our E-dS model is only crude, it highlights an important problem. Simply put, how does the Universe know about itself in such a globally coordinated manner, given that it takes signals travelling at the speed of light much longer than the age of the Universe to cross it? *What we've got here is a failure to communicate.* We will see later in this course how modern cosmology can address this puzzle and facilitate communication.

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 proper distance to redshift z formula is (for E-dS models):

$$l(z)R(z) = \frac{2c}{H_0(1+z)^{3/2}}\left(\sqrt{1+z}-1\right) \tag{714}$$

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: $z = 5/4$.)

10.3 The Friedmann-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 Friedmann-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 change of notation. *The variable r will henceforth denote a comoving radial coordinate in expanding spatial geometries.* To avoid any confusion, we will no longer use l as a comoving coordinate.

The Friedmann-Robertson-Walker, or FRW, metric is the most general metric for a homogeneous isotropic cosmology. To understand its mathematical form, we start simply. The ordinary metric for a planar two-dimensional space ("2-space") without curvature may be written in cylindrical coordinates as

$$ds^2 = d\varpi^2 + \varpi^2 d\phi^2 \tag{715}$$

where the radial ϖ and angular ϕ polar coordinates are related to ordinary Cartesian x and y coordinates by the familiar formulae:

$$x = \varpi\cos\phi, \qquad y = \varpi\sin\phi \tag{716}$$

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 and still retain the symmetry of every point being mathematically equivalent. The metric for such a spherical surface is

$$ds^2 = a^2 d\theta^2 + a^2\sin^2\theta d\phi^2, \tag{717}$$

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 θ in simple geometrical terms as the colatitude angle measured downward from the z axis, so that this ϖ is actually the same cylindrical radius ϖ seen in equation (715). But don't expect to get back to (715) from (717) via this simple coordinate change $\sin\theta = \varpi/a$. Instead, we find that the spherical surface metric becomes:

$$ds^2 = \frac{d\varpi^2}{1-\varpi^2/a^2} + \varpi^2 d\phi^2, \tag{718}$$

an entirely different space from the flat planar surface (715). The space of (718) is the same original spherical surface space we started with by any other name; the metric resembles equation (715) for a plane only in the limit $a \to \infty$. 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

$$ds^2 = \frac{d\varpi^2}{1+\varpi^2/a^2} + \varpi^2 d\phi^2, \tag{719}$$

which upon substituting $\sinh\chi = \varpi/a$ reverts to

$$ds^2 = a^2 d\chi^2 + a^2\sinh^2\chi\, d\phi^2 \tag{720}$$

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.

Exercise. Show that the Ricci scalar R for the spherical metric (717) is $-2/a^2$, and for the saddle metric (720) is $2/a^2$. These are constants of equal magnitudes across their respective surfaces, but with opposite signs. The underlying homogeneous and isotropic symmetry is the same in both cases.

10.3.1 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

$$-c^2 d\tau^2 = -c^2 dt^2 + R^2(t)\left(\frac{dr^2}{1-r^2/a^2} + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2\right) \tag{721}$$

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-ball" (Cartesian coordinates w, x, y, z) is given by

$$w^2 + x^2 + y^2 + z^2 = a^2 = \text{constant} > 0. \tag{722}$$

Thus, on the three-dimensional surface ("3-sphere"), a small change in w^2 is constrained to satisfy:

$$d(w^2) = -d(x^2 + y^2 + z^2) \equiv -d(r^2), \text{ whence } \quad dw = -rdr/w \tag{723}$$

with $r^2 \equiv x^2 + y^2 + z^2$. Therefore,

$$(dw)^2 = \frac{r^2(dr)^2}{w^2} = \frac{r^2(dr)^2}{a^2 - r^2} \tag{724}$$

Next, the line element in Cartesian 4-space is

$$ds^2 = dx^2 + dy^2 + dz^2 + dw^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2 + dw^2. \tag{725}$$

Adding together dr^2 and $(dw)^2$ as given by (724), there follows immediately the desired line element associated with the 3-sphere of a 4-ball, analogous to our expression for the ordinary spherical surface metric (717):

$$ds^2 = \frac{dr^2}{1 - r^2/a^2} + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2, \tag{726}$$

as seen in equation (721).

An alternative form for (726) is sometimes useful without the singular denominator. Set $r = a \sin\chi$. Then

$$ds^2 = a^2 d\chi^2 + a^2 \sin^2\chi d\theta^2 + a^2 \sin^2\chi \sin^2\theta d\phi^2 \tag{727}$$

Exercise. Show that the 3-sphere (surface) of a 4-ball 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-sphere of a 5-ball looks like either in the form of (726) or (727)?

What if we began, not with the 3-sphere of a Euclidian 4-ball, but with the 3D surface of a 4D hyperboloid,

$$x^2 + y^2 + z^2 - w^2 = a^2 = constant \tag{728}$$

in a flat Minkowski-like space? The line element is

$$ds^2 = dx^2 + dy^2 + dz^2 - dw^2. \tag{729}$$

Why not? Special relativity is not the only venue for hyperbolic geometry.

An identical calculation constraining dw to stay within the 3D surface now yields

$$(dw)^2 = \frac{r^2(dr)^2}{w^2} = \frac{r^2(dr)^2}{a^2 + r^2}. \tag{730}$$

The four-dimensional Minkowski metric (729) constrained within our 3D surface takes the form

$$ds^2 = dx^2 + dy^2 + dz^2 - dw^2 = \frac{dr^2}{1 + r^2/a^2} + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2, \tag{731}$$

This is the corresponding negatively curved (saddle-like) 3D surface. With $r = a \sinh\chi$, its alternative form is

$$ds^2 = a^2 d\chi^2 + a^2 \sinh^2\chi d\theta^2 + a^2 \sinh^2\chi \sin^2\theta \, d\phi^2. \tag{732}$$

For our Universe, a appears to be immeasurably large, but with χ compensatingly small so that $a\chi$ is finite. In other words, space is neither positively nor negatively curved. Space is very close to 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 choose this route, 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:

$$-c^2 d\tau^2 = -c^2 dt^2 + R^2(t)\left(\frac{dr^2}{1-kr^2} + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2\right), \tag{733}$$

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

$$g_{00} = -1, \quad g_{rr} = R^2(t)/(1-kr^2), \quad g_{\theta\theta} = R^2(t) r^2, \quad g_{\phi\phi} = R^2(t) r^2 \sin^2\theta. \tag{734}$$

is often what people mean when they write "the FRW metric." Because a is absent in this formulation, this can be mathematically convenient in calculations. In my view, the form (721) is more physically transparent.

Be careful. In (733), R has dimensions of length, and r is *dimensionless.* In (721), 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." 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 (721). 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

$$\int_0^r \frac{dr'}{\sqrt{1 - r'^2/a^2}} = a \sin^{-1}(r/a) \quad (a^2 > 0), \qquad \text{or} \qquad |a| \sinh^{-1}(r/|a|) \quad (a^2 < 0).$$

(In the above, a is positive when $a^2 > 0$, and $|a|$ is the positive norm when $a^2 < 0$.)

By way of a specific example, let us rework equation (712) 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

$$a \sin^{-1}(r/a) = c\int_t^{t_0} \frac{dt'}{R(t')} = c\int_R^1 \frac{dR'}{\dot{R}' R'} = c\int_0^z \frac{dz'}{(1+z')\dot{R}(z')} \equiv a \sin^{-1}(r[z]/a)$$

Note that we have used $R = 1/(1+z)$, so that the current value of R is 1, which is appropriate for the form of the metric in equation (721). $\dot{R}$ is given by equation (697) quite generally, but we need to relate E and a. *That* requires the field equations. In §10.4, we show that $2E = -c^2/a^2$.

Let us next look at a flat, pure vacuum universe, but one with finite ρ_V and H_V so that we have something to work with! Then $\dot{R} = H_V R$ and

$$r = c\int_R^1 \frac{dR'}{\dot{R}' R'} = \frac{c}{H_V}\int_R^1 \frac{dR'}{R'^2} = \frac{c}{H_V}\left(\frac{1}{R} - 1\right) = \frac{cz}{H_V} \tag{735}$$

How interesting. There is no horizon in such a universe. As $z \to \infty$, $r \to \infty$. 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 in a maximally symmetric space within redshift z_m is formally given by

$$V = 4\pi \int_0^r R^3 \frac{r'^2 dr'}{\sqrt{1 - r'^2/a^2}}, \tag{736}$$

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, let us say, a prescribed group of photon sources, then $R = R_0$, a constant, and life is simple. If we are interested in the net volume occupied of the sources *at the time of the photon 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 ($R = R_0 = 1$)

$$V = 4\pi \int_0^r \frac{r'^2 dr'}{\sqrt{1 - r'^2/a^2}} = 2\pi \left(a^3 \sin^{-1} \frac{r}{a} - a^2 r \sqrt{1 - \frac{r^2}{a^2}} \right). \tag{737}$$

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!):

$$V(z) = \frac{32\pi}{3} \left(\frac{c}{H_0} \right)^3 \left(1 - \frac{1}{\sqrt{1+z}} \right)^3. \tag{738}$$

10.4 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. The field equations can be modified by the addition of what is known as the cosmological constant, a term whose effect is very important on the largest scales. This cosmological constant is mathematically identical to the effect of a vacuum energy density. We may therefore, without loss of generality, keep the original form of the field equations, but allow for the possibility that the energy density ρc^2 contains a contribution from the vacuum.

The student should read and study subsections 10.4.1 and 10.4.3, but may regard the rather heavy 10.4.2 as optional (off syllabus). Remember, however, the identification of $2E$ with $-c^2/a^2$!

10.4.1 The effect of a cosmological constant

Begin first with the field equations *including the cosmological constant,* equation (294):

$$R_{\mu\nu} - \frac{g_{\mu\nu}}{2} R = -\frac{8\pi G}{c^4} T_{\mu\nu} + \Lambda g_{\mu\nu} \tag{739}$$

Recall the stress energy tensor for a perfect fluid:

$$T_{\mu\nu} = P g_{\mu\nu} + (\rho + P/c^2) U_\mu U_\nu. \tag{740}$$

We may arrange the right side source term of (739) as follows:

$$-\frac{8\pi G}{c^4}T_{\mu\nu}+\Lambda g_{\mu\nu}=-\frac{8\pi G}{c^4}\left[\tilde{P}g_{\mu\nu}+\left(\tilde{\rho}+\frac{\tilde{P}}{c^2}\right)U_\mu U_\nu\right], \tag{741}$$

where

$$\tilde{P}=P-\frac{c^4\Lambda}{8\pi G},\quad \tilde{\rho}=\rho+\frac{c^2\Lambda}{8\pi G} \tag{742}$$

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

$$T_{\mu\nu\ (V)}=P_V g_{\mu\nu}+(\rho_V+P_V/c^2)U_\mu U_\nu, \tag{743}$$

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. This 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 famously commented that "To drop the cosmical constant would be to knock the bottom out of space." At the time, this was probably viewed as Eddington in his customary curmudgeon mode. Today, his insight seems quite prescient. Contemporary physicists think less in terms of a cosmological constant, placing greater conceptual emphasis instead on the notion of a vacuum energy density. Many questions then arise. What is the reason for its existence? Why does it have the value that it does? If ρ_V is not strictly constant, there must be new physics to be discovered. Are the best observational data supportive of a truly constant value for ρ_V? The value of ρ_V probably emerged from the same type of "cancellation process" that has somehow produced finite values for the masses for the fundamental particles. How can we calculate or even estimate this? These are some of the most difficult questions physics.

Moving forward, for now we must place these fascinating issues to the side, and continue our development of large scale models of the Universe. We do so without the formal appearance of a cosmological constant, but with the understanding that we may add the appropriate contributions to the density ρ and pressure P to account precisely for the effects of Λ.

10.4.2 Formal analysis

Let us recall the field equations written in terms of the source function $S_{\mu\nu}=T_{\mu\nu}-g_{\mu\nu}T/2$"

$$R_{\mu\nu}=-\frac{8\pi G}{c^4}S_{\mu\nu}, \tag{744}$$

and the FRW metric in the compact form of equation (734). We shall lose no generality in doing so, as we may always set $k = 1/a^2$ at the end of the calculation, without invalidating any step along the way. In comoving coordinates, the only component of the covariant 4-velocity U_μ 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

$$T_{00} = \rho c^2, \quad T_{ij} = P g_{ij}. \tag{745}$$

This means

$$T^0_0 = -\rho c^2, \quad T^i_j = \delta^i_j P, \quad T = T^\mu_\mu = -\rho c^2 + 3P. \tag{746}$$

We shall also need

$$S_{00} = T_{00} - \frac{g_{00}}{2}T = \frac{1}{2}(\rho c^2 + 3P), \quad S_{ij} = \frac{1}{2}g_{ij}(\rho c^2 - P) \tag{747}$$

To calculate R_{00}, begin with the Ricci tensor from equation (315):

$$R_{\mu\kappa} = \frac{1}{2}\frac{\partial^2 \ln|g|}{\partial x^\kappa \partial x^\mu} - \frac{\partial \Gamma^\lambda_{\mu\kappa}}{\partial x^\lambda} + \Gamma^\eta_{\mu\lambda}\Gamma^\lambda_{\kappa\eta} - \frac{\Gamma^\eta_{\mu\kappa}}{2}\frac{\partial \ln|g|}{\partial x^\eta}$$

where g is the determinant of the metric tensor $g_{\mu\nu}$ given by (734). Defining

$$f(r) = 1/(1 - kr^2),$$

we have

$$|g| = R^6(t)\, r^4\, f(r) \sin^2\theta \tag{748}$$

For diagonal metrics, $\Gamma^a_{ab} = \partial_b(\ln g_{aa})/2$ and $\Gamma^b_{aa} = -(\partial_b g_{aa})/2g_{bb}$ (no sum on a). Therefore $\Gamma^\lambda_{00} = 0$, and our expression for R_{00} simplifies to

$$R_{00} = \frac{1}{2}\frac{\partial^2 \ln|g|}{\partial x^0 \partial x^0} + \Gamma^\eta_{0\lambda}\Gamma^\lambda_{0\eta} = \frac{1}{2}\frac{\partial^2 \ln|g|}{\partial x^0 \partial x^0} + \Gamma^r_{0r}\Gamma^r_{0r} + \Gamma^\theta_{0\theta}\Gamma^\theta_{0\theta} + \Gamma^\phi_{0\phi}\Gamma^\phi_{0\phi} \tag{749}$$

With

$$\frac{1}{2}\frac{\partial^2 \ln|g|}{\partial x^0 \partial x^0} = \frac{3\ddot{R}}{Rc^2} - \frac{3\dot{R}^2}{R^2c^2}, \quad \Gamma^r_{0r} = \Gamma^\theta_{0\theta} = \Gamma^\phi_{0\phi} = \frac{\dot{R}}{Rc}, \tag{750}$$

the 00 component of (744) becomes

$$\ddot{R} = -\frac{4\pi G R}{3}\left(\rho + \frac{3P}{c^2}\right) \tag{751}$$

which differs from our Newtonian equation (696) only by the additional term, seemingly very small, of $3P/c^2$ as an additional 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 in this chapter.

The rr component of the field equations is a bit more involved. With $f' \equiv df/dr$, we prepare a working table in advance of all the results we will need:

$$R_{rr} = \frac{1}{2}\frac{\partial^2 \ln|g|}{\partial r^2} - \frac{\partial \Gamma^\lambda_{rr}}{\partial x^\lambda} + \Gamma^\eta_{r\lambda}\Gamma^\lambda_{r\eta} - \frac{1}{2}\Gamma^\eta_{rr}\frac{\partial \ln|g|}{\partial x^\eta} \tag{752}$$

$$\Gamma^r_{rr} = \frac{1}{2g_{rr}}\frac{\partial g_{rr}}{\partial r} = \frac{f'}{2f}, \quad \Gamma^0_{rr} = -\frac{1}{2cg_{00}}\frac{\partial g_{rr}}{\partial t} = \frac{fR\dot{R}}{c}, \quad \Gamma^r_{0r} = \frac{1}{2cg_{rr}}\frac{\partial g_{rr}}{\partial t} = \frac{\dot{R}}{Rc} \tag{753}$$

$$\Gamma^\phi_{\phi r} = \frac{1}{2g_{\phi\phi}}\frac{\partial g_{\phi\phi}}{\partial r} = \Gamma^\theta_{\theta r} = \frac{1}{2g_{\theta\theta}}\frac{\partial g_{\theta\theta}}{\partial r} = \frac{1}{r} \tag{754}$$

$$\frac{\partial \ln|g|}{\partial r} = \frac{f'}{f} + \frac{4}{r}, \quad \frac{1}{2}\frac{\partial^2 \ln|g|}{\partial r^2} = \frac{1}{2}\frac{f''}{f} - \frac{1}{2}\frac{(f')^2}{f^2} - \frac{2}{r^2}, \quad \frac{\partial \Gamma^0_{rr}}{\partial x^0} = \frac{1}{c^2}\left(f\dot{R}^2 + fR\ddot{R}\right) \tag{755}$$

$$\frac{\partial \ln|g|}{\partial t} = \frac{6\dot{R}}{R}, \quad \frac{\partial \Gamma^r_{rr}}{\partial r} = \frac{f''}{2f} - \frac{(f')^2}{2f^2} \tag{756}$$

Putting it all together:

$$R_{rr} = \frac{1}{2}\frac{\partial^2 \ln|g|}{\partial r^2} - \frac{\partial \Gamma^r_{rr}}{\partial r} - \frac{\partial \Gamma^0_{rr}}{\partial x^0} + (\Gamma^r_{rr})^2 + \left(\Gamma^\theta_{\theta r}\right)^2 + \left(\Gamma^\phi_{\phi r}\right)^2 + 2\Gamma^0_{rr}\Gamma^r_{r0} - \frac{\Gamma^r_{rr}}{2}\frac{\partial \ln|g|}{\partial r} - \frac{\Gamma^0_{rr}}{2c}\frac{\partial \ln|g|}{\partial t}$$

$$= \frac{f''}{2f} - \frac{(f')^2}{2f^2} - \frac{2}{r^2} - \left[\frac{f''}{2f} - \frac{(f')^2}{2f^2}\right] - \frac{1}{c^2}\left(f\dot{R}^2 + fR\ddot{R}\right) + \frac{(f')^2}{4f^2} + \frac{2}{r^2} + \frac{2f\dot{R}^2}{c^2} - \frac{f'}{4f}\left(\frac{f'}{f} + \frac{4}{r}\right) - \frac{3f\dot{R}^2}{c^2}$$

Thus, with $f'/f = 2krf$,

$$R_{rr} = -\frac{2f\dot{R}^2}{c^2} - \frac{fR\ddot{R}}{c^2} - 2kf = -\frac{8\pi G}{c^4}S_{rr} = -\frac{4\pi G}{c^4}g_{rr}(\rho c^2 - P) = -\frac{4\pi R^2 G f}{c^4}(\rho c^2 - P)$$

or

$$2\dot{R}^2 + R\ddot{R} + 2kc^2 = 4\pi G R^2(\rho - P/c^2) \tag{757}$$

Notice that r has disappeared, as it must. (Why must it?) Eliminating $\ddot{R}$ from (757) via equation (751) and simplifying the result leads to

$$\dot{R}^2 - \frac{8\pi G\rho R^2}{3} = -kc^2 \tag{758}$$

This is exactly the Newtonian equation (697) with the constant $2E$ identified with $-kc^2$. *But be careful.* The Newtonian version (697) was formulated with R dimensionless. If we use the standard FRW convention that k is either 1, 0, or -1, then remember that R has been rescaled to have dimensions of length in equation (758). To recover the Newtonian form is just a matter of replacing k by $1/a^2$, with dimensions of an inverse length squared. We then have $2E = -c^2/a^2$, with the same dimensions of $1/t^2$. Remember that we must allow $a^2 < 0$ in the case of negative curvature. The important point is that we now know that equation (697) is valid in full general relativity. The dynamical Newtonian energy constant is therefore to be identified with the geometrical curvature of the space.

There is one final item to consider in our analysis. We have not yet made explicit use of the equation for the conservation of energy-momentum, based on (211) and the Bianchi Identities:

$$T^{\mu\nu}_{\ \ ;\mu} = \frac{1}{\sqrt{|g|}}\frac{\partial(\sqrt{|g|}T^{\mu\nu})}{\partial x^\mu} + \Gamma^\nu_{\mu\lambda}T^{\mu\lambda} = 0. \tag{759}$$

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 (758): it was a lengthy derivation after all. The answer is an interesting YES and NO.

The $\nu = 0$ component of (759) reads

$$T^{\mu 0}_{;\mu} = \frac{1}{\sqrt{|g|}} \frac{\partial(\sqrt{|g|}T^{00})}{\partial x^0} + \Gamma^0_{\mu\lambda} T^{\mu\lambda} = 0. \tag{760}$$

Only affine connections of the form Γ^0_{ii} (spatial index i, no sum) are present, and with

$$\Gamma^0_{ii} = -\frac{1}{2g_{00}} \frac{\partial g_{ii}}{\partial x^0} = \frac{1}{2} \frac{\partial g_{ii}}{\partial x^0}, \quad T^{00} = \rho c^2, \quad T^{ii} = P g^{ii} = \frac{P}{g_{ii}}, \quad \text{(NO } i \text{ SUMS)}$$

equation (760) becomes:

$$\frac{1}{R^3} \frac{\partial(R^3 \rho c^2)}{\partial t} + 3P \frac{\dot{R}}{R} = 0 \to \dot{\rho} + 3\left(\rho + \frac{P}{c^2}\right) \frac{\dot{R}}{R} = 0. \tag{761}$$

This is simply the relativistic form of the first law of thermodynamics, $dE = -PdV$, which reverts to mass conservation in the nonrelativistic limit.

Exercise. Justify the last statement, and show that a pure vacuum energy density universe trivially satisfies (761).

If we now use (761) to substitute for P in (751), after a little rearranging we easily arrive at the result

$$\frac{d}{dt}\left(\dot{R}^2\right) = \frac{8\pi G}{3} \frac{d}{dt}\left(\rho R^2\right) \tag{762}$$

which in turn immediately integrates to (758)! This is surely a much more efficient route to (758), except...except that we cannot relate the integration constant that emerges from (762) to the spatial curvature constant k. True, we have a faster route to our final equation. But we only have dynamical content. Equation (758) is, after all, just a statement of energy conservation. Without explicitly considering the Ricci R_{rr} component, we lose the geometrical connection between Newtonian E (an untethered integration constant) and $-kc^2$ (a geometrical curvature constant). We therefore have *consistency* between the energy conservation and Ricci approaches, but not true *equivalence.* A subtle and interesting distinction.

10.4.3 The Ω_0 parameter

We begin with the dynamical cosmological equation in its "$2E$" form, with no loss of generality,

$$\dot{R}^2 - \frac{8\pi G \rho R^2}{3} = 2E, \tag{763}$$

where E is simply an energy integration constant. This may be determined, in terms of directly observable quantities, by using 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:

$$H_0^2 \left(1 - \frac{8\pi G \rho_0}{3H_0^2}\right) \equiv H_0^2(1 - \Omega_0) = 2E = -\frac{c^2}{a^2}, \tag{764}$$

where we have introduced the parameter Ω_0

$$\Omega_0 = \frac{8\pi G \rho_0}{3H_0^2}. \tag{765}$$

The Universe is therefore positively curved (closed) or negatively curved (open) according to whether the measured value of Ω_0 is larger or smaller than unity. Defining the critical mass density ρ_c by

$$\rho_c = \frac{3H_0^2}{8\pi G}, \tag{766}$$

and the critical *energy* density in the Universe is then $\rho_c c^2$. The parameter $\Omega(t)$ is defined just as Ω_0, with $\rho(t)$ and $H(t)$ replacing their current t_0 values.

A puzzle associated with classical cosmologies that were based on the assumption that the Universe is comprised of standard matter, was that the measured value of Ω_0 seemed to be about 0.05. This is disquieting because the two terms on the left side of equation (763) are then neither very close, nor vastly different from one another. Just different. But over the lifetime of the Universe, this interval of being just different is a mere fleeting instant of time. In the early Universe, the two terms were both large and remained very close in magnitude, like an Einstein-de Sitter model. Thereafter, the $\dot{R}$ term dominates in a continuously expanding Universe, with $\Omega(t) \ll 1$ over almost all time. So why are we privileged to live in a very special epoch in which $\Omega(t)$ just happens to be in transition from unity to miniscule?

It is now understood that the constituent components of our Universe involve the energy density of the vacuum. The currently best measured value of H_0 (in standard astronomical units) is 67.6 km s^{-1} Mpc^{-1},[21] or $H_0 \simeq 2.2\times10^{-18}$ s^{-1}. This number implies a critical density of 8.6×10^{-27} kg (about 5 hydrogen atoms) per cubic meter, far less than what is actually observed in standard baryonic matter[22]. 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 as earlier measurements suggested, 5% comes from ordinary baryonic matter in the form of gas and stars. We will denote the current vacuum contribution to Ω_0 as Ω_{V0} and the matter contribution as Ω_{M0}. In addition, there is a contribution from radiation, $\Omega_{\gamma 0}$, and while it is quite negligible now, in the early universe ($z \geq 5500$) it was completely dominant. The Universe went first through a radiation-dominated phase, followed by a matter-dominated phase, and then at about a redshift of 0.3, it started to switch to its current vacuum-dominated phase. This latter transition is still ongoing.

While an $\Omega = 1$ Universe remains an $\Omega = 1$ Universe throughout all time, the relative contribution from each of constituents changes, and the question of fine-tuning has not quite disappeared. We now just happen to be in the extended neighbourhood of an inflection point where Einstein-de Sitter behaviour is giving rise to exponential growth. It is not obvious what, if anything, to make of that.

10.5 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

[21] 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 pc $= 3.0856 \times 10^{13}$ km, 1 Mpc $= 3.0856 \times 10^{19}$ km. Stars are typically separated by 1 pc in galaxies, galaxies from one another by about 1Mpc.

[22] The absolute value of the *pressure* in the vacuum is some 5.4×10^{-10} Pa (for "Pascals", the SI pressure unit). This would be an ultra-vacuum pressure in a terrestrial lab, but is rather substantial by astronomical standards: in magnitude, it is roughly the ordinary 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 science fiction 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 ρ_{M0}. The dynamical equation of the Universe may then be written

$$\dot{R}^2 = 2E + \frac{8\pi G\rho R^2}{3} = H_0^2\left(1 - \Omega_{M0} + \frac{\Omega_{M0}}{R}\right) \tag{767}$$

or

$$\int_0^R \frac{dR'}{\sqrt{1 - \Omega_{M0} + (\Omega_{M0}/R')}} = \int_0^R \frac{R'^{1/2}dR'}{\sqrt{(1-\Omega_{M0})R' + \Omega_{M0}}} = H_0 t \tag{768}$$

The nature of the integral depends upon whether Ω_{M0} is less than, equal to, or greater than, 1. The case $\Omega_{M0} = 1$ is simple and leads immediately to

$$R = \left(\frac{3H_0 t}{2}\right)^{2/3} \tag{769}$$

our Einstein-de Sitter solution (706). According to (763) with $E = 0$, the average density is then given explicitly by

$$\rho = \frac{\dot{R}^2}{R^2}\frac{3}{8\pi G} = \frac{1}{6\pi G t^2} \tag{770}$$

Consider next the case $\Omega_{M0} > 1$, that of a Closed Universe. Then (768) may be written:

$$\int_0^R \frac{R'^{1/2}dR'}{\sqrt{1 - (1-\Omega_{M0}^{-1})R'}} = \Omega_{M0}^{1/2} H_0 t \tag{771}$$

Set

$$(1 - \Omega_{M0}^{-1})R' = \sin^2\theta' \tag{772}$$

Then (771) becomes

$$\frac{2}{(1-\Omega_{M0}^{-1})^{3/2}}\int_0^\theta \sin^2\theta'\, d\theta' = \frac{1}{(1-\Omega_{M0}^{-1})^{3/2}}\int_0^\theta (1 - \cos 2\theta')\, d\theta' = \Omega_{M0}^{1/2} H_0 t \tag{773}$$

or

$$\frac{2\theta - \sin 2\theta}{2(1-\Omega_{M0}^{-1})^{3/2}} = \Omega_{M0}^{1/2} H_0 t \tag{774}$$

With $\eta = 2\theta$, our final solution is in the form of a parameterisation:

$$R = \frac{1 - \cos\eta}{2(1-\Omega_{M0}^{-1})}, \quad H_0 t = \frac{\eta - \sin\eta}{2\sqrt{\Omega_{M0}(1-\Omega_{M0}^{-1})^3}} \tag{775}$$

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 [21].) Precisely analogous expressions emerge for the case $\Omega_{M0} < 1$, which are easily verified from (768):

$$R = \frac{\cosh\eta - 1}{2(\Omega_{M0}^{-1} - 1)}, \quad H_0 t = \frac{\sinh\eta - \eta}{2\sqrt{\Omega_{M0}(\Omega_{M0}^{-1} - 1)^3}} \tag{776}$$

This is the solution for negative curvature, an Open Universe.

Exercise. Show that for either equation (775) or (776),

$$\frac{d\eta}{dt} = \frac{c}{|a|R}$$

where $|a|$ denotes a modulus. We will use this result in the problems.

Exercise. Show that the equations analogous to (769), (775) and (776) for the FRW metric in the form (733) are respectively

$$R = \omega^{1/3} t^{2/3} \quad (\text{Einstein} - \text{deSitter})$$

$$R = \frac{\omega}{c^2}\left(\frac{1-\cos\eta}{2}\right), \quad ct = \frac{\omega}{c^2}\left(\frac{\eta - \sin\eta}{2}\right) \quad (\text{closed})$$

$$R = \frac{\omega}{c^2}\left(\frac{\cosh\eta - 1}{2}\right), \quad ct = \frac{\omega}{c^2}\left(\frac{\sinh\eta - \eta}{2}\right) \quad (\text{open})$$

where $\omega = 8\pi G\rho_{M0}R_0^3/3$. Show that $cdt = Rd\eta$ for the last two cases.

Exercise. Show that as $\eta \to 0$, both (775) and (776) reduce to $R = \Omega_0^{1/3}[3H_0 t/2]^{2/3}$ (the same of course as [769] for $\Omega_0 = 1$), and at late times (776) 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 Ω_0. Figure (10) shows this behaviour quite clearly.

What if $\Omega_0 = 0$? Show then that $R = H_0 t$ for all times. Wait. Can a universe *really* be expanding if there is nothing in it? Expanding with respect to exactly what, please? See Problem Set.

10.6 Our Universe

10.6.1 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 at the time this was a shocking result, for years there had been mounting evidence that something was seriously amiss with standard cosmological models. Without a process to increase the rate of the Universe's expansion, the large measured 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.

10.6.2 A Universe of ordinary matter and vacuum energy

It is possible to find a rather simple mathematical function for the scale factor $R(t)$ of our Universe. With $E = 0$ for a Euclidian space, (763) is

$$\dot{R} = \left(\frac{8\pi G\rho}{3}\right)^{1/2} R \tag{777}$$

The energy density ρc^2 is a combination of nonrelativistic matter ρ_M, which decays as $1/R^3$, and a vacuum energy density ρ_V, which remains constant. With ρ_{M0} the current value of ρ_M and $R_0 = 1$, we have $\rho_M = \rho_{M0}/R^3$ and therefore

$$\rho R^2 = \rho_V R^2 + \frac{\rho_{M0}}{R}. \tag{778}$$

Substituting (778) into (777) leads to

$$\frac{R^{1/2}\, dR}{\sqrt{\rho_{M0} + \rho_V R^3}} = \sqrt{\frac{8\pi G}{3}}\, dt. \tag{779}$$

Integration then yields

$$\int \frac{R^{1/2}\, dR}{\sqrt{\rho_{M0} + \rho_V R^3}} = \frac{2}{3}\int \frac{d(R)^{3/2}}{\sqrt{\rho_{M0} + \rho_V R^3}} = \frac{2}{3}\frac{1}{\sqrt{\rho_V}} \sinh^{-1}\left(\sqrt{\frac{\rho_V}{\rho_{M0}}} R^{3/2}\right) = \sqrt{\frac{8\pi G}{3}}\, t \tag{780}$$

Equation (780) then tells us:

$$R^{3/2} = \sqrt{\frac{\rho_{M0}}{\rho_V}} \sinh\left(\frac{3}{2}\sqrt{\frac{8\pi G\rho_V}{3}}\, t\right). \tag{781}$$

In terms of the Ω parameters, we have

$$\Omega_M = \frac{8\pi G\rho_{M0}}{3H_0^2}, \quad \Omega_V = \frac{8\pi G\rho_V}{3H_0^2}, \tag{782}$$

henceforth dropping the 0 subscript on the Ω's with the understanding that these parameters refer to current time. The dynamical equation of motion (777) at the present epoch tells us directly that

$$\Omega_M + \Omega_V = 1. \tag{783}$$

In terms of the observationally accessible quantity Ω_M, the scale factor$R$ becomes

$$R = \frac{1}{(\Omega_M^{-1} - 1)^{1/3}} \sinh^{2/3}\left(\frac{3}{2}\sqrt{1-\Omega_M}\; H_0 t\right), \tag{784}$$

or, with $\Omega_M = 0.31$,

$$R = 0.7659 \sinh^{2/3}(1.246 H_0 t). \tag{785}$$

This gives a current age of the Universe t_0 of

$$H_0 t_0 = \frac{2\sinh^{-1}(\sqrt{\Omega_M^{-1} - 1})}{3\sqrt{1-\Omega_M}} = \frac{\sinh^{-1}[(0.7659)^{-3/2}]}{1.246} = 0.9553, \tag{786}$$

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.)[23]

Exercise. Show that the redshift z is related to the time t since the big bang by

$$z = 1.306 \sinh^{-2/3}\left(\frac{1.19t}{t_0}\right) - 1,$$

with inverse

$$\frac{t}{t_0} = 0.84012 \sinh^{-1}\left(\frac{1.4919}{[1+z]^{3/2}}\right) \to \frac{1.2534}{(1+z)^{3/2}},$$

with the final limit corresponding to large z. If a civilisation develops 5 billion years after the big bang and we detect their signals(!), at what redshift would they be coming from?

Figure [22] is a comparison of a classical open model, with $\Omega_{M0} = 0.31$, with our vacuum-dominated Universe for the same value of Ω_{M0}. Notice how the extended transition from concave to convex temporal growth causes our Universe to grow nearly linearly with time over much of its history. The open model is some 12% younger than the now accepted vacuum-dominated model, and would not be old enough to accommodate the oldest known stars.

Figure [23] summarises the $\Omega_V - \Omega_M$ parameter space. Accelerating/decelerating and open/closed boundaries are also shown. The pioneering supernova data have now been superceded by statistical analyses of the cosmic microwave radiation background (discussed in more detail in later sections). These more recent results both confirm the original supernova date interpretation, and tightly constrain the allowed parameter range.

Figure [24] is a plot of equation (786) showing that age of vacuum/matter universes as a function of Ω_M. The smaller Ω_M is, the older the inferred age of the universe. As $\Omega_M \to 0$, $H_0 t_0$ approaches $-(\ln \Omega_M)/3$, a phenomenon that Sir Arthur Eddington, an early proponent of such models on philosopical grounds, dubbed the "logarithmic eternity."

[23] Possible notational confusion: $\sinh^{-1}$ always denotes the inverse hyperbolic sine function, whereas $\sinh^{\pm 2/3}$ denotes the hyperbolic sine function to the $+2/3$ or $-2/3$ power.

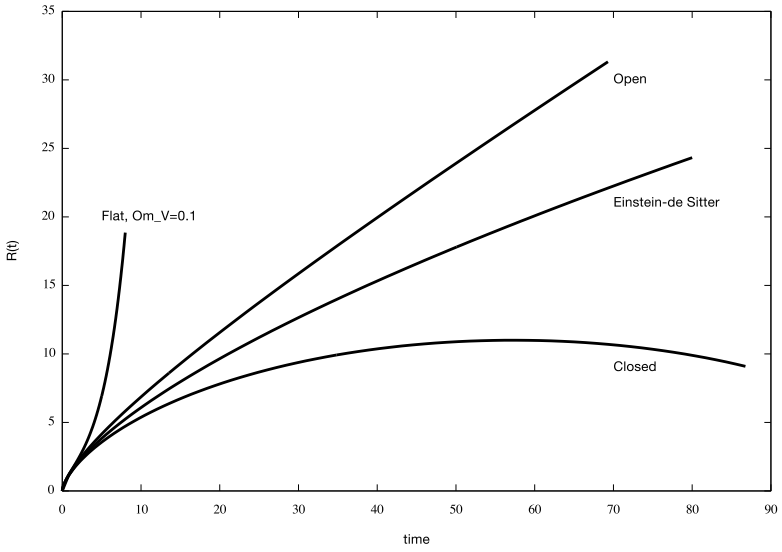


Figure 21: $R(t)$ versus time, units of $\Omega_M^{-1/2} H_0^{-1}$, for four model universes. 'Closed' is eq. (775); 'Einstein- de Sitter' eq. (769); 'Open' eq. (776). 'Closed' has $\Omega_M = 1.1$, for 'Einstein–de Sitter' $\Omega_M = 1$, 'Open' has $\Omega_M = 0.9$. The curve labelled 'Flat, Om_V= 0.1', eq. (784), 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.

Exercise. The $\sinh^{2/3}(x)$ function can be awkward to work with when inside an integral. There is, however, an accurate numerical approximation,

$$\sinh^{2/3}(x) \simeq \frac{x^{2/3}}{1 - x^2/9},$$

with a maximum error of about 2% over the range of x relevant for the age of the Universe, corresponding to $0 \leq x \leq 1.19$. (This maximum error occurs at $t = t_0$.) Using this approximation in our expression (785) for $R(t)$, determine the current size of the horizon in our Universe, and compare it with the E-dS value $2c/H_0$. (Answer: $2.49c/H_0$.) Explain qualitatively why the horizon is larger than the E-dS value.

For a yet more accurate approximation, replace $1 - x^2/9$ with $1 - x^2/9 + 4x^4/405$, which is essentially indistiguishable from $\sinh^{2/3}(x)$ over the parameter range relevant for the lifetime of the Universe. Make a plot a judge for yourself!

10.6.3 Radiation-dominated era

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 $\rho_\gamma c^2$. Under these conditions,

$$\rho_\gamma \propto 1/R^4.$$

This follows from the fact that the equilibrium energy density of radiation is proportional to T_γ^4, where T_γ is the radiation temperature. The temperature T_γ has the same R-scaling as a photon with energy $h\nu$: both the frequency ν as well as T_γ (which measures average energy per particle) decrease as $1/R$. In these early times, the dynamical equation of motion (763)

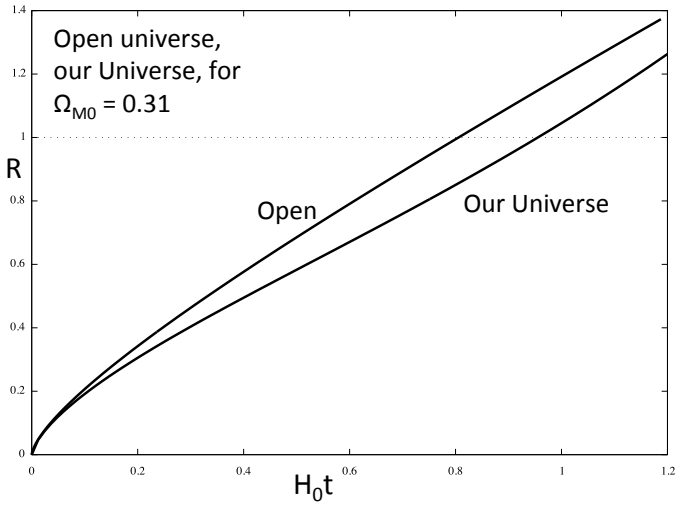


Figure 22: Comparison of $R(t)$ versus H_0t for our Universe eq. (784) and an open universe eq. (776), both with $\Omega_M = 0.31$. Crossing of the dotted line $R = 1$ (at the current age) occurs earlier for the positive energy open universe, which is about 12% younger than our Universe. This would be too young for the oldest stars.

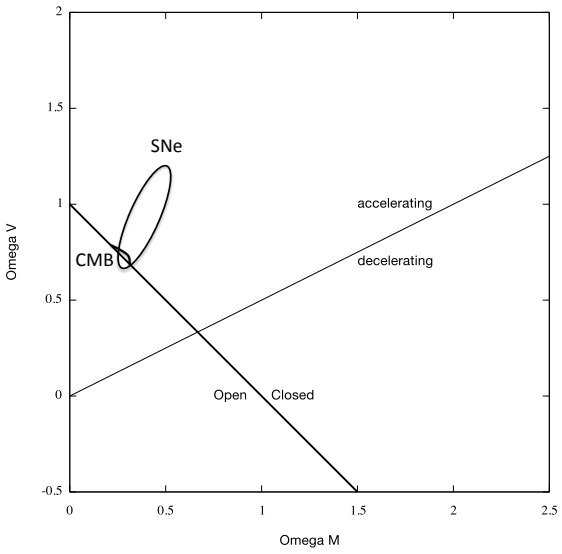


Figure 23: Parameter plane of Ω_V versus Ω_M assuming no radiation contribution. Regions of open/closed geometry and currently accelerating/decelerating dynamics are shown. Also shown are approximate zones of one standard deviation uncertainties for the distant supernova data (SNe) and—you have to squint—for 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.

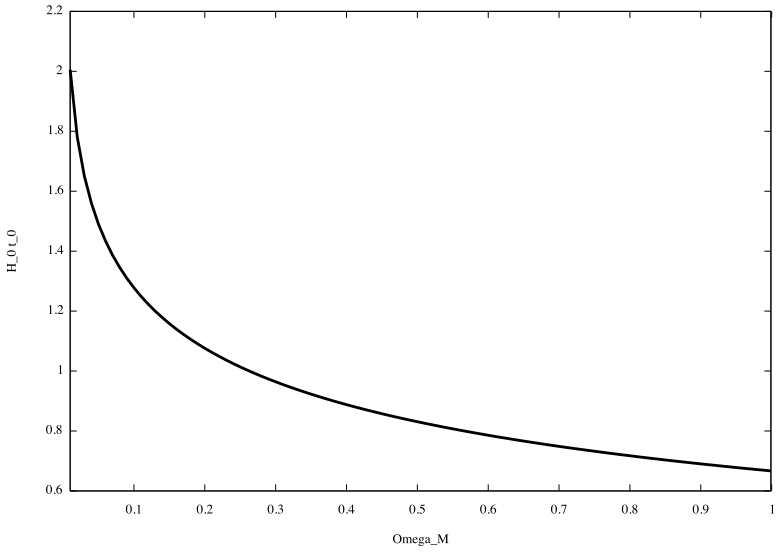


Figure 24: The current age of the Universe $H_0 t_0$ as a function of Ω_M for flat, matter plus vacuum energy models with negligible radiation, equation (786). As Ω_M approaches unity, the model recovers the Einstein–de Sitter value $H_0 t_0 = 2/3$; as $\Omega_M \to 0$, $H_0 t_0$ becomes proportional to $-\ln \Omega_M$ (Show!), and we recover the "logarithmic eternity" that was first highlighted by Sir Arthur Eddington.

is dominated by the two terms on the left side of the equation, both of which are very large (becoming infinite as $t \to 0$) compared with the constant on the right side. Then, using $\rho_\gamma R^4 =$ constant, the dynamical equation of motion at early times may then be written

$$R^2 \dot{R}^2 = \frac{8\pi G \rho_\gamma R^4}{3} = (\text{constant}). \tag{787}$$

This leads to a simple power law time dependence for the scale factor:

$$R(t) \propto t^{1/2} \quad (\text{radiation dominated universe}). \tag{788}$$

Knowing only that $R \sim t^{1/2}$, if we now return to the dynamical equation of motion, we may solve explicitly for $\rho_\gamma c^2$:

$$\rho_\gamma c^2 = \frac{3c^2}{8\pi G} \left(\frac{\dot{R}^2}{R^2} \right) = \frac{3c^2}{32\pi G t^2}. \tag{789}$$

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 4×10^{25} 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. We shall have more to say about this topic in §(10.8).

10.7 Observational foundations of cosmology

10.7.1 The first detection of cosmological redshifts

To some, the name of Vesto Melvin Slipher may conjure up images of a 1930's J. Edgar Hoover FBI G-man in a fedora who went after the bad guys. But Vesto was instead a mild-mannered, very 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 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 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 variables 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 true 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 astronomers is that they have a well-defined relationship between their oscillation time period and their absolute mean luminosity. So to determine an abosulte distance, first find a distant Cepheid variable. Measure its oscillation period. Determine thereby its true luminosity. Next, measure the star's flux, which is simply the energy per unit area per unit time crossing your detector. 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. This is all very simple, but for the kinds of distances we are discussing here one would need a superb quality, huge (100-inch) telescope. To such an instrument, only Hubble and a small handful of other astronomers had access[24].

[24]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.

10.7.2 The cosmic distance ladder

As observations improved through the 1930's, the linear relation between velocity and distance, which subsequently became known as the *Hubble Law,* $v = H_0 r$, continued to be more firmly established. There are two major difficulties 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 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 π 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}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.

10.7.3 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:

$$H(t) = \frac{\dot{R}}{R} = H_0 + (t - t_0)\left(\frac{\ddot{R}_0}{R_0} - H_0^2\right) + ... \tag{790}$$

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

$$H(t) = H_0 + H_0^2(t_0 - t)(1 + q_0) \tag{791}$$

where we have defined the *deceleration parameter* q_0 by

$$q_0 \equiv -\frac{\ddot{R}_0 R_0}{\dot{R}_0^2} \to -\frac{\ddot{R}_0}{H_0^2} \quad \text{for } R_0 = 1. \tag{792}$$

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 (791) becomes

$$H(z) = H_0[1 + z(1 + q_0)]. \tag{793}$$

For the classical matter-dominated cosmological models, q_0 is defined in such a way that it ought be a positive quantity. If $q_0 < 0$, the Universe would be accelerating, a seemingly outlandish possibility not taken seriously until after 1998. Consider the integral for the comoving coordinate of a photon, $r = \int cdt/R(t)$. With $R_0 = 1$ and $t - t_0 \equiv \delta t$, carry out the Taylor expansion in δt of

$$1 + z = \frac{1}{R(t)} = \frac{1}{R(t_0 + \delta t)} = 1 - \delta t \dot{R}_0 + (\delta t)^2 (\dot{R}_0^2 - \frac{\ddot{R}_0}{2}) + ... \tag{794}$$

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 (794), the next order refinement is

$$-\delta t = t_0 - t = \frac{1}{H_0}\left[z - \left(1 + \frac{q_0}{2}\right) z^2\right]. \tag{795}$$

Using our expansion for $1/R$ we find,

$$l(z) \equiv c \int_t^{t_0} \frac{dt'}{R} = c(t_0 - t) + \frac{cH_0}{2}(t - t_0)^2 + ... \tag{796}$$

and substituting with the help of (795) through order z^2:

$$l(z) = \frac{c}{H_0}\left[z - \frac{z^2}{2}(1 + q_0) + ...\right] \tag{797}$$

Equation (797) is general for any FRW model. q_0 embodies the leading order deviations from a simple Euclidian model in which $l(z) \propto z$.

10.7.4 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

$$\mathcal{F}(z) = \frac{LR^2(t)}{4\pi R_0^4 l^2(z)} = \frac{L}{4\pi(1+z)^2 R_0^2 l^2(z)} \tag{798}$$

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

$$d_L(z) = (1+z) R_0 l(z), \tag{799}$$

which reduces to the conventional Euclidian distance at small z.

Explanation: the two factors of $1 + z$ in the denominator of (798) 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:

$$\mathcal{F}(z) = \frac{LH_0^2}{16\pi c^2(1+z)(\sqrt{1+z}-1)^2} \qquad \text{(Einstein} - \text{de Sitter)} \tag{800}$$

The simple Euclidian value of $L/4\pi r^2$ is recovered at small z by recalling $cz = v = H_0 r$, whereas for large z, all 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 historical reasons. It is a logarithmic measure of the flux. Explicitly:

$$\mathcal{F} = \mathcal{F}_0 10^{-0.4m} \tag{801}$$

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 $F_0 = 2.52 \times 10^{-8}$ 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 (798) 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 Ω_V: we live in an accelerating Universe! (See figure [25].)

Exercise. Show that

$$\mathcal{F}(z) = \frac{LH_0^2}{4\pi c^2 z^2}\left[1 + (q_0 - 1)z + ...\right]$$

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.

With the determination of Ω_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 Ω_0 of the current mass density ρ_0 to the critical mass density $3H_0^2/8\pi G$.
- The value of the cosmological constant or equivalently, the ratio Ω_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 $\ddot{R}$ (751) and $\dot{R}^2$ (763). A quick summary:

- $H_0 \simeq 67.6$ km s^{-1} Mpc^{-1}.

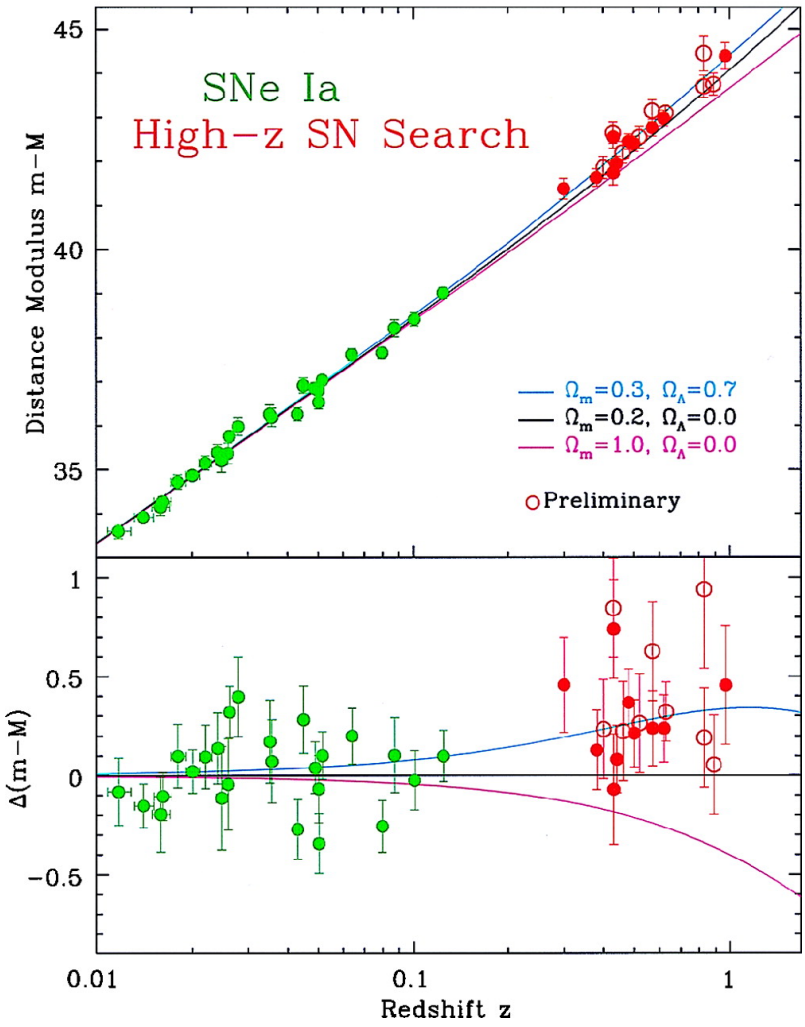


Figure 25: Evidence for an accelerating Universe from type Ia supernovae. The top figure shows a redshift-magnitude plot for three different FRW models, $\Omega_{M0} = 1$, $\Omega_V = 0$, an Einstein-de Sitter model, in magenta; $\Omega_{M0} = 0.2$, $\Omega_V = 0$, an open model, in black; and $\Omega_{M0} = 0.3$, $\Omega_V = 0.7$ 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.

- $t_0 \simeq 13.8$ 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 Supernova 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 [23] speaks for itself.

10.8 The Cosmic Microwave Background Radiation (CMB)

10.8.1 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 $1/R^3$. The Universe must, at some point in the past, have become radiation dominated. This, we shall see, occurred at a redshift $z \sim 1000$.

You should now be thinking of the radiation as actual 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:

$$n(\nu, T)\, d\nu = \frac{(8\pi\nu^2/c^3)d\nu}{\exp(h\nu/kT_\gamma) - 1} \tag{802}$$

If radiation is in thermal equilibrium with matter, there is no choice of how many photons there are at a given frequency. Mother Nature determines that precisely. 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 total number density is not given *exactly* by (802), the photons are simply not in thermal equilbrium with their surroundings, and their numbers will change as this thermal equilibrium is inevitably approached. So does the unavoidable decline in photon number density with ogoing, universal expansion mean that we lose a thermal equilibrium distribution for the radiation?

At about the time that the radiation energy density was becoming subdominant relative to matter, the photons also lost good coupling with the matter. Thermal equilibrium could then no longer be maintained via photon-electron collisions or induced exchanges with atoms. Imagine a situation in which the photons "move along" with the expansion of the Universe, but are otherwise unaffected by interactions. 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, the number density and effective temperature of the photons together 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 concurrently evolving and ever-cooling photon temperature.

To see this, start with the spectrum (802) at the time t of last scattering, by which we mean the last moment thermal equilibrium was maintained by collisions. We follow a group of photons with frequency ν and small 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 ν' 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 determine 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 ν everywhere in the Planck formula by $R'\nu'/R$. Carrying this through,

$$n(\nu', t')d\nu' = \frac{R^3}{R'^3} n(\nu, t)d\nu = \frac{R^3}{R'^3} \frac{(8\pi/c^3)(R'\nu'/R)^2 d(\nu' R'/R)}{\exp(hR'\nu'/RkT_\gamma) - 1} = \frac{8\pi\nu'^2\, d\nu'/c^3}{\exp(h\nu'/kT'_\gamma) - 1} \qquad (803)$$

where $T'_\gamma = T_\gamma R/R'$, a much *cooler* temperature than T_γ. The remarkable point is that the distribution (803) is *still* a blackbody at time t', but with a new temperature T'_γ 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 precise blackbody. Adding to the confusion, there were puzzling observations, largely based on strikingly non-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 whence 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, no rocket exhaust of course, and it was established beyond any question that the background radiation was almost a perfect blackbody at $T_\gamma = 2.725K$.

Figure [26] summarises the data. The error bars are shown at 400 times their actual value, just to be visible! (The earlier rocket observations had by this time been 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 km s^{-1} relative to the CMB, a surprisingly large value. This translates to a measured radiation temperature that is about 2×10^{-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 tiny seeds, collapse, and then 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. Compress the matter and you will also compress the radiation, and thereby 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 indeed found the temperature fluctuations (figure [27]). But when COBE (and later satellites in yet greater detail) did find them, they were an order of magnitude smaller than expected. So how does structure form in the Universe? The consensus answer is that there is more 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 only a gravitational imprint on the microwave background. Dark Matter feels the force of gravity, and so far as is known, nothing else. It had been postulated long before COBE as something that was needed to hold galaxies together. (Were visible light truly representative of the bulk of the matter, internal stellar motions in galaxies would be far too large to be keep the stars gravitationally bound.) COBE offered yet 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.

10.8.2 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 the 1940s. 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[25].

Gamow drew attention to the fact that helium, with about 25% of the mass of the observable Universe, is neither an overwhelming nor a tiny constituent. Gamow was convinced,

[25]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.

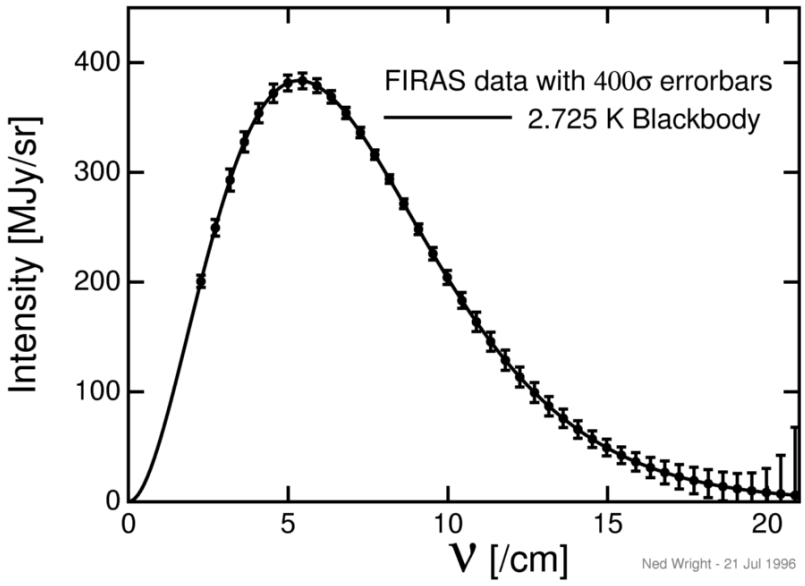


Figure 26: COBE satellite data showing a perfect fit to a blackbody at 2.725 K. To be visible, the error bars are shown at 400 times their actual value! When shown at an American Astronomical Society Meeting, this plot triggered a spontaneous standing ovation. Units of ν are cm^{-1}, i.e., the wavelength in cm is the reciprocal of the number on the x axis.

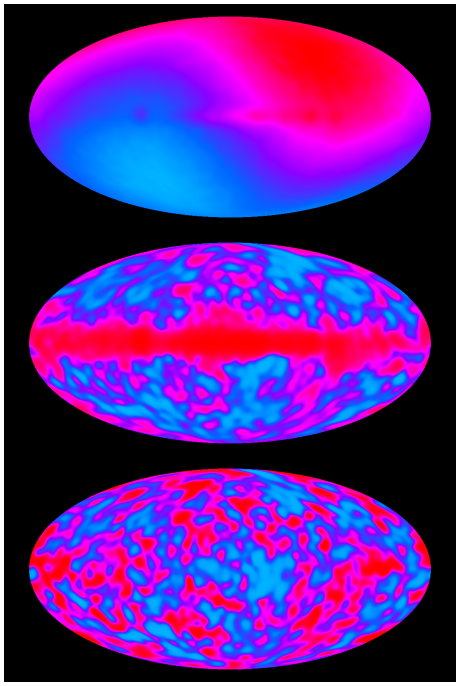


Figure 27: COBE temperature fluctuations. *Top:* Temperature dipole variation dominated by motion of the galaxy. *Middle:* Residual with kinematic dipole subtracted off. *Bottom:* Further residual with galactic foreground subtracted off. These represent primordial fluctuations in the CMB.

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 bit turns out to be false. Heavier elements are made over the course of stellar evolution, and also possibly in neutron star mergers as well.) 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 evolves 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 simple:

$$n + p \rightarrow D + \gamma$$

where γ denotes 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 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 σ (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). Roughly speaking, at the time of nucleosynthesis t, about one reaction per proton should have occurred. This is inferred from the order unity mass fraction of helium that we now observe. The heart of the Gamow argument is thus:

$$nv\sigma t \sim 1, \tag{804}$$

where t is the age of the Universe at the time of nucleosynthesis. How simple!

The product σv is just a laboratory number nearly independent of v (because the cross section σ depends inversely on v), and, as Gamow knew from nuclear physics experiments, is equal to about 4.6×10^{-26} m^3 s^{-1}. As for the time t, we are interested in the epoch when the temperature $T_\gamma \sim 10^9$K. Using equation (789) for the density of relativistic particles, and following Gamow by assuming (not quite correctly, but let it go) that this was all radiation with energy density $a_\gamma T_\gamma^4$ (a_γ is the standard blackbody radiation constant), the time of interest t is

$$t = \frac{c}{T_\gamma^2}\left(\frac{3}{32\pi G a}\right)^{1/2} \tag{805}$$

which amounts to 230 s for $T_\gamma = 10^9$K. This gives n very close to 10^{23} 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. (Recall that 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_γ should be near $10^9/10^8 = 10$ K! The millimeter wave 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). Gamow could have, but did not, pursue this result aggressively, and the prediction was gradually forgotten. Thus, the CMB might have been detected in the late 1940s, as part of the great postwar flowering of physics. But it wasn't.

The argument would be put slightly differently now, but in essence it is correct, a brilliant piece of intuitive reasoning.

Exercise. Repeat the Gamow argument with modern cosmological numbers. Keep $T_\gamma = 10^9$K and the σv 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 $(7/8)a_\gamma T_\gamma^4$ 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_γ 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 very 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 its view 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 instrument builders who designed and built the contemporary Dicke radiometer. The 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 own detector. Their device was designed to receive long range 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 they 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. In 2019, however, Jim Peebles received his own Nobel for his cumulative contributions to modern comology.

10.8.3 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, as we have noted. Recall that in 1992, matters were finally lain 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 within galaxies and large galactic velocities within clusters of galaxies. These velocities are unsustainable unless most of the gravitating mass of the galaxy or cluster is in the form of matter which does not emit light: dark matter. The COBE results are yet further evidence for the presence of dark matter in the Universe.

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, Ω_{M0}, Ω_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 there were no qualitatively new physical findings. The current story remains essentially the one revealed to us by WMAP (figure [28]), now on a yet stronger footing.

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 full meaning and implications will take many years to elucidate.

10.9 Thermal history of the Universe

10.9.1 Prologue

The recorded history of civilisation began when Sumerian merchants inscribed their grain inventories on clay tablets. (By the way, 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. 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 these inscriptions as temperature fluctuations, and the clay tablet is the cosmic microwave background.

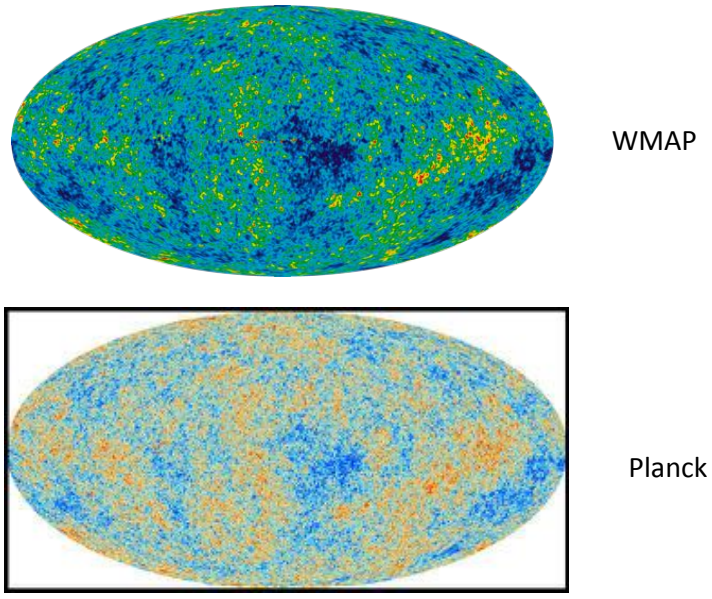


Figure 28: WMAP (top) and Planck (bottom) temperature fluctuations, with different colour scales. Think of these as the Universe's baby pictures. Note the much higher angular resolution compared with COBE; Planck is another factor of 3 better than WMAP. The angular information encoded on these small scales is valuable as a means of tightly constraining cosmological parameters.

The temperature fluctuations trace the initial density fluctuations, which as we have seen, are coupled to the radiation by electron-photon scattering, also known as *Thompson scattering.* While this scattering remained vigorous, the CMB dutifully recorded and re-recorded the ever evolving changes in the density. Then, matter and radiation suddenly became decoupled. 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 σ_T be the Thompson cross section for a photon to collide with an electron, a constant number equal to 6.7×10^{-29} 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_γ 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 stages. 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} K through 3000K, the time when hydrogen ions recombine with electrons to form neutral H atoms. In the second stage, in Chapter 11, 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, one of the best arguments 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 quantities 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:

$$m_P = \left(\frac{hc}{G}\right)^{1/2} = 5.456 \times 10^{-8} \text{ kg} \tag{806}$$

$$l_P = \left(\frac{hG}{c^3}\right)^{1/2} = 4.051 \times 10^{-35} \text{ m} \tag{807}$$

$$t_P = l_P/c = \left(\frac{hG}{c^5}\right)^{1/2} = 1.351 \times 10^{-43} \text{ s} \tag{808}$$

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 5×10^{-4} GeV in rest energy; the Planck mass energy is $m_Pc^2 = 3.06\times10^{19}$ GeV. By these Planckian standards, we will always be working with very low mass particles, very large distances, and very long times.

10.9.2 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.[26] There is also a population of protons and neutrons, which are energetically unimportant at these temperatures. But keep track of them in the inventory! 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 out 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,

$$\Delta mc^2 \equiv (m_n - m_p)c^2 = 1.293 \text{ MeV} = 2.072 \times 10^{-13} \text{ J}. \tag{809}$$

This is very small compared with either m_nc^2 or m_pc^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:*

$$\frac{P_i}{P_j} = \frac{g_i}{g_j} \exp\left[-\left(\frac{E_i - E_j}{kT}\right)\right] \tag{810}$$

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, 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 1.5×10^{10}K. (By way of comparison, for electrons and positrons, $m_ec^2/k = 5.9 \times 10^9$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

[26] By using the word "freely," we ignore the particle rest masses.

$f(E_i)/f(E_j)$. 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 δ, that

$$f(E) \propto \exp(-\beta E)$$

where β is an as yet undetermined constant. We can determine β 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 simply 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, 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 10^9K, the remaining neutrons are salvaged by being safely packed into stable ^{4}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

$$\frac{4m_p \times (n/2)}{m_p(n+p)} = \frac{2}{1+p/n} \simeq 0.33. \tag{811}$$

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.

10.9.3 Neutrino and photon temperatures

This subsection is non-examinable.

As the temperature slides from 5×10^9 K to below 10^9, the electron positron pairs annihilate into photons,

$$e^+e^- \to 2\gamma$$

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,

$$E + PV - TS + \mu\cancel{N} = 0,$$

for a gas with zero chemical potential μ. The latter is true of a relativistic gas that freely creates and annihilates its own particles. So for the entropy per unit volume s,

$$s \equiv \frac{S}{V} = \frac{E}{VT} + \frac{P}{T} \tag{812}$$

With ρc^2 denoting the energy per unit volume and $P = \rho c^2/3$ for a relativistic gas,

$$s = \frac{4\rho c^2}{3T} \tag{813}$$

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

$$sR^3 = \frac{4a_\gamma (TR)^3}{3}\left(1 + 2 \times \frac{7}{8}\right) = \frac{11a_\gamma (TR)^3_{before}}{3} \tag{814}$$

Afterwards, and after the reestablishment of thermal equilibrium, it is all photons:

$$sR^3 = \frac{4a_\gamma (TR)^3_{after}}{3} \tag{815}$$

Since the number of photons produced is equal to the number of relativistic electrons and positrons that annihilated, this process conserves entropy, sR^3. It is reversible: one could recompress the adiabatic expansion backwards to produce anew the pairs. In other words,

$$\frac{4(TR)^3_{after}}{3} = \frac{11(TR)^3_{before}}{3} \tag{816}$$

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^- annihilation we are now considering ignores the neutrinos. This is because 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_ν to photon temperature T_γ now follows from:

$$T_{\nu\ (after)} = T_{\nu\ (before)} = T_{\gamma\ (before)} = \left(\frac{4}{11}\right)^{1/3} T_{\gamma\ (after)} = 0.7138\ T_{\gamma\ (after)}. \tag{817}$$

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 1.7×10^{-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.

10.9.4 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

$$\frac{P_i}{P_j} = \frac{g_i}{g_j} \exp\left(-\frac{E_i - E_j}{kT}\right) \tag{818}$$

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, (818) becomes, upon adding up all possible free electron states,

$$\frac{n_e n_p}{n_0} = \exp\left(-\frac{\Phi}{kT}\right) \frac{4\pi}{h^3} \int_0^\infty p^2\, e^{-p^2/2m_e kT}\, dp \tag{819}$$

where Φ is the ionisation potential of hydrogen, m_e the electron mass, 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

$$\int_0^\infty p^2\, e^{-p^2/2m_e kT}\, dp = (2m_e kT)^{3/2} \int_0^\infty x^2 e^{-x^2}\, dx = (2m_e kT)^{3/2} \times \frac{\pi^{1/2}}{4} \tag{820}$$

The Saha equation becomes

$$\frac{n_e n_p}{n_0} = \left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} \exp\left(-\frac{\Phi}{kT}\right) \tag{821}$$

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

$$\frac{x^2}{1-x} = \frac{1}{n_H}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} \exp\left(-\frac{\Phi}{kT}\right) = \left(\frac{2.415 \times 10^{21} T^{3/2}}{n_H}\right) \exp\left(-\frac{1.578 \times 10^5}{T}\right) \tag{822}$$

in SI units. The value $x = 0.5$ is attained at a redshift of about 1400, a temperature of 3800 K. (This is naturally close to the formal decoupling redshift, $z \simeq 1100$.) This is remarkable, because it is much less than the formal Boltzmann ionisation temperature of 1.58×10^5K. At redshifts much less than 1400, the Universe becomes transparent to photons, and 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 radiation produced by the accretion process that gave rise to these galaxies. Alas, to pursue this active area of current astrophysical research would take us too far afield at this point.

"Curiouser and curiouser!" cried Alice ..."now I'm opening out like the largest telescope that ever was!"

— Alice's Adventures in Wonderland, by L. Carroll

11 Supplemental Topics

This chapter contains additional material on several advanced topics in general relativity. It is off-syllabus and non-examinable. I hope you will pick and choose material that is of interest you. This is work in preparation, and your critical comments are welcome.

11.1 The Oppenheimer-Snyder problem

In section (6.9), we presented a calculation of the redshift coming from the surface of a collapsing spherical dust cloud, en route to forming a black hole, as measured by a distant observer. The collapse problem is a more comprehensive investigation carried out by Oppenheimer and Snyder in 1939, which included an explicit solution for the global form of the metric, both interior and exterior to the dust cloud. This calculation relies on concepts developed for cosmology, as well as for Schwarzschild black holes.

Consider a self-gravitating sphere whose surface is at radial coordinate a at time $t = 0$. The sphere encloses a uniform energy density $\rho(t)c^2$, and is subject only to its own internal gravitational forces. In particular, there is no pressure: this is a cloud of noninteracting dust particles. The metric describing the interior of such a dust sphere is easy to write down, by application of Birkhoff's theorem. The trick is to imagine adding an infinite, spherically symmetric dust envelope exterior to our collapsing dust sphere. This envelope contracts uniformly and continuously with the interior sphere, and does not affect the sphere's internal dynamics. There is, in effect, no gravitational field inside the interior sphere due to the exterior envelope. But the metric for this combined extended mass distribution is clearly just a standard FRW cosmology, which we write in the following form:

$$-c^2 d\tau^2 = -c^2 dt^2 + R^2(t)\left(\frac{dr^2}{1 - Kr^2/c^2} + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2\right) \tag{823}$$

where $K > 0$ is a constant, with units of $(\text{time})^{-2}$. This metric must be the *interior* solution for the dust sphere in isolation as well. Recall that r and t are comoving coordinates in (823), and R is a dimensionless scale factor.

The equation for the rate of change of the scale factor $\dot{R}$ is the usual dynamical equation of cosmology:

$$\dot{R}^2 - \frac{8\pi G\rho R^2}{3} = -K \tag{824}$$

We choose the origin of time $t = 0$ to coincide with the start of the collapse, when the sphere's surface has radial coordinate a. (Note: this is *not* the same as the proper radius.

Why not?) With $R(0) = 1$, $\dot{R}(0) = 0$, this fixes the constant K:

$$K = \frac{8\pi G\rho(0)}{3} = \frac{2GM}{a^3} = \frac{c^2 R_S}{a^3} \equiv \epsilon \frac{c^2}{a^2}, \tag{825}$$

where we have defined $\epsilon = R_S/a$ and

$$Mc^2 \equiv \frac{4\pi\rho(0)c^2a^3}{3} \tag{826}$$

is the (conserved) total energy of the sphere. The Schwarzschild radius R_S is defined using this value of M. ϵ is a small quantity for an initial Newtonian-scale cloud radius a large compared to R_S, but in what follows we shall not assume this, except when explicitly noted. Using $\rho(t)R^3 = \rho(0)$, equation (824) becomes

$$\dot{R}^2 = K(R^{-1} - 1) = \epsilon(c^2/a^2)(R^{-1} - 1) \tag{827}$$

with $\dot{R} < 0$.

Exterior to the dust sphere, Birkhoff's theorem once again comes into play. Since the total mass of this spherical distribution is not changing with time, the only possible radial metric is the static Schwarzschild solution:

$$-c^2 d\tau^2 = -c^2 d\bar{t}^2 \left(1 - 2GM/\bar{r}c^2\right) + \left(1 - 2GM/\bar{r}c^2\right)^{-1} d\bar{r}^2 + \bar{r}^2 d\bar{\theta}^2 + \bar{r}^2 \sin^2\bar{\theta} d\bar{\phi}^2 \tag{828}$$

We use the notation $\bar{t}, \bar{r}$, etc. to distinguish these fixed coordinates, which go over to standard spherical Minkowski coordinates at large distances from the sphere, from the comoving t, r coordinates of (823). The problem now is to find a *single* set of coordinates for both the interior and exterior of the sphere, valid at all times. We shall do this by expressing the interior sphere metric (823) in terms of the exterior coordinates of (828).

The metric at the surface of the sphere follows from continuity arguments: it must have the form of a standard Schwarzschild metric at the time the instantaneous radius of the sphere, $aR(t)$ passes through a given fixed radial coordinate $\bar{r}$. Writing the metric in terms of ϵ, we have:

$$-c^2 d\tau^2 = -c^2 d\bar{t}^2 \left(1 - \epsilon/R\right) + \frac{d\bar{r}^2}{1 - \epsilon/R} + \bar{r}^2 d\bar{\theta}^2 + \bar{r}^2 \sin^2\bar{\theta} d\bar{\phi}^2 \quad \text{(surface metric)} \tag{829}$$

At the surface of the sphere, the relationship between $\bar{t}$ and t is exactly that given by equation (395) from §6.9, which relates the time interval $d\bar{t}$ of an observer at infinity to the time interval dt of an observer at the surface. Here, dt here replaces $d\tau$, and aR replaces r:

$$\frac{d\bar{t}}{dt} = \frac{\sqrt{1 - \epsilon}}{1 - \epsilon/R} \tag{830}$$

It is useful to incorporate the time implicitly via the function $R(t)$:

$$\frac{d\bar{t}}{dR} = \frac{1}{\dot{R}} \frac{d\bar{t}}{dt} = -\frac{a/c}{1 - \epsilon/R} \sqrt{\frac{1 - \epsilon}{\epsilon(R^{-1} - 1)}}. \tag{831}$$

In integral form this is:

$$\bar{t} = \frac{a}{c} \sqrt{\frac{1 - \epsilon}{\epsilon}} \int_R^1 \frac{x}{x - \epsilon} \frac{dx}{\sqrt{x^{-1} - 1}} \tag{832}$$

where the integration constant has been determined by demanding that $\bar{t} = 0$ at the start of the collapse, when $R = 1$. The goal now is to generalise this result so that it is a functional relationship valid not only between $\bar{t}$ and the surface radius $r = a$, but for any interior radius r as well. Finding the interior metric in $\bar{t}$, $\bar{r}$ coordinates reduces to implementing the coordinate transformation.

Formally, we require the functions f and g, where

$$\bar{r} = f(r,t), \quad \bar{t} = g(r,t). \tag{833}$$

The spherical angles $\theta = \bar{\theta}$ and $\phi = \bar{\phi}$ are unaffected. It is then obvious that

$$\bar{r} = f(r,t) = rR(t) \tag{834}$$

so that the coefficient of solid angle terms matches. We therefore require only the function $g(r,t)$, knowing its value at $\bar{r} = rR$ via equation (832).

Next, we use the relations (833) in the metric (823) and enforce the symmetry-maintaining constraint that there be no term proportional to $d\bar{r}\, d\bar{t}$ in the interior metric. We have

$$d\bar{r} = Rdr + r\dot{R}dt, \quad d\bar{t} = \partial_r g\, dr + \partial_t g\, dt \tag{835}$$

Solving for dr, dt:

$$dr = (\partial_t g\, d\bar{r} - r\dot{R}\, d\bar{t})/D, \quad dt = (Rd\bar{t} - \partial_r g\, d\bar{r})/D, \quad D = (R\partial_t g - r\dot{R}\partial_r g) \tag{836}$$

Substituting these relations into (823), and demanding that the coefficient of $d\bar{r}\, d\bar{t}$ vanish, leads to the equation

$$\frac{\partial g}{\partial t} - \frac{c^2 - Kr^2}{rR\dot{R}}\frac{\partial g}{\partial r} = 0, \tag{837}$$

a linear partial differential equation for g. The solution to this class of equation is that the function $g(r,t)$ is constant along $r = r(t)$ paths in the rt plane defined by

$$\frac{dr}{dt} = -\frac{c^2 - Kr^2}{rR\dot{R}}, \quad g = g(r(t), t). \tag{838}$$

These paths are known as *characteristics.* Note that g need not be the same constant from one path to another, but g must remain constant along each individual characteristic path. To prove this claim, note that

$$dg \equiv dt\left[\frac{\partial g}{\partial t} + \frac{dr}{dt}\frac{\partial g}{\partial r}\right], \tag{839}$$

so that if equation (837) is enforced along the $r(t)$ paths defined by (838), then g is indeed constant ($dg = 0$) along a characteristic. Another way to state the same thing is that $g(r,t)$ depends only upon the combination of r and t that constitutes the integration constant of the characteristic path of the ordinary differential equation (838).

The characteristic ordinary differential equation may be written in terms of $R = R(t)$:

$$\frac{dr}{dR} = \frac{1}{\dot{R}}\frac{dr}{dt} = -\frac{c^2 - Kr^2}{rR\dot{R}^2} = -\frac{c^2 - Kr^2}{rK(1-R)} = -\frac{(a^2 - \epsilon r^2)}{\epsilon r(1-R)}, \tag{840}$$

whence it is readily solved:

$$(1-\epsilon r^2/a^2)^{1/2}(1-R) = \text{constant}. \tag{841}$$

We next implicitly define the function $F(r,t)$ by

$$(1-\epsilon r^2/a^2)^{1/2}(1-R) = (1-\epsilon)^{1/2}(1-F) \tag{842}$$

Then $F = R(t)$ at the surface $r = a$ for all time. Moreover, off this surface, F will have by explicit construction the correct (r,t) dependence to be a constant of the motion along each characteristic. Therefore, we may choose the time transformation to depend solely upon F: $\bar{t}(r,t) = g(F)$. In particular, if we simply replace R by F in equation (832):

$$\bar{t} = g(r,t) = g(F) = \frac{a}{c}\sqrt{\frac{1-\epsilon}{\epsilon}}\int_{F(r,t)}^{1}\frac{x}{x-\epsilon}\frac{dx}{\sqrt{x^{-1}-1}}, \tag{843}$$

we arrive at precisely the function we are seeking. It agrees for all t with $\bar{t}$ at the surface of the sphere, and has the proper (r,t) dependence in the interior. The integral may in fact be explictly evaluated:

$$\frac{c\bar{t}}{a} = \sqrt{\frac{1-\epsilon}{\epsilon}}\left[(1+2\epsilon)\cos^{-1}\sqrt{F} + \sqrt{\frac{\epsilon^3}{1-\epsilon}}\ln\left|\frac{1+Y}{1-Y}\right| + \sqrt{F(1-F)}\right] \tag{844}$$

where

$$Y^2 = \frac{\epsilon}{F}\frac{1-F}{1-\epsilon} \tag{845}$$

When $\epsilon \ll 1$ with no further approximations, then to leading order in ϵ we have the simple expressions,

$$F - R \simeq -\frac{\epsilon}{2}(1-R)(1-r^2/a^2), \quad F(1-F) - R(1-R) \simeq \frac{\epsilon}{2}(1-r^2/a^2). \tag{846}$$

Using equations (834) and (843) for $\bar{r}$ and $\bar{t}$, it is now a straightforward, if somewhat cumbersome, procedure to transform the interior metric (823 into $\bar{t}, \bar{r}, \bar{\theta}, \bar{\phi}$ coordinates. The interior metric then takes the form

$$-c^2 d\tau^2 = -Bc^2 d\bar{t}^2 + Ad\bar{r}^2 + \bar{r}^2(d\bar{\theta}^2 + d\bar{\phi}^2) \tag{847}$$

where

$$B = \frac{R}{F}\sqrt{\frac{1-\epsilon r^2/a^2}{1-\epsilon}}\frac{(1-\epsilon/F)^2}{1-\epsilon r^2/a^2 R} \tag{848}$$

and

$$A = (1-\epsilon r^2/a^2 R)^{-1} \tag{849}$$

The derivation of A and B is left as a problem for the reader.

Formally, this solves our problem. The $\bar{t}$ coordinate does not, however, cover all of spacetime! First, there is a logarithmic singularity at $F = \epsilon$. If we treat this by taking a principal value for the divergent integral, then two different values of t are mapped to the same $\bar{t}$. Yet worse, once $F < 0$ (e.g., when $R = \epsilon^2$ and $\epsilon \ll 1$), $\bar{t}$ is not defined at all. These regions of spacetime are of course perfectly viable, and well-decribed in interior (r,t) coordinates. A complete description of space time may be constructed by using $(\bar{r},\bar{t})$ coordinates exterior to the cloud and continuously through a small neighbourhood near the surface. The coodinate transformation (843) then patches smoothly to (r,t) coordinates, which are valid until the surface reaches the $r = 0$ singularity.

11.2 The energy and angular momentum of black hole spacetimes

In section (5.8) we showed that the 4-momentum P^ν and angular momentum components J^{jk} for a given spacetime geometry take the forms:

$$P^\nu = \int_V \tau^{0\nu} d^3x \quad \text{and} \quad J^{ij} = \frac{1}{c}\int_V (x^i\tau^{0j} - x^j\tau^{0i})\, d^3x, \tag{850}$$

where $\tau^{\mu\nu}$ is the energy-momentum pseudo tensor of the combined matter, radiation and gravitational fields. In this section, we shall evaluate these expressions explicitly for the Kerr black hole metric (420), showing that the total energy is Mc^2, and the total angular momentum is Mac.

Combining equations (297), (563) and (578), the field equations may be written:

$$\begin{aligned}
-\frac{16\pi G}{c^4}\tau^{\mu\nu} &= \partial_\lambda\partial^\lambda\bar{h}^{\mu\nu} - \partial_\lambda\partial^\nu\bar{h}^{\lambda\mu} - \partial_\lambda\partial^\mu\bar{h}^{\lambda\nu} + \eta^{\mu\nu}\partial_\lambda\partial_\rho\bar{h}^{\lambda\rho} \\
&= \partial_\lambda\partial^\lambda\bar{h}^{\mu\nu} - \eta^{\nu\lambda}\partial_\rho\partial_\lambda\bar{h}^{\rho\mu} - \partial_\lambda\partial^\mu\bar{h}^{\lambda\nu} + \eta^{\mu\nu}\partial_\lambda\partial_\rho\bar{h}^{\lambda\rho} \\
&= \partial_\lambda\left(\partial^\lambda\bar{h}^{\mu\nu} - \eta^{\nu\lambda}\partial_\rho\bar{h}^{\rho\mu} - \partial^\mu\bar{h}^{\lambda\nu} + \eta^{\mu\nu}\partial_\rho\bar{h}^{\lambda\rho}\right) \\
&\equiv \partial_\lambda\Theta^{\lambda\mu\nu}.
\end{aligned} \tag{851}$$

At the expense of obscuring the $\mu\nu$ symmetry, the index manipulation in the second term on the second line allows for a flux pseudo tensor $\Theta^{\lambda\mu\nu}$ that is manifestly antisymmetric in λ and μ. This means that in the integrals of equation (850), we may substitute

$$\tau^{0\nu} = -\frac{c^4}{16\pi G}\partial_\lambda\Theta^{\lambda 0\nu} = -\frac{c^4}{16\pi G}\partial_i\Theta^{i0\nu}, \tag{852}$$

i.e., only the spatial contributions of the $\Theta^{\lambda 0\nu}$ divergence for $\tau^{\mu\nu}$ enter into the calculation. This leads immediately, via Gauss's theorem, to a surface integral at infinity, which requires only the leading asymptotic order of $\bar{h}^{\mu\nu}$ for its evaluation, a considerable savings of effort.

To express the integrand of J^{ij} in the form of an exact divergence, we perform the following manipulations:

$$\begin{aligned}
-\frac{16\pi G}{c^4}(x^i\tau^{0j} - x^j\tau^{0i}) &= \partial_i(x^j\Theta^{i0k} - x^k\Theta^{i0j}) + \Theta^{k0j} - \Theta^{j0k} \\
&= \partial_i(x^j\Theta^{i0k} - x^k\Theta^{i0j}) + \partial^k\bar{h}^{0j} - \partial^j\bar{h}^{0k} \\
&= \partial_i(x^j\Theta^{i0k} - x^k\Theta^{i0j} + \delta^{ik}\bar{h}^{0j} - \delta^{ij}\bar{h}^{0k}) \\
&= \partial_i[x^j(\partial^i\bar{h}^{0k} - \partial^0\bar{h}^{ik}) - x^k(\partial^i\bar{h}^{0j} - \partial^0\bar{h}^{ij}) + \delta^{ik}\bar{h}^{0j} - \delta^{ij}\bar{h}^{0k}]
\end{aligned} \tag{853}$$

To arrive at the final equality, we have dropped the term

$$(\delta^{ij}x^k - \delta^{ik}x^j)\partial_\mu\bar{h}^{\mu 0},$$

anticipating its vanishes identically when contracted with n_i over a spherical surface integral at infinity. Assembling now our results for P^μ and J^{jk}, we have:

$$P^0 = -\frac{c^4}{16\pi G}\int n_i\Theta^{i00}r^2 d\Omega = -\frac{c^4}{16\pi G}\int n_i(\partial^i\bar{h}^{00} - \partial^0\bar{h}^{i0} - \partial_\mu\bar{h}^{i\mu})r^2 d\Omega \tag{854}$$

$$P^j = -\frac{c^4}{16\pi G}\int n_i \Theta^{i0j} r^2 d\Omega = -\frac{c^4}{16\pi G}\int n_i(\partial^i \bar{h}^{0j} - \partial^0 \bar{h}^{ij} - \delta^{ij}\partial_\mu \bar{h}^{0\mu}) r^2 d\Omega \tag{855}$$

$$J^{jk} = -\frac{c^4}{16\pi G}\int [x^j(\partial^i \bar{h}^{0k} - \partial^0 \bar{h}^{ik}) - x^k(\partial^i \bar{h}^{0j} - \partial^0 \bar{h}^{ij}) + \delta^{ik}\bar{h}^{0j} - \delta^{ij}\bar{h}^{0k}] r^2 d\Omega \tag{856}$$

To proceed further, we require explict expressions for the Kerr metric $\bar{h}_{\mu\nu}$ in Cartesin (x, y, z) coordinates. From a simple inspection of equation (420), it is apparent that as $r \to \infty$, to leading order in $1/r$, h_{00} and h_{ij} will be identical to their Schwarzschild counterparts, as given by equation (326). (With $h = 0$, there is of course no distinction between $h_{\mu\nu}$ and $\bar{h}_{\mu\nu}$.) In particular, a^2 enters only via terms of order a^2/r^2 and higher. On the other hand, the Kerr metric has a leading order off-diagonal term $h_{0i} = h_{i0}$ which embodies the angular momentum effects at large distances from the hole. With $r^2 = x^2 + y^2 + z^2$ valid at large distances, standard spherical geometry may be used and

$$d\phi = \frac{1}{r^2 \sin^2\theta}(xdy - ydx)$$

. Then, equation (420) may be formally written:

$$-c^2 d\tau^2 \text{ (Kerr)} = -c^2 d\tau^2 \text{ (Schwarzschild)} - \frac{4GMa}{c^2 r^3}(cdt)(xdy - ydx), \qquad r \to \infty \tag{857}$$

At large distances from the hole, the Kerr metric reduces to a simple sum of the Schwarzschild metric plus off-diagonal angular momentum terms linear in a. Equation (326) gives

$$\bar{h}_{00} = \bar{h}^{00} = \frac{2GM}{rc^2}, \qquad \bar{h}_{ij} = \bar{h}^{ij} = \frac{2GMn_i n_j}{rc^2}, \quad \frac{n_i \equiv x_i}{r}, \tag{858}$$

From equation (857) we have:

$$\bar{h}_{0y} = -\bar{h}^{0y} = -\frac{2GMax}{c^2 r^3}, \qquad \bar{h}_{0x} = -\bar{h}^{0x} = \frac{2GMay}{c^2 r^3}. \tag{859}$$

Since the Kerr metric tensor is time independent, the energy P^0 takes the form

$$P^0 = -\frac{c^4}{16\pi G}\int n_i(\partial^i \bar{h}^{00} - \partial_j \bar{h}^{ij}) r^2 d\Omega \tag{860}$$

Then,

$$n_i \partial^i \bar{h}^{00} = \partial_r \bar{h}^{00} = -\frac{2GM}{r^2 c^2}, \qquad n_i \partial_j \bar{h}^{ij} = \frac{2GM}{r^2 c^2} \tag{861}$$

which leads to the welcome result for the "dressed mass" (material mass plus gravitational field) of a Kerr black hole of

$$P^0 = Mc^2. \tag{862}$$

For the Kerr angular momentum vector, taken to lie along the z-axis, the integral becomes

$$\begin{aligned} J^{xy} &= -\frac{c^3}{16\pi G}\int r^2 d\Omega n_i(x\partial^i \bar{h}^{0y} - y\partial^i \bar{h}^{0x} + \delta^{iy}\bar{h}^{0x} - \delta^{ix}\bar{h}^{0y}) \\ &= -\frac{c^3}{16\pi G}\int r^2 d\Omega(x\partial_r \bar{h}^{0y} - y\partial_r \bar{h}^{0x} + y\bar{h}^{0x}/r - x\bar{h}^{0y}/r) \end{aligned} \tag{863}$$

With $x^2 + y^2 = r^2 \sin^2\theta$, we have

$$x\partial_r \bar{h}^{0y} - y\partial_r \bar{h}^{0x} = -\frac{4GMa\sin^2\theta}{c^2 r^2}, \qquad \frac{1}{r}(y\bar{h}^{0x} - x\bar{h}^{0y}) = -\frac{2GMa\sin^2\theta}{c^2 r^2} \tag{864}$$

Substituting the above into our expression for J^{xy} leads to

$$J^{xy} = \frac{3}{8\pi} Mac \int \sin^2\theta d\Omega = Mac, \tag{865}$$

the angular momentum of a Kerr black hole.

11.3 Particle production and event horizons

In our opening discussion in Chapter 9 on gravitational radiation, we noted that there is currently no formulation of gravitation that is compatible with quantum field theory. When studying the quantum theory of electromagnetic radiation, it is helpful to begin with a semi-classical approach, in which an atom interacting with an external radiation field is treated via quantum theory, but the field is viewed as a classical continuum rather than as an ensemble of photons. The equivalent approach for general relativistic gravity was first carried out in 1974 by Stephen Hawking, who showed that quantum field theory, in the presence of a classical Schwarzschild geometry, predicts that particles will be spontaneously produced at the the event horizon. More precisely, the vacuum expectation value of the particle number operator at the event horizon is identical to that of a population in thermal equilibrium: an event horizon has a finite, well-defined temperature! This calculation, controversial for many years, is now regarded as one of the landmark achievements of twentieth-century theoretical physics. Its implications for fundamental physics are still being understood, and are very much a topic of current research.

In the wake of Hawking's achievement, researchers began to study quantum field theory in more simple geometries for the insights this may afford into the nature of particle production by spacetime curvature. Event horizons do not require an anchoring black hole for their existence. They can arise even in ordinary Minkowski space, if the metric is written in coordinates appropriate to a uniformly accelerating observer.[27] Do such flat-space event horizons also result in particle production? In 1976, Unruh showed that indeed they do.

The concept of Unruh radiation is just as remarkable as, and in mnay ways even more fundamental than, Hawking radiation. Begin with empty Minkowski space. There is no radiation of course, or it wouldn't be empty Minkowski space. Perform a coordinate transformation to accelerating coordinates. Suddenly, thermal radiation is present. Wait. If the stress energy tensor $T_{\mu\nu}$ vanishes to start with, a coordinate transformation should not create a $T_{\mu\nu}$: a zero tensor in one coordinate system transforms to zero in another. Right?

The issue here is far from trivial, or for that matter even settled. The received explanation is that it all comes down to the radiation detector (Birrell & Davies 1982). If a tree falls in a forest and no one is present, is there sound? Does radiation exist without a detector to detect it? While the coordinate change appears to bring about the radiation field, it is the detector—the observer—that is key to the argument. An accelerating observer detects the radiation produced by the event horizon the motion itself creates. But if there is an accelerating detector present, then from the point-of-view of a nonaccelerating observer, it is the detector itself that is producing the radiation. We are confronted squarely with the

[27] It may seem surprising that a photon travelling at c can never overtake a fleeing accelerating observer always travelling at less than c, but it is so. The gap shrinks as $t \to \infty$, but never quite reaches zero.

reality that the horizon radiation has no independent existence, outside of the presence of a detector to measure it! Neither, apparently, would the leptons and quarks, which are calculated to be produced as well. The entwining of observer, detector and reality brings us to the ontological heart of quantum mechanics. It is the observer that conjures reality into being. The falling tree in our forest analogy produces compressional waves in the air, with or without a conscious registration of sound, but it is the very notion of the existence of particles that is at issue here. Nowhere is this dilemma more stark than in Unruh radiation. The final word on this fascinating problem is yet to come.

The heart of the mathematical Unruh calculation is that, in a formal sense, the inertial vacuum state of a constant velocity observer and the so-called *Rindler* vacuum state of an accelerating observer are not the same quantum superposition of states in the vector space in which these vacua live: the multi-particle *Fock space*. In short, the reason that particles are seen by an acclerating observer is that, in order to express his Rindler vacuum in terms of the Minkowski Fock space states, the superposition requires the appearance of multi-particle vector states. In the calculation to follow, we shall not proceed along quite such formal lines, however. Rather, we shall fix the vacuum state right at the start, and focus instead upon the particle operators which act it. *We invoke the creation and annihilation operators designed for separate Rindler and Minkowski bases to understand the particle production.* In relating these bases to one another, the calculation will come down to a well-defined, if somewhat delicate, Fourier integral.

11.3.1 Minkowski and Rindler metrics

In this Chapter, $c = 1$.

We will need several different coordinate representations of Minkowski spacetime. Start with the standard Minkowski metric in Cartesian coordinates:

$$-d\tau^2 = -dt^2 + dx^2 + dy^2 + dz^2 \tag{866}$$

Introducing outgoing and ingoing photon coordinates,

$$\bar{u} = t - x, \quad \bar{v} = t + x, \tag{867}$$

the Minkowski metric may also be written as

$$-d\tau^2 = -d\bar{u}d\bar{v} + dy^2 + dz^2. \tag{868}$$

Transform now to accelerating observer ($'$) coordinates (Problem Set 1):

$$x = x' \cosh(at'), \quad t = x' \sinh(at') \tag{869}$$

the metric (866) becomes

$$-d\tau^2 = -(ax')^2 dt'^2 + dx'^2 + dy'^2 + dz'^2 \tag{870}$$

Note the coordinate (not structural) singularity at $x' = 0$, corresponding to the location of an observer-dependent event horizon. The character of the spacetime has not changed from Minkowski, but we shall henceforth follow custom and refer to these as Rindler coordinates and the Rindler metric.

The final set of coordinates required are known a *Lass* coordinates. These are defined by the transform $x' \to X'$ via

$$x' = e^{aX'}/a, \tag{871}$$

with t' untouched. The corresponding *null Lass* coordinates are

$$\bar{u}' = t' - X', \ \bar{v}' = t' + X'. \tag{872}$$

Note that the X' Lass coordinate covers only part of the spacetime: $x' > 0$, hence $x > 0$.

The respective metrics are

$$-d\tau^2 = e^{2aX'}(-dt'^2 + dX'^2) + dy'^2 + dz'^2 \qquad \text{(Lass)}, \tag{873}$$

and

$$-d\tau^2 = -e^{a(\bar{v}'-\bar{u}')}d\bar{u}'d\bar{v}' + dy'^2 + dz'^2 \qquad \text{(null Lass)}. \tag{874}$$

The simple coordinate transformation

$$\bar{u} = -e^{-a\bar{u}'}/a, \ \bar{v} = e^{a\bar{v}'}/a \leftrightarrow \bar{u}' = -\ln(-a\bar{u})/a, \ \bar{v}' = \ln(a\bar{v})/a \tag{875}$$

takes us back to the null (unaccelerated) Minkowski metric, $-d\bar{u}d\bar{v} + dy^2 + dz^2$. But in going from $\bar{u}', \bar{v}'$ back to Minkowski, once again only part of the spacetime is covered, namely $-\infty < \bar{u} < 0$, and $0 < \bar{v} < \infty$.

11.3.2 Scalar field equation

Our study of particle production will be restricted to the simplest case of a massles, scalar (spin 0) field Φ. In ordinary Minkowski spacetime, the behaviour of such fields is very well-understood. However, in the presence of an event horizon, notably an horizon in Rindler spacetime, the behaviour of such quantum fields is extremely subtle.

In standard Cartesian Minkowski space, the field Φ satisfies the Klein-Gordon equation

$$\partial_\mu \partial^\mu \Phi = 0. \tag{876}$$

In either curved spacetimes (or generalised coordinates), this becomes

$$g^{\mu\nu}(\partial_\mu \Phi)_{;\nu} = \frac{1}{\sqrt{g}}\partial_\mu \sqrt{g}\,(g^{\mu\nu}\partial_\nu \Phi) = 0, \tag{877}$$

where g is, as usual, $|\det g_{\mu\nu}|$. We will use the notation $\Box\Phi$ to indicate the operator in equation (877).

In Cartesian Minkowski coordinates x^μ, the equation $\Box\Phi = 0$ has classical solutions of the form

$$\Phi = A\exp(\pm ik^\mu x^\mu) \tag{878}$$

where the wave four-vector k^μ takes the form

$$k^\mu = (\omega, k_x, k_y, k_z), \quad \omega^2 = (k_x^2 + k_y^2 + k_z^2) \equiv k^2 \tag{879}$$

where $\omega > 0$ is an angular frequency. Remarkably, in Lass coordinates the field equation takes on *precisely the same form,* even though the metric (873) is quite different, a consequence of the metric's conformal structure. We may regard solutions of the form $\exp(\pm ik^\mu X'_\mu)$, as harmonic excitations in the Lass vacuum, particles in their on right, on an equal footing with those in the Minkowski vacuum. We are, in effect, describing the same underlying field with a different set of basis functions.

We consider next Rindler coordinates. Assuming an $e^{i\omega t'}$ time dependence and choosing the x' axis to coincide with the direction of the spatial wavevector k^i, the field equation for the spatial amplitude $\Phi(x')$ takes the form

$$\frac{\omega^2}{a^2 x'}\Phi + \frac{d}{dx'}x'\frac{d\Phi}{dx'} = 0, \tag{880}$$

with solution

$$\Phi(x') = \text{constant} \times (x')^{\pm i\omega/a} \sim \exp(\pm i\omega \ln(x')/a) \sim \exp(\pm i\omega X'). \tag{881}$$

These solutions thus correspond to $\exp(\pm i\omega\bar{u}')$ and $\exp(\pm i\omega\bar{v}')$, a result consistent with the $\Box\Phi = 0$ field equation,

$$\partial_{\bar{u}'}\partial_{\bar{v}'}\Phi = 0, \tag{882}$$

as well as with our above conclusions for the Lass basis functions. An identical equation also holds of course in standard Minkowski $(\bar{u}, \bar{v})$ coordinates, with solutions $\exp(\pm i\omega\bar{u})$ and $\exp(\pm i\omega\bar{v})$. The branch cut singularity evinced in (881) at $x' = 0$ is at this point purely formal, since the domain of x' is $x' > 0$. Nevertheless, the distinct mathematical form of these field solutions is in fact crucial to the emergence of particles from the vacuum, as we shall presently see.

11.3.3 Review of quantum field theory

The Bogoliubov formalsim allows us to go between two different bases in a quantised field theory, by expressing the creation and destruction of one in terms of the another.

We begin with a brief review of the foundations of quantum fields. (See the text of Schwarz 2017 for a thorough treatment.) In classical theory, any real solution to the standard Minkowski field equation may be written as a Fourier superposition over wavenumbers:

$$\Psi = a_j \exp(ik_j^\mu x_\mu) + a_j^* \exp(-ik_j^u x_\mu) \quad \text{(classical)} \tag{883}$$

where, as usual, we employ the convention that the repeated j index is summed over unless otherwise stated. (Here j is a label for the wavenumber mode, not a spatial index. μ is a spacetime coordinate index.) The Fourier coefficient a_j^* is the complex conjugate of a_j. In quantum field theory (QFT), these same coefficients are upgraded to creation and annihilation operators. The formalism of creation and annihilation operators is possible, including an explicit representation of the operators, because the classical field theory correspnds to the continuum limit of a collection of harmonic operators: in effect, masses on springs. QFT is simply the quantum mechanics of this collection of harmonic oscillators, with Heisenberg's raising and lowering operators taken over directly. The lowest excited state of the oscillators is a field particle, higher excited states are many field particles. Creation and destruction of these field particles corresponds to respective applications of the raising and lowering operators, as shown by Dirac in his famous 1928 photon paper.

In QFT, the Fourier amplitude a_j becomes a destruction operator, lowering an excited state by the removal of a particle of wavenumber k_j^μ. The Fourier ampliude a_j^* becomes a creation operator, creating a particle of the same wavenumber from a lower state. Since this amplitude is now an operator, the complex conjugate becomes a hermitian conjugate and is written $a^\dagger$. The a-operators satisfy the commutation relation

$$[\, a_i, a_j^\dagger \,] \equiv a_i a_j^\dagger - a_j^\dagger a_i = \delta_{ij} \tag{884}$$

All a_j commute with all other a_j, and the same for the $a_j^\dagger$.

The field Ψ, now a formal operator, needs to have objects to operate upon. These objects are close analogues to the familar state vectors in Hilbert space one encounters in the formalism of ordinary quantum mechanics. Here these states are "upgraded" to many-particle states in a *Fock space.* We use Dirac bra and ket notation: the vacuum state is denoted $|0\rangle$, the state with n_j particles of wavenumber k_j^μ is $|n_j\rangle$, the state with n_i particles with wavenumber k_i^μ and n_j particles with wavenumber k_j^μ is $|n_i, n_j\rangle$, and so forth, throughout the entire multiparticle Fock space. Then the creation and annihilation operators must behave in such a way that

$$a^\dagger|n_j\rangle = C_+|n_j+1\rangle, \quad a|n_j\rangle = C_-|n_j-1\rangle, \tag{885}$$

where C_+ and C_- are constants to be determined.

We require an operator $\hat{N}_j$ that returns the number of particles in wavenumber state j. This is simply

$$\hat{N}_j \equiv a_j^\dagger a_j = a_j a_j^\dagger - 1. \tag{886}$$

To see how this works, we proceed by induction. Note first that $\hat{N}_j$ returns zero when applied to the vaccum state $|0\rangle$. Dropping the j subscript, if we next assume that $\hat{N}|n\rangle = n|n\rangle$, then application of $\hat{N}$ to the state

$$a^\dagger|n\rangle = C_+|n+1\rangle$$

gives

$$\hat{N}a^\dagger|n\rangle = a^\dagger a a^\dagger|n\rangle = a^\dagger(1+a^\dagger a)|n\rangle = (n+1)a^\dagger|n\rangle, \tag{887}$$

which proves our claim. It is now a simple matter to determine C_+:

$$C_+^2 = \langle n+1|C_+^2|n+1\rangle = \langle n+1|aa^\dagger|n+1\rangle = \langle n+1|1+a^\dagger a|n+1\rangle, = n+1 \tag{888}$$

and the reader should now have little difficulty showing that a similar calculation on $|n-1\rangle$ gives

$$C_-^2 = n. \tag{889}$$

To reiterate: the vaccum state is not unique, not all vacuums are the same. Different observers will interpret the vacuum differently even within the same classical spacetime manifold. We shall henceforth denote the Minkowski vacuum state by $|0\rangle_M$ and the vacuum of an accelerating observer described by Rindler coordinates as $|0\rangle_R$.

11.3.4 Bogoliubov formalism

The equation

$$\Psi = (e^{-ik_j^\mu x_\mu} a_j^\dagger + e^{ik_j^\mu x_\mu} a_j)|0\rangle_M \tag{890}$$

states that the Ψ field is a superposition of particles states, labeled by a (summed) fiducial index j, that are created and destroyed from the vacuum state $|0\rangle$. The destruction operators a_j acting on the shown vacuum return a value of zero of course, which is not very interesting, but we use this particular example to establish our notation.

The Rindler vacuum has its own accelerating observer's X' coordinates, which are not the same as x, but are just as valid, satisfying the same harmonic oscillator field equation, $\Box\Phi = 0$:

$$\Psi = (e^{-ik_j'^\mu X'_\mu} b_j^\dagger + e^{ik_j'^\mu X'_\mu} b_j)|0\rangle_R. \tag{891}$$

We use a distinct b notation for the Rindler creation and annihilation operators because they are distinct operators. *The essence of the Bogoluibov approach is that a_i and b_i are expressible as superpositions of one another:*

$$a_i = \alpha_{ij} b_j + \beta_{ik} b_k^\dagger, \quad a_i^\dagger = \alpha_{ij}^* b_j^\dagger + \beta_{ik}^* b_k. \tag{892}$$

Note the possibility that a creation (annihilation) operator in one vacuum contains an annihilation (creation) operator from another! One observer's vacuum may thereby contain particles as measured by another observer.

The Rindler number operator for particles with momentum state i in a state n is

$$N_i(n) = {}_R\langle n | b_i^\dagger b_i | n \rangle_R. \tag{893}$$

If we apply the number operator $a_i^\dagger a_i$ to the Minkowski vacuum $|0\rangle_M$, it will of course vanish. But what if we apply this number operator to the Rindler vacuum $|0\rangle_R$? Does

$${}_R\langle 0 | a_i^\dagger a_i | 0 \rangle_R = 0 \ ? \tag{894}$$

We may answer this question by expressing the $a's$ (Minkowski) in terms of the $b's$ (Rindler) via (892). In the cross multiplication of $\hat{N}_i = a_i^\dagger a_i$ (no sum i), the only possible surviving term on the Rindler vacuum expectation is

$${}_R\langle 0 | \beta_{ij}^* \beta_{ik} b_j b_k^\dagger | 0 \rangle_R = {}_R\langle 0 | \beta_{ij}^* \beta_{ik} (\delta_{jk} + b_k^\dagger b_j) | 0 \rangle_R = {}_R\langle 0 | \beta_{ij}^* \beta_{ij} | 0 \rangle_R = \beta_{ij}^* \beta_{ij} \tag{895}$$

as all other terms ($b^\dagger b, bb, b^\dagger b^\dagger$) are eliminated. The vacuum expectation value for the particle occupation number is then equal to (summed over j, not i):

$$N_i(0)_R = {}_R\langle 0 | a_i^\dagger a_i | 0 \rangle_R = \beta_{ij}^* \beta_{ij} \equiv \sum_j |\beta_{ij}|^2 \tag{896}$$

The presence of β-mixing coefficients are critical. These are present when one vacuum's pure annihilation operator contains a piece of another's creation operator. When this occurs, the so-called vacuum will not be a state of zero particle number.

11.3.5 Invariant inner product

To determine the individual contributions from a superposition of basis functions, an invariant inner product operation must be defined. For functions f_1 and f_2 which satisfy the (covariant) KG equation, a suitable formal inner product is:

$$(f_1, f_2) = -i \int \sqrt{\gamma}\, d^3x\, n^\mu [f_1 (f_2^*)_{;\mu} - f_2^* (f_1)_{;\mu}] \tag{897}$$

where the integration is over a spacelike hypersurface Σ, and the overall $-i$ phase factor is chosen for convenience. The inner product is constructed so that it is independent of the particular choice for Σ. n^μ is a unit vector normal to Σ ($n^\mu n^\nu g_{\mu\nu} = -1$), and γ is the absolute value of the determinant of the spacelike metric for Σ.

Here, we shall be interested in metrics for which $g_{0i} = 0$ vanish, in which case we may choose a purely timelike normal $n^\mu = (n^0, 0, 0, 0)$, and μ then takes on only its 0 time index.

The n^μ normalisation condition implies $n^0 = 1/\sqrt{-g_{00}}$. To motivate this formulation for the inner product, consider the 4-current J^μ defined in terms of the scalar fields ψ and ϕ by

$$J^\mu = \psi^* \partial^\mu \phi - \phi \partial^\mu \psi^*, \tag{898}$$

where ϕ, ψ and their complex conjugates all satisfy the covariant KG equation. Then,

$$J^\mu_{;\mu} = \frac{1}{\sqrt{g}} \partial_\mu (\sqrt{g} J^\mu) = (\partial_\mu \psi^*)\partial^\mu \phi - (\partial_\mu \phi)\partial^\mu \psi^* + \psi^* (\partial^\mu \phi)_{;\mu} - \phi (\partial^\mu \psi^*)_{;\mu} = 0, \tag{899}$$

because the first two terms on the right cancel and the last two term vanish by the KG equations. This means

$$\partial_t \int \sqrt{|g|} J^0 d^3x = - \int_\infty \sqrt{|g|} J^j n_j dS \to 0, \tag{900}$$

where the integral on the right is taken over a surface at infinity (j here is a spatial index!). In this sense, the J^0 integral is an invariant conserved quantity. With our $g_{0i} = 0$ assumption, it is easy to see that $g = |g_{00}\gamma|$. (In fact, this determinant relationship is true more generally as we showed in Problem Set 1.) Hence, the integrand beomes

$$|g|^{1/2} J^0 = |g_{00}\gamma|^{1/2} g^{00} J_0 = -|\gamma / g_{00}|^{1/2} J_0 = -|\gamma|^{1/2} n^0 J_0. \tag{901}$$

Up to an unimportant overall phase, $-\int d^3x |\gamma|^{1/2} n^0 J_0$ is the same invariant inner product of equation (897).

Consider now how plane wave solutions to the KG equation behave in the inner product. We have

$$f_1 = e^{ik_1^\mu x_\mu}, \quad f_2 = e^{ik_2^\mu x_\mu}, \ \partial_t f_2^* = i\omega_2 e^{-ik_2^\mu x_\mu}, \ \partial_t f_1 = -i\omega_1 e^{ik_1^\mu x_\mu} \tag{902}$$

$$f_1 \partial_t f_2^* = i\omega_2 e^{i(k_1^\mu - k_2^\mu)x_\mu}, \ f_2^* \partial_t f_1 = -i\omega_1 e^{i(k_1^\mu - k_2^\mu)x_\mu}, \tag{903}$$

$$f_1 \partial_t f_2^* - f_2^* \partial_t f_1 = i(\omega_1 + \omega_2) e^{-i(k_1^\mu - k_2^\mu)x_\mu}. \tag{904}$$

This leads to

$$(f_1, f_2) = (e^{ik_1^\mu x_\mu}, e^{ik_2^\mu x_\mu}) = (\omega_1 + \omega_2) e^{i(\omega_1 - \omega_2)t} (2\pi)^3 \delta^3(\mathbf{k_1} - \mathbf{k_2}). \tag{905}$$

Using this result, it is convenient to renormalise the f-functions to unit norm:

$$f_i \equiv \frac{e^{ik_i^\mu x_\mu}}{\sqrt{16\pi^3 \omega_i}} \qquad \text{(renormalised)} \tag{906}$$

Note that

$$(e^{-ik_1^\mu x_\mu}, e^{ik_2^\mu x_\mu}) = (\omega_1 - \omega_2) e^{i(\omega_1 + \omega_2)t} (2\pi)^3 \delta^3(\mathbf{k_1} + \mathbf{k_2}) \tag{907}$$

Since the $\omega > 0$, the above must vanish. It is now straightforward to show that, with f_i defined as in (906),

$$(f_i, f_j) = -(f_i^*, f_j^*) = \delta_{ij}, \quad (f_i, f_j^*) = 0 \tag{908}$$

Note the sign change between the f and f^* inner products.

If we now form a second orthonormal g-basis from a superposition of f-basis functions,

$$g_i = \alpha_{ik} f_k + \beta_{ik} f_k^* \tag{909}$$

then by equation (908) the α and β coefficients are

$$(g_i, f_j) = \alpha_{ij}, \quad (g_i, f_j^*) = -\beta_{ij}. \tag{910}$$

Therefore, using the inner product rules (908):

$$(g_i, g_j) = \delta_{ij} = \sum_k (\alpha_{ik} f_k + \beta_{ik} f_k^*, \alpha_{jl} f_l + \beta_{jl} f_l^*),$$

$$= \sum_k \alpha_{ik}[(f_k, f_l)\alpha_{jl}^* + \beta_{ik}(f_k^*, f_l^*)\beta_{jl}^*]$$

or

$$\delta_{ij} = \sum_k (\alpha_{ik}\alpha_{jk}^* - \beta_{ik}\beta_{jk}^*). \tag{911}$$

Setting $i = j$, we obtain the important result

$$1 = \sum_k (|\alpha_{ik}|^2 - |\beta_{ik}|^2) \to \int \left(|\alpha_{\omega\omega'}|^2 - |\beta_{\omega\omega'}|^2|\right) d\omega' \tag{912}$$

where the expression following the arrow corresponds to associating the label k with a continuous frequency distribution. Following equation (896), we anticipate that if their is a relation between α and β which takes the form $|\alpha_{\omega\omega'}|^2 = n^2|\beta_{\omega\omega'}|^2$ where n^2 is a constant other than unity, then (912) implies:

$$N_i(0) = {}_R\langle 0|a_i^\dagger a_i|0\rangle_R = \int |\beta_{\omega\omega'}|^2 d\omega' = \frac{1}{n^2 - 1} \tag{913}$$

In other words, the vacuum expectation value for the particle number operator would be finite! (More typically, β vanishes for the vacuum state and n^2 is formally infinite.) Let us see how a finite n^2 arises between the Rindler and Minkowski bases functions.

11.3.6 Minkowski and Rindler vacuums

In Minkowski coordinates, the one-dimensional outgoing basis functions are proportional to $e^{-i\omega\bar{u}}$; in Rindler coordinates, the outgoing basis functions correspond to $e^{-i\omega\bar{u}'}$:

$$f_{\omega'} = \frac{e^{ik'^\mu x_\mu}}{\sqrt{2\pi\omega'}} = \frac{e^{-i\omega'\bar{u}}}{\sqrt{2\pi\omega'}} \quad \text{(Minkowski)} \tag{914}$$

$$g_\omega = \frac{e^{ik^\mu X'_\mu}}{\sqrt{2\pi\omega}} = \frac{e^{-i\omega\bar{u}'}}{\sqrt{2\pi\omega}} \quad \text{(Rindler)} \tag{915}$$

The Bogoliuobov coefficients are:

$$\alpha_{\omega\omega'} = (g_\omega, f_{\omega'}), \qquad \beta_{\omega\omega'} = -(g_\omega, f_{\omega'}^*), \tag{916}$$

The integrations are carried out in Minkowski coordinates, so that the integration variable is $\bar{u}$, not x, and $\bar{u}$ depends upon time as well as x. An effective way to handle this is to do the original integral over $\delta(t - t_0)dx\,dt$, where t_0 is simply a fiducial time. Then

$$\int \delta(t - t_0)dx\,dt = \int \delta\left(\frac{\bar{u} + \bar{v}}{2} - t_0\right) \left|\frac{\partial(x,t)}{\partial(\bar{u},\bar{v})}\right| d\bar{u}\,d\bar{v} = \int d\bar{u} \tag{917}$$

where $|\partial(x,t)/\partial(\bar{u},\bar{v})|$ is the full Jacobian. The Jacobian is just 1/2, and the delta function integration over $\bar{v}$ returns a factor of 2 and nothing else, so this is all simply reduces to $d\bar{u}$! Finally:

$$\partial_t = \partial_t \bar{u}\, \partial_{\bar{u}} + \partial_t \bar{v}\, \partial_{\bar{v}} = \partial_{\bar{u}}, \tag{918}$$

since $\partial_t \bar{u}$ is unity and nothing depends on $\bar{v}$. Note that while $\bar{u}'$ ranges from $+\infty$ to $-\infty$, the variable $\bar{u}$ runs from $-\infty$ to 0 (equation [875]).

Expressions for $\alpha_{\omega\omega'}, \beta_{\omega\omega'}$ are:

$$\alpha_{\omega\omega'} = -i\int_{-\infty}^{0} [g_\omega\, \partial_{\bar{u}} f^*_{\omega'} - f^*_{\omega'}\, \partial_{\bar{u}} g_\omega] d\bar{u} = -2i\int_{-\infty}^{0} g_\omega\, \partial_{\bar{u}} f^*_{\omega'}\, d\bar{u}, \tag{919}$$

$$\beta_{\omega\omega'} = i\int_{-\infty}^{0} [g_\omega\, \partial_{\bar{u}} f_{\omega'} - f_{\omega'}\, \partial_{\bar{u}} g_\omega] d\bar{u} = 2i\int_{-\infty}^{0} g_\omega\, \partial_{\bar{u}} f_{\omega'} d\bar{u}, \tag{920}$$

where we have integrated the second term by parts in each case. With

$$g_\omega \partial_{\bar{u}} f^*_{\omega'} = i\frac{e^{i\omega'\bar{u} - i\omega\bar{u}'}}{2\pi}\sqrt{\frac{\omega'}{\omega}} \tag{921}$$

$$g_\omega \partial_{\bar{u}} f_{\omega'} = i\frac{e^{-i\omega'\bar{u} - i\omega\bar{u}'}}{2\pi}\sqrt{\frac{\omega'}{\omega}} \tag{922}$$

we see that $\beta_{\omega\omega'} = \pm i\alpha_{(\omega)(-\omega')}$. (The sign ambiguity is present because we must choose a branch of the square root, and we do so below. It is not a concern if we are interested only in the modulus.) Thus, to obtain $\beta_{\omega\omega'}$, we need only calculate the $\alpha_{\omega\omega'}$ coefficient. This is given by (875), (919) and (921):

$$\alpha_{\omega\omega'} = \frac{1}{\pi}\sqrt{\frac{\omega'}{\omega}}\int_{-\infty}^{0} e^{-i\omega\bar{u}' + i\omega'\bar{u}} d\bar{u} = \frac{1}{\pi}\sqrt{\frac{\omega'}{\omega}}\int_{-\infty}^{0} e^{i\omega\ln(-\bar{u}a)/a} e^{i\omega'\bar{u}} d\bar{u} \tag{923}$$

Rearranging the integral a bit, with $x = -\bar{u}$,

$$\alpha_{\omega\omega'} = \frac{a^{i\omega/a}}{\pi}\sqrt{\frac{\omega'}{\omega}}\int_{0}^{\infty} e^{i\omega\ln(x)/a} e^{-i\omega' x} dx = \frac{a^{i\omega/a}}{\pi}\sqrt{\frac{\omega'}{\omega}}\int_{0}^{\infty} x^{i\omega/a} e^{-i\omega' x} dx \tag{924}$$

Finally, with $y = \omega' x$,

$$\alpha_{\omega\omega'} = \frac{a^{i\omega/a}}{\pi}\sqrt{\frac{\omega'}{\omega}}(\omega')^{-1-i\omega/a}\int_{0}^{\infty} y^{i\omega/a} e^{-iy} dy \tag{925}$$

To calculate the $\beta_{\omega\omega'}$ coefficient, substitute $\omega' \to -\omega'$ into the above, noting that the integral (which is a complex-valued Γ-function) does not change. It is the single factor $\omega'^{-i\omega/a}$ that makes the calculation interesting, since none of the other terms change the modulus when ω' changes its sign. The integral (924) defines an analytic function only in the *lower* complex plane, which determines the appropriate branch of $-\omega' = e^{-i\pi}\omega'$. Hence

$$|\beta_{\omega\omega'}|^2 = |\alpha_{(\omega)(-\omega')}|^2 = e^{-2\pi\omega/a}|\alpha_{\omega\omega'}|^2. \tag{926}$$

Returning to equation (913), the particle occupation number is found to be

$$N_i(0) = \frac{1}{e^{2\pi\omega/a} - 1}. \tag{927}$$

Astonishingly, the is the particle occupation number of a thermal population of scalar particles! Putting c back in the equation and identifying the argument of the exponent with $\hbar\omega/k_B T_U$, defines the "Boltzmann-Unruh temperature T_U" of the particle distribution:

$$T_U = \frac{\hbar a}{2\pi k_B c} = 4.055 \times 10^{-21} a \quad \text{K}, \tag{928}$$

where a is expressed in m s^{-2}. ($\hbar$ and k_B are the Planck and Boltzmann constants). We conclude that in an otherwise empty Minkwoski space, an accelerating observer would measure a thermal background of particles due to the presence of an horizon at $x' = 0$. The temperature is barely above absolute zero for accelerations significantly below $\sim 10^{20}$m s^{-2}. Nevertheless, there are currently proposals for measuring the Unruh-effect radiation in the laboratory via the quantum mechanical Berry phase[28].

Let us recap our earlier discussion, now that the calculation is complete. The interpretation of the Unruh effect is at the time of this writing not entirely free of controversey, with some authors questioning whether radiation is actually produced. The detector itself is key, bringing to the fore the ontological entanglement between quantum physics and the presence of an observer. The current consensus is that the accelerated Rindler observer registers thermal particle absorption on their detector, whereas the inertial Minkowski observer views this as an emission from the detector with an accompanying radiation reaction. More generally, the Unruh effect is conceptually linked to, and its initial investigation motivated by, particle production at event horizons, to which we now turn.

11.3.7 The Hawking effect

We shall restrict our attention to the Hawking effect in its simplest form, the production of massless scalars particles at the event horizon of a Schwarzschild black hole. We begin with the geodesic equation for a radially moving, freely-falling observer (§6.5.2). It is identical to the Newtonian case:

$$\frac{d^2 r}{d\tau^2} = -\frac{GM}{r^2} \tag{929}$$

Consider next a static observer just outside the event horizon at $r = R_S = 2GM/c^2$. The proper acceleration felt by a stationary observer is $d^2 l/d\tau^2$, where the proper length dl is given by

$$dl = \frac{dr}{(1 - R_S/r)^{1/2}},$$

and thus

$$\frac{d^2 l}{d\tau^2} = \frac{1}{(1 - R_S/r)^{1/2}} \frac{d^2 r}{d\tau^2},$$

since the $dr/d\tau$ terms vanish for a stationary observer. In order to ensure that the second-order τ derivative also vanishes, an additional external forcing of $+GM/r^2$ must be added to the right side of the coordinate equation (929) for $d^2 r/d\tau^2$, which then translates to a *proper* counter-accleration of

$$a(r) = \frac{d^2 l}{d\tau^2} = \frac{GM}{R_S^2 \left(1 - R_S/r\right)^{1/2}} = \frac{c^4}{4GM \left(1 - R_S/r\right)^{1/2}} \tag{930}$$

[28] Martin Martinez, E.; Fuentes, I.; Mann, R. B. (2011), "Using Berry's Phase to Detect the Unruh Effect at Lower Accelerations," Physical Review Letters, 107 (13): 131301

Hence, a static observer hovering just above the event horizon feels a proper physical acceleration, not of GM/R_S^2, but of $(1 - R_S/r)^{-1/2}$ times this value, which becomes huge near R_S. (If the static observer could stand on a scale, this is the weight that would be shown.) More significantly for our purposes, an inertial, freely-falling, inward-moving observer passing the same location would perceive the static observer accelerating upward at this rate. To this locally Minkowski observer, this looks just like the Unruh problem! Hence, the detector of such a static observer would record a thermal spectrum with Unruh temperature

$$T_U = \frac{\hbar a(r)}{2\pi kc} = \frac{\hbar c^3}{8\pi kGM\left(1 - R_S/r\right)^{1/2}}. \tag{931}$$

This, however, is the *local* radiation temperature, not the radiation temperature detected by an observer at infinity. The latter radiation, received at large distances, is redshifted by an offsetting factor of

$$\sqrt{-g_{00}} = \left(1 - R_S/r\right)^{1/2}.$$

In other words, the so-called *Hawking temperature* of particle radiation observed by a detector an infinite distance from a Schwarzschild black hole is

$$T_H = \frac{\hbar c^3}{8\pi k_B GM} = 6.17 \times 10^{-8} \left(\frac{M\odot}{M}\right) \text{ K} \tag{932}$$

This is Hawking's celebrated 1974 result: a black hole radiates particles with a thermal spectrum at this temperature when viewed in the far field of its event horizon. Energy is, of course, not spontaneously generated at a black hole event horizon. The source is the black hole energy itself, Mc^2. The emission of a particle of energy $\hbar\omega$ is accompanied by a reduction of Mc^2 by an identical amount. That is to say, black holes *evaporate.*

11.3.8 Hawking luminosity and Bekenstein-Hawking entropy

Let us assume that the thermal luminosity of a black hole's event horizon is dominated by spin 1 photons as opposed to the simple massless spin 0 particles of our calculation. The emitted luminosity from the event horizon of a Schwarzschild black hole would then be

$$L_H = 4\pi R_S^2 \sigma T_H^4 = \frac{\hbar c^6}{15360\pi G^2 M^2}, \tag{933}$$

where σ is the Stefann-Boltzmann constant. In the real Universe, there is thermal background of photons at 2.3K, which means that any known black hole would actually be far too cold to emit, but would instead *grow* by the net absorption of cosmic microwave backgrounds photons alone! (To exceed a Hawking temperature of 2.7K, the black hole mass would need to be below 4.5×10^{22}kg, about 60% of the Earth's moon.) Let us ignore this practicality, and assume that the black hole is radiating into a vacuum at absolute zero. The evaporation time is then straightforward to calculate, a task we shall leave to the reader:

$$t_{\text{ev}} = \frac{5120\pi G^2 M^3}{\hbar c^4} = 2.1 \times 10^{67} \left(\frac{M}{M_\odot}\right)^3 \text{ y}, \tag{934}$$

some 57 orders of magnitude greater than the age of the Universe! One would need a mass less than $\sim 1.7 \times 10^{11}$ kg, about the size of a large mountain, with a correspnding $T_H = 7 \times 10^{11} K$, to bring the evaporation lifetime of a black hole within the age of the Universe. There is

no known mechanism to produce such a small mass black hole in the course of stellar or galactic evolution, but gravitational collapse in the early Universe may, in principle, lead to such tiny black hole masses. If this occurred, such sources would today be approaching the $M \to 0$ limit, and therefore radiating high energy gamma rays explosively, according to equations (932) and (933). Astronomical gamma ray sources are, in fact, routinely observed, but none to date exhibit the spectral properties that would make them suitable candidates for evaporating primordial black holes.

Accompanying this finite temperature there must be a finite entropy, which may be found from the first law of thermodynamics, $dE = TdS$. Applied to a black hole of mass M, $E = Mc^2$ and $T = T_H$, so that a vast amount of energy is associated with a tiny temperature. Hence, the entropy is

$$S = \int \frac{dE}{T} = \frac{8\pi k_B G}{\hbar c} \int M dM = \frac{4\pi k G M^2}{\hbar c} = \frac{kAc^3}{4G\hbar}. \tag{935}$$

where A is the area of the event horizon, $4\pi R_S^2$. Hawking (1971) had already established that, during dynamical black hole interactions, the total area of all event horizons must always increase, never decrease (e.g., by the emission of gravitational radiation). This must-always-increase property is curiously reminscent of entropy in thermodynamic interactions. Motivated by this finding, Bekenstein (1972) argued that the area/entropy correspondence was no mere "analogy," there really was a true black hole entropy, and it was directly proportional to the event horizon surface area! On dimensional grounds, the dimensionless entropy S/k_B must then be proportional to $Ac^3/g\hbar$, but neither Bekenstein nor anyone else could deduce the proportionality constant in the absence of any theory. But it didn't take long. Stephen Hawking's classic 1974 paper did just that. Equation [935] is often referred to as the Bekenstein-Hawking entropy, hereafter $S_{\rm BH}$.

In dimensionless form,

$$\frac{S_{\rm BH}}{k} = \frac{Ac^3}{4G\hbar} = \frac{A}{4l_P^2} = 1.05 \times 10^{77} \left(\frac{M}{M_\odot}\right)^2 \tag{936}$$

where $l_P = (c^3/G\hbar)^{1/2}$ is the Planck length. This is a huge number. By way of comparison, the dimensionless Sackur-Tetrode entropy for the Sun is about 3.5×10^{58}, a factor of $\sim 10^{-18.5}$ smaller than the Hawking entropy associated with a black hole formed from the Sun! The entropy of the Universe appears to be utterly dominated by the event horizons of its constituent black holes.

The concept of entropy was introduced into physics in the course of the development of thermodynamics by Clausius in the mid-nineteenth century. Later (1877), an underyling and profound statistical interpretation of entropy was advanced by Boltzmann via his famous statistical formula $S = k \ln \Omega$, where Ω is the number of internal microstates that give rise to a specified macrostate. We have just seen that there is a thermodynamic entropy associated with black holes. Is there an underlying statistical interpretation associated with this entropy in the form of counting internal microstates? The fact that $S_{\rm BH}/k_B$ is, up to a factor of order unity, the area of the event horizon area measured in square Planck lengths, is very suggestive.

In 1996, Strominger and Vafa showed that, provided one is willing to work in five dimensions, and with a rather extreme manifestation of black hole (one carrying electric and a more exotic "axion" charge), string theory provides the required mathematical framework for carrying out just such a statistical calculation. This is done by counting the degeneracy of the relevant solution of the governing field equations for such a black hole. Because of the special symmetries evinced by this extreme black hole solution, it the simplest possible problem of its class. The Boltzmann-like statistical entropy that emerges from such a calculation

agrees precisely with the Bekenstein-Hawking entropy (936). It is widely believed that the entropy of less extreme astrophysical black holes has an analogous statistical basis, although at the time of this writing, only the Strominger-Vafa case has been explicitly calculated.

The association of event horizons and accelerated observers with thermal particle production, the understanding that the quantum vacuum is a thermal state, the well-defined and enormous entropy of black holes, the underlying statistical basis for Hawking entropy in string theory—these are results of enormous significance from late twentieth-century physics Today, the quest to understand the behaviour of quantum fields in curved spacetimes is very much at the forefront of modern physics. There are many unanswered questions. Is quantum information destroyed at the event horizon of a black hole, in contradiction to a foundational principle of quantum mechanics? A rigorous answer to this question will likely require a fully-developed theory of quantum gravity, whose beginnings perhaps may have been previewed by the black hole-entropy connection we have been begun to explore here.

11.4 The Raychaudhuri equation

In 1955, A. K. Raychaudhuri derived a very general and elegant equation describing how the 4-velocity field of a test particles evolves as they move through spacetime. More precisely, the *Raychaudhuri equation* (RE) describes how the 4-velocity divergence, $\theta \equiv U^{\mu}{}_{;\mu}$, changes with time in the presence of shear, vorticity and spacetime curvature. The equation itself is a purely mathematical statement for any spacetime. In general relativity theory, the curvature may in turn be directly expressed related to components of the stress energy tensor. By construction, the 4-divergence measures how a comoving 4-volume of particles changes with time. The power of the RE is its generality: it places no restrictions upon the symmetry of the flow, and only very weak restrictions upon the equation of state. The equation plays a central role in the Penrose-Hawking singularity theorems of general relativity, in essence by demanding that if all future-directed geodesics are converging in the spacetime (in particular the null geodesics), they can never subsequently diverge. Indeed, the area of the hypersurface pierced by a bundle of such converging geodesics will vanish in a finite amount of time. And that is what leads to a singularity.

We start our derivation by recalling equation (235), which defines the curvature tensor $R^{\mu}{}_{\lambda\sigma\kappa}$ in terms of the commutator of the covariant derviatives.

$$U^{\mu}{}_{;\sigma;\kappa} - U^{\mu}{}_{;\kappa;\sigma} = R^{\mu}{}_{\lambda\sigma\kappa}U^{\lambda} \tag{937}$$

Here, U^{μ} is taken to be a timelike vector, which we are free to think of as a 4-velocity. Contracting μ and σ, and multiplying by U^{κ} gives

$$U^{\kappa}\theta_{;\kappa} - U^{\kappa}U^{\mu}{}_{;\kappa;\mu} = R_{\lambda\kappa}U^{\lambda}U^{\kappa}, \tag{938}$$

or

$$U^{\kappa}\theta_{;\kappa} - (U^{\kappa}U^{\mu}{}_{;\kappa})_{;\mu} + U^{\kappa}{}_{;\mu}U^{\mu}{}_{;\kappa} = R_{\lambda\kappa}U^{\lambda}U^{\kappa}. \tag{939}$$

We write this more compactly as

$$\dot{\theta} - (\dot{U}^{\mu})_{;\mu} + U^{\kappa}{}_{;\mu}U^{\mu}{}_{;\kappa} = R_{\lambda\kappa}U^{\lambda}U^{\kappa}, \tag{940}$$

where we have identified the standard covariant time derivative $D/D\tau$ along the 4-velocity curve using the dot notation $\dot{X} = U^{\mu}X_{;\mu}$. (Here, X is a tensor of arbitrary rank, indices suppressed). Notice that this equation for $\dot{\theta}$ is purely geometrical: we have have not yet used general relativity.

We next introduce two important spacelike tensors, $\sigma_{\mu\nu}$ and $\omega_{\mu\nu}$, which are a measure of the flow shear and vorticity, respectively. Both are linear in the space and time covariant derivatives of U_μ, and in fact reduce to standard expressions for the shear and vorticity (e.g. Landau & Lifschitz 1959) in the local rest frame. While their interpretation in terms of flow properties enhances our understanding of the equations to follow, this is not at all essential to our formal results. These tensors are defined as follows:

$$\sigma_{\mu\nu} = \frac{1}{2}\left(U_{\mu;\nu} + U_{\nu;\mu}\right) + \frac{1}{2}\left(\dot{U}_\mu U_\nu + \dot{U}_\nu U_\mu\right) - \frac{\theta}{3}\left(g_{\mu\nu} + U_\mu U_\nu\right) \tag{941}$$

$$\omega_{\mu\nu} = \frac{1}{2}\left(U_{\mu;\nu} - U_{\nu;\mu}\right) + \frac{1}{2}\left(\dot{U}_\mu U_\nu - \dot{U}_\nu U_\mu\right) \tag{942}$$

The tensor $\sigma_{\mu\nu}$ is constructed so that *i)* it is linear in the first order velocity derivatives; *ii)* it is symmetric in its indices; *iii)* its trace vanishes; and *iv)* it satisfies $U^\mu\sigma_{\mu\nu} = 0$. Up to an overall arbitrary normalisation, these four constraints uniquely determine $\sigma_{\mu\nu}$. Note that $\sigma_{0\nu} = \sigma_{\nu 0} = 0$ in the local rest frame of the flow, and in this frame $\sigma_{\mu\nu}$ reduces to the standard, purely spacelike shear tensor σ_{ij}.

The tensor $\omega_{\mu\nu}$ follows along similar lines. It is constructed so that *i)* it is linear in the velocity derivatives; *ii)* it is anti-symmetric in its indices; *iii)* its trace vanishes; and *iv)* it satisfies $U^\mu\omega_{\mu\nu} = 0$. Up to an overall arbitrary normalisation, these four constraints uniquely determine $\omega_{\mu\nu}$. In common with $\sigma_{\mu\nu}$, all 0-indexed components of $\omega_{\mu\nu}$ vanish in a local rest frame, and $\omega_{\mu\nu}$ then reduces to the standard spatial vorticity tensor ω_{ij}.

It is easily shown that

$$\sigma_{\mu\nu} + \omega_{\mu\nu} = U_{\mu;\nu} + \dot{U}_\mu U_\nu - \frac{\theta}{3}(g_{\mu\nu} + U_\mu U_\nu), \tag{943}$$

$$\sigma_{\mu\nu} - \omega_{\mu\nu} = U_{\nu;\mu} + \dot{U}_\nu U_\mu - \frac{\theta}{3}(g_{\mu\nu} + U_\mu U_\nu), \tag{944}$$

and therefore the product

$$(\sigma_{\mu\nu} + \omega_{\mu\nu})(\sigma^{\mu\nu} - \omega^{\mu\nu}) = \sigma_{\mu\nu}\sigma^{\mu\nu} - \omega_{\mu\nu}\omega^{\mu\nu} \tag{945}$$

may be evaluated by direct multiplication. Note in particular the direct terms,

$$(U_{\mu;\nu} + \dot{U}_\mu U_\nu)(U^{\nu;\mu} + \dot{U}^\nu U^\mu) = U_{\mu;\nu}U^{\nu;\mu},$$

$$\frac{\theta^2}{9}(g_{\mu\nu} + U_\mu U_\nu)(g^{\mu\nu} + U^\mu U^\nu) = \frac{\theta^2}{3},$$

and the cross terms,

$$-\frac{\theta}{3}(g_{\mu\nu} + U_\mu U_\nu)(U^{\nu;\mu} + \dot{U}^\nu U^\mu) = -\frac{\theta}{3}(U_{\mu;\nu} + \dot{U}_\mu U_\nu)(g^{\mu\nu} + U^\mu U^\nu) = -\frac{\theta^2}{3}.$$

(In arriving at these results, we have made repeated use of the constancy of $U^\mu U_\mu$.) Therefore,

$$\sigma_{\mu\nu}\sigma^{\mu\nu} - \omega_{\mu\nu}\omega^{\mu\nu} = U_{\mu;\nu}U^{\nu;\mu} - \frac{\theta^2}{3}. \tag{946}$$

Using this in (940) brings us to the RE in its most general form:

$$\dot{\theta} = -(\sigma_{\mu\nu}\sigma^{\mu\nu} - \omega_{\mu\nu}\omega^{\mu\nu}) - \frac{\theta^2}{3} + (\dot{U}^{\mu})_{;\mu} + R_{\lambda\kappa}U^{\lambda}U^{\kappa} \tag{947}$$

Next, we introduce the physics of general relativity by relating the Ricci tensor $R_{\lambda\kappa}$ to components of a local stress-energy source tensor. The Einstein field equation is, in its source function form,

$$R_{\lambda\kappa} \equiv -8\pi G(T_{\lambda\kappa} - g_{\lambda\kappa}T/2) = -8\pi G S_{\lambda\kappa}. \tag{948}$$

For an ideal fluid, the stress-energy tensor is

$$T_{\lambda\kappa} = Pg_{\lambda\kappa} + (P+\rho)U_{\lambda}U_{\kappa}, \qquad S_{\lambda\kappa} = (\rho - P)g_{\lambda\kappa}/2 + (P+\rho)U_{\lambda}U_{\kappa}, \tag{949}$$

and

$$S_{\lambda\kappa}U^{\lambda}U^{\kappa} = (\rho + 3P)/2, \tag{950}$$

thus the RE becomes:

$$\dot{\theta} = -(\sigma_{\mu\nu}\sigma^{\mu\nu} - \omega_{\mu\nu}\omega^{\mu\nu}) - \frac{\theta^2}{3} - 4\pi G(\rho + 3P) + (\dot{U}^{\mu})_{;\mu}. \tag{951}$$

For an ideal fluid, the term $S_{\lambda\kappa}U^{\lambda}U^{\kappa}$ is generally positive definite. More generally, fluids for which

$$S_{\lambda\kappa}t^{\lambda}t^{\kappa} > 0, \tag{952}$$

where t^{κ} is any timelike vector, are said to satisfy the *Strong Energy Condition.* (See Carroll Ch. [4] for a discussion of other energy conditions used in general relativity.) It is possible, in principle, to violate such a constraint: a purely vacuum energy density, for example, satisfies $\rho + 3P = -2\rho$. The collapse of the vacuum not being the focus of the present investigation, we shall assume that the material in question satisfies the Strong Energy Criterion, and thus $S_{\lambda\kappa}U^{\lambda}U^{\kappa} > 0$.

To understand the crucial role of the RE in the singularity theorems of general relativity, consider a region of neighbouring geodesics, known as a *congruence,* which by definition satisfy $\dot{U}^{\mu} = 0$. The congruence is constructed so that the U^{μ} are orthogonal to well-defined hypersurfaces in the spacetime manifold, which means in essence that U^{μ} is a gradient flow. Under these conditions $\omega_{\mu\nu}$ vanishes, for the same reason that the curl of a gradient vanishes. There is no vorticity associated with gradient flow. Using this hypersurface-orthogonal congruence to probe the spacetime, the RE becomes:

$$\dot{\theta} = -\sigma_{\mu\nu}\sigma^{\mu\nu} - \frac{\theta^2}{3} - 8\pi G S_{\lambda\kappa}U^{\lambda}U^{\kappa}. \tag{953}$$

Since $\sigma^{\mu\nu}$ is a symmetric, spacelike tensor, $\sigma_{\mu\nu}\sigma^{\mu\nu}$ must be nonnegative (it may be zero). (It is the sum of the squares of the spatial diagonal elements in the appropriate rest frame, and it is frame-independent.) Hence,

$$\dot{\theta} < -\frac{\theta^2}{3} \tag{954}$$

This inequality may be integrated:

$$\theta < \frac{\theta_0}{1 + \tau\theta_0/3} \tag{955}$$

where θ_0 is the integration constant corresponding to θ at $\tau = 0$. If this is negative, the geodesics are always converging, and θ becomes infinite in a finite amount of time! In other words, a caustic forms.

This is the intuitive heart of the singularity theorems, though there is still a technical gap to fill from the formation of caustics in a congruence to an actual singularity in the spacetime manifold. There is also a RE for *null* geodesics, which is very similar in form to (947), but with only two spatial dimensions entering the calculation. Then, the factor of 1/3 is replaced by 1/2 with no other formal changes. (The null RE also requires a somewhat weaker energy constraint for convergence of geodesics.) Null geodesic caustics once again form according to (955), but with the factor of 3 replaced by 2. Setting up an initial condition of converging geodesics, Penrose introduced the notion of a *trapped surface,* from which all all future-directed null geodesics point inwards[29]. This occurs, for example, in the spherical Oppenheimer-Snyder collapse problem within those regions of spacetime for which the cloud radius is less than $2GM/c^2$. The work of Penrose, Hawking and others showed that once such a trapped surface forms, the RE combined with more technical global arguments lead to the conclusion that such a spacetime must be incomplete. In this sense, singularities are inevitable. With this establishment of this powerful result in the 1960's, black holes permanently entered the domains of observational astronomy and fundamental physics alike.

11.5 Effects of electric charge

11.5.1 The Reissner-Nordström metric

The Schwarzschild metric is not the only spherically symmetric black hole solution to the field equations. If an electric charge q is swallowed by the hole, its effect is not lost from the universe. There is the direct electric field which is retained, but more subtly, the charge alters the form of the spacetime metric. An electromagnetic stress tensor $T_{\mu\nu}$ from the charge is everywhere present, permeating the spacetime and determining, with the hole's mass, the external geometry. The presence of the stress tensor means that the metric is not a *vacuum* solution to the Einstein field equations. Rather, the metric is a solution to the full, inhomogeneous form of the field equations, and is known as the Reissner-Nordström (RN) metric. It was discovered in the very early days of general relativity (Reissner 1916, Nordström 1918), not long after Schwarzschild's great contributions. In practise, such solutions are not relevant for observational astronomy, since a charged black hole would rapidly neutralise in any actual astrophysical environment. The metric, however, evinces structure that is physically interesting in its own right.

The mathematical calculation for the RN metric is a rather straightforward extension of the Schwarzschild case. The general form of the metric tensor is the same as equation (309):

$$-c^2 d\tau^2 = -B(r)c^2 dt^2 + A(r)dr^2 + r^2(\sin^2\theta\, d\theta^2 + d\phi^2) \tag{956}$$

where B and A are functions to be determined from the field equations. The electromagnetic field tensor $F_{\mu\nu}$ is defined in terms of of vector potential $\mathcal{A}_\mu$ (not be confused with $A(r)$) by

$$F_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu. \tag{957}$$

[29] A trapped surface does not require exceptionally high densities if the volume is large. A gas with the density of Earth's sea level atmosphere spread throughout a sphere the size of the outer solar system would form a trapped surface at its boundary.

(See §2.5.2, and recall that the covariant curl may calculated with standard partial derivatives.) The black hole of interest has a mass M and a static charge q. Absent a magnetic vector potential, there is only one independent component of the field tensor:

$$F_{rt} = \partial_r \mathcal{A}_t = -\partial_r \mathcal{A}^t = E_r = -F_{tr} = -F^{rt} = F^{tr}, \tag{958}$$

(Be careful with signs.) We require the divergence of the electric field to vanish outside the hole. Self-consistently, this must be done with the RN metric itself. Using the covariant form of equation (85),

$$\frac{1}{\sqrt{g}} \frac{\partial(\sqrt{g} F^{\mu\nu})}{\partial x^\nu} = 0 = \frac{1}{\sqrt{g}} \frac{\partial(\sqrt{g} F^{tr})}{\partial r} = 0, \tag{959}$$

where g is, as usual, the absolute value of the determinant of $g_{\mu\nu}$, or $r^4 AB$. We find

$$F^{tr} = \frac{q}{r^2 \sqrt{AB}}, \tag{960}$$

with A and B to be determined.

Our next task is to find a suitable form for the electromagnetic stress tensor $T^{\mu\nu}_{\text{em}}$. This may be obtained from equation (99), replacing $\eta^{\alpha\beta}$ with $g^{\mu\nu}$:

$$T^{\mu\nu}_{\text{em}} = \frac{1}{4\pi} \left(F^\mu_{\ \gamma} F^{\nu\gamma} - \frac{1}{4} g^{\mu\nu} (F_{\delta\gamma} F^{\delta\gamma}) \right) \tag{961}$$

For ease of index manipulation, we henceforth drop the "em" subscript. Then,

$$4\pi T^{tt} = g_{rr} (F^{tr})^2 - \frac{1}{2} g^{tt} g_{tt} g_{rr} (F^{tr})^2 = \frac{A}{2} (F^{tr})^2 \tag{962}$$

Lowering indices and using (960), we find

$$T_{tt} = \frac{q^2 B}{8\pi r^4}. \tag{963}$$

A very similar calculation yields

$$T_{rr} = -\frac{q^2 A}{8\pi r^4}. \tag{964}$$

These are the only two components required to find A and B, as we know from the Schwarzschild calculation.

Because the trace $T^\mu_{\ \mu}$ vanishes, the feld equations take the form

$$R_{\mu\nu} = -\frac{8\pi G}{c^4} T_{\mu\nu} \tag{965}$$

The Ricci tensor components R_{tt} and R_{rr} are given by equations by (316) and (317). Note that even with the sources (963) and (964) present,

$$\frac{R_{rr}}{A} + \frac{R_{tt}}{B} = 0, \tag{966}$$

just as for a Schwarzschild black hole. This leads once again to $AB = 1$, and equation (960) then shows that the electric field outside the hole is unchanged from its Minkowski form in these quasi-spherical coordinates.

The tt field equation may be written:

$$-AR_{tt} = \frac{B''}{2} + \frac{B'}{r} = \frac{Gq^2}{c^4 r^4}. \tag{967}$$

The solution of this simple differential equation consists of the solution to the homogeneous equation (i.e., with zero on the right side), plus a particular solution of the inhomogeneous equation. The homogeneous solution is $1 - R_S/r$, where R_S is the Schwarzschild radius $2GM/c^2$, since this behaviour must be recovered in the limit $q \to 0$. The particular solution is easily found to be $Gq^2/(c^4 r^2)$. Hence

$$B(r) = 1 - \frac{2GM}{c^2 r} + \frac{Gq^2}{c^4 r^2}. \tag{968}$$

Therefore, the Reissner-Nordström metric for a charged black hole is

$$-c^2 d\tau^2 = -\left(1 - \frac{2GM}{c^2 r} + \frac{Gq^2}{c^4 r^2}\right) c^2 dt^2 + \frac{dr^2}{1 - 2GM/c^2 r + Gq^2/c^4 r^2} + r^2(\sin^2\theta\, d\theta^2 + d\phi^2) \tag{969}$$

(To return to SI units, simply replace the charge q with $q/\sqrt{4\pi\epsilon_0}$.)

11.5.2 Singular metric structure: worm holes and white holes

In common with Schwarzschild geometry, the RN metric exhibits a coordinate singularity when

$$B(r) = \frac{1}{A} = 1 - \frac{2GM}{c^2 r} + \frac{Gq^2}{c^4 r^2} = 0. \tag{970}$$

It is convenient to introduce two characteristic radii: the gravitational radius $r_g = GM/c^2 = R_S/2$, and the charge radius $r_q = qG^{1/2}/c^2$. Then, there are singularities in the RN metric at the locations

$$r_\pm = r_g \pm \sqrt{r_g^2 - r_q^2}. \tag{971}$$

There are now potentially two event horizons!

When $r_q \ll r_g$ and the hole is only slightly charged, then r_+ reduces to the familiar Schwarzschild radius $R_S = 2GM/c^2$, and $r_- = r_q^2/R_S$. Crossing from $r > r_+$ to $r < r_+$ is no different from the Schwarzschild experience. The role of the space r and time t coordinates are reversed. A test mass must move inexorably in the direction of future time, which means continuing smoothly toward smaller r. The mass can pause at fixed t, but not at fixed r.

However, when the inward transition into the domain $r < r_-$ occurs, a very different geometry materialises. The metric reverts back to "normal form," with r the space variable and t the time variable. The $r = 0$ singularity is once again a place, not a future time. It can be avoided. One could make a life for oneself in $r < r_-$ land. Should the urge to travel proves irresistable, however, there is nothing to stop a return journey into $r > r_-$ land. But be careful of the change in time zone! Now r is the time variable once again, and future time is toward *increasing* r, one's unavoidable destiny. This includes an outward ejection from r_+, once that boundary is reached.

Back into a normal, asymptotically flat spacetime? How can that possibly be? There is no returning from within an event horizon, which sounds very much like what has just been described. The resolution of this apparent contradiction is that the asymptotically flat spacetime the test mass has emerged into is *not* the original spacetime domain from whence it fell into the RN black hole, but a different domain altogether. *The original RN quasi-spherical coordinates are an incomplete coverage of the manifold, just as we found to be the case in our discussion of Kruskal coordinates for the Schwarzschild geometry.* The outward recrossings of (first) the r_- boundary, and (second) the r_+ boundary occur in regions of spacetime different from those of initial ingress. The phenomenon we have just described is known as a *white hole*, a sort of inverted black hole from which matter can only escape, never enter. The bridging structure between the black and white holes is known as a *worm hole.* We are certainly at the speculative edge of physics! There is as yet no astronomical evidence that is supportive of the existence of white holes, in sharp contrast to the now overwhelming case that can be made for black holes. Moreover, as we have noted, charged black holes are not thought to be possible in practise. As a matter of principle however, is this senario one that the laws of physics permits? We currently have no basis for its prohibition.

11.5.3 Naked singularities

If $r_q > r_g$, the character of the metric changes completely. There are now *no* coordinate singularities at all in the metric, though of course the true singularity at $r = 0$ remains. No coordinate singularities means no event horizons. Thus one may approach, in principle, arbitraily close to the true singularity, and return to the comfort of asymptotically flat spacetime. This type of black hole is known as a *naked* singularity.

How large a charge does this involve? Setting $r_g = r_q$ and using SI units, we obtain for the formal critical charge q_c,

$$q_c = \sqrt{4\pi G\epsilon_0}M \tag{972}$$

which translates to about 10^{20} Coulombs for a $1M_\odot$ black hole. This is an enormous amount of charge by any standard! The electric field at the Schwarzschild radius $2GM/c^2$ (simply a fiducial reference here, not an event horizon) would be 10^{23} V/m. By way of comparison, the physicist Julian Schwinger showed in 1951 that a static electric field would spontaneously create a current of $e^+\,e^-$ pairs out of the vacuum, once the strength surpassed 10^{18} V/m. Thus, our result is a breakdown of self-consistency as well as of the integrity of the vacuuum. Roger Penrose has conjectured that naked singularities simply cannot exist, full stop. This is not for practical reasons of construction, but because the laws of physics, when properly applied, will not allow a self-consistent solution under any circumstances. Penrose's conjecture is known as the *Cosmic Censorship Hypothesis:* naked spacetime singularities are forbidden. They must be veiled by an event horizon.

Further discussion of such matters is consigned to the problems. We need say nothing more.

11.6 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 result in 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 $\rho\boldsymbol{v}$. Hence

$$\frac{d}{dt}\int_V \rho\, dV = \int_V \frac{\partial\rho}{\partial t}\, dV = -\int_{\partial V} \rho\boldsymbol{v}\cdot\boldsymbol{dA} = -\int_V \boldsymbol{\nabla}\cdot(\rho\boldsymbol{v})\, dV$$

where ∂V is the volume's boundary, and we have used the divergence theorem. The volume V is arbitrary, so we must have

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot(\rho\boldsymbol{v}) = 0 \tag{973}$$

which is the equation of mass conservation. Newton's equation of motion states that if a mass element of fluid ρdV is accelerating, then it is acted on by a gravitational force given by $-\rho dV\boldsymbol{\nabla}\Phi$, where Φ is the associated potential function. In other words, the force equation reads after cancellation of ρdV,

$$\left[\frac{\partial}{\partial t} + (\boldsymbol{v}\cdot\boldsymbol{\nabla})\right]\boldsymbol{v} = -\boldsymbol{\nabla}\Phi \tag{974}$$

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

$$\boldsymbol{v} = \frac{\dot{R}}{R}\boldsymbol{r} \tag{975}$$

a familiar Hubble law. The mass equation then becomes

$$\frac{\partial\rho}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^3\rho\frac{\dot{R}}{R}\right) = 0 \tag{976}$$

With the background ρ independent of position,

$$\frac{d\ln\rho}{dt} + \frac{3\dot{R}}{R} = 0 \tag{977}$$

whence ρR^3 is a constant, as we know. As for the force equation, a straightforward calculation yields

$$\frac{\partial\boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{v} = \frac{\ddot{R}}{R}\boldsymbol{r} \tag{978}$$

and, as the discussion of §10.1.1 shows that $\ddot{R}/R = -4\pi G\rho/3$, our solution for the background expansion is correct. The Poisson equation

$$\nabla^2\Phi = \frac{1}{r}\frac{\partial^2(r\Phi)}{\partial r^2} = 4\pi G\rho \tag{979}$$

is likewise solved by our solution. (Try $\Phi = 2\pi G\rho r^2/3$.)

We are interested in the fate of small disturbances $\delta\rho$, $\delta\boldsymbol{v}$ and $\delta\Phi$ expressed as small additions to this background:

$$\rho \to \rho + \delta\rho, \quad \boldsymbol{v} \to \boldsymbol{v} + \delta\boldsymbol{v}, \quad \Phi \to \Phi + \delta\Phi \tag{980}$$

Using our fundamental equations, we replace our dynamical variables as shown, and because the δ-quantities are small, we retain them only through linear order. We ignore quadratic and higher order tems. The mass equation is then

$$\frac{\partial\delta\rho}{\partial t} + \boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{\delta v}) + \boldsymbol{\nabla}\boldsymbol{\cdot}(\boldsymbol{v}\delta\rho) = 0 \tag{981}$$

Note that the gradient of ρ vanishes and that the equilibrium $\boldsymbol{v}$ satisfies $\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{v} = -\partial_t \ln\rho$. It is then straightforward to show that the perturbed linearised mass conservation equation simplifies to

$$\left[\frac{\partial}{\partial t} + (\boldsymbol{v}\cdot\boldsymbol{\nabla})\right]\frac{\delta\rho}{\rho} + \boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{\delta v} = 0 \tag{982}$$

The linearised equation of motion

$$\left[\frac{\partial}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\right]\boldsymbol{\delta v} + (\boldsymbol{\delta v}\cdot\boldsymbol{\nabla})\boldsymbol{v} = -\boldsymbol{\nabla}\delta\Phi \tag{983}$$

becomes

$$\left[\frac{\partial}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\right]\boldsymbol{\delta v} + \frac{\dot{R}}{R}\boldsymbol{\delta v} = -\boldsymbol{\nabla}\delta\Phi \tag{984}$$

Next, we change to comoving coordinates. This is straightforward. Let $\boldsymbol{r} = R(t')\boldsymbol{r'}$ and $t = t'$. Then $\boldsymbol{r'}$ (or x'_i in index notation) is a comoving spatial coordinate. The partial derivative transformation is (sum over repeated i):

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t'} + \frac{\partial x'_i}{\partial t}\frac{\partial}{\partial x'_i}, \qquad \boldsymbol{\nabla} = \frac{1}{R}\boldsymbol{\nabla}', \qquad \frac{\partial x'_i}{\partial t} = -\frac{x_i}{R^2}\dot{R} = -\frac{v_i}{R}, \tag{985}$$

so that

$$\frac{\partial}{\partial t} + (\boldsymbol{v}\cdot\boldsymbol{\nabla}) = \frac{\partial}{\partial t'}, \tag{986}$$

which is a time derivative following a fluid element of the unperturbed expansion. Then, our two equations for mass conservation and dynamics are

$$\frac{\partial}{\partial t'}\frac{\delta\rho}{\rho} + \frac{1}{R}\boldsymbol{\nabla}'\boldsymbol{\cdot}\boldsymbol{\delta v} = 0 \tag{987}$$

$$\frac{\partial\boldsymbol{\delta v}}{\partial t'} + \frac{\dot{R}}{R}\boldsymbol{\delta v} = -\frac{1}{R}\boldsymbol{\nabla}'\delta\Phi \tag{988}$$

Taking $\boldsymbol{\nabla}'\boldsymbol{\cdot}$ of (988) and using (987) leads to

$$\frac{\partial^2}{\partial t'^2}\frac{\delta\rho}{\rho} + \frac{2\dot{R}}{R}\frac{\partial}{\partial t'}\frac{\delta\rho}{\rho} = \frac{1}{R^2}\nabla'^2\delta\Phi = 4\pi G\rho\left(\frac{\delta\rho}{\rho}\right) \tag{989}$$

where the last equality is the linearised Poisson equation $\nabla^2\delta\phi = 4\pi G\delta\rho$. For an Einstein-de Sitter univese, 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:

$$\ddot{\delta} + \frac{4}{3t}\dot{\delta} - \frac{2}{3t^2}\delta = 0 \tag{990}$$

This differential equation has two very simple linearly independent solutions, one where δ 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 pertubations 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[30]. The questions of where those might come from are the topics of the next section.

11.7 Inflationary Models

11.7.1 "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 (177). 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

$$1 - \frac{8\pi G\rho}{H^2} \equiv 1 - \Omega = \frac{2E}{\dot{R}^2} \tag{991}$$

Now if E happens to be zero, Ω is unity for all time. Fair enough. But the measured value of Ω, 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 Ω is very close to unity if E is small and $\dot{R}$ large, or Ω is very small indeed, proportional to $1/R\dot{R}^2$ throughout most of the vast history of a matter dominated universe. But Ω passes though 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. Physicists are uneasy with coincidences like that.

To see how the concept of inflation can resolve both of these problems, consider the integral that is done on page (177) 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

[30] One also needs a great deal of Dark Matter, a topic we must alas leave untouched.

proportional to

$$\int_0^t \frac{dt'}{R(t')} = \int_0^R \frac{dR}{R\dot{R}} \tag{992}$$

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 $\dot{R} \geq 0$ 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 ρ_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[31]. The rapid large expansion also has the effect of killing off the $2E/\dot{R}$ term in equation (991). 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, (991) shows us that Ω must then be equal to unity to great accuracy. This is just what observations show.

11.7.2 The stress energy tensor of a scalar field

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 centre-stage appearance in a leading role.

We start simply, 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

$$\Box\Phi - \mu^2\Phi = 0, \quad \mu^2 = (mc/\hbar)^2 \tag{993}$$

where as usual $\Box = \partial^\alpha\partial_\alpha$, m is the mass of the particle (the "quantum of the field"), and $\hbar$ is Planck's constant over 2π. To solve the Einstein Field Equations, an explicit form for the stress energy tensor $T^{\mu\nu}$ is of course needed. For ordinary matter, this is not a problem:

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

The question is, what is $T^{\mu\nu}$ for a *quantum field* obeying equation (993)? Is there something we can identify as the energy density? The pressure? Once we have a $T^{\mu\nu}$ from a field equation, it should be treated on an equal footing with the standard $T^{\mu\nu}$ given above, and used as a 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

[31] 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.

for whatever legitimate use we would like. We may certainly use the radiation stress tensor of §2.6 on the right side of the Einstein Field Equations. To remind the reader, we define the field tensor $F_{\alpha\beta}$ in terms of the 4-potential A^μ:

$$F_{\alpha\beta} = \frac{\partial A_\beta}{\partial x^\alpha} - \frac{\partial A_\alpha}{\partial x^\beta}. \tag{994}$$

The stress tensor is then

$$T^{\alpha\beta} = \frac{1}{4\pi}\left(F^\alpha_{\ \gamma}F^{\beta\gamma} - \frac{1}{4}\eta^{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta}\right). \tag{995}$$

(Jackson [1998] or W72 are good references if further details are needed. Beware of different sign conventions for the metric tensor and normalisation factors in different texts.) This is the only conserved tensorial combination that is quadratic in the derivatives of the 4-potential, and in principle one may derive the stress tensor directly this way. The overall normalisation constant of $T^{\alpha\beta}$ can then be determined by identifying the interaction term between the fields and the particles via the Lorentz equation of motion. (More precisely, by identifying the rate at which work is done by the fields upon the particle sources.) A true tensor that is quadratic in the derivatives of the potentials and is conserved: these are the essential features we seek in our stress tensor.

With this thought in mind, it is a rather simple matter to find the $T_{\alpha\beta}$ for the Klein-Gordon field, and even to pick out its corresponding ρ and P. Multiply (993) by $\partial_\beta\Phi$. For future convenience, we write the $\mu^2\Phi$ term in terms of the derivative of a potential function $V(\Phi)$, and set $dV(\Phi)/d\Phi = \mu^2\Phi$, the potential derivative. For the K-G equation, $V = \mu^2\Phi^2/2$. We will shortly consider other forms, as well. Our field equation becomes:

$$\frac{\partial\Phi}{\partial x^\beta}\frac{\partial^2\Phi}{\partial x^\alpha \partial x_\alpha} = \frac{dV(\Phi)}{d\Phi}\frac{\partial\Phi}{\partial x^\beta} = \frac{\partial V}{\partial x^\beta} = \frac{\partial(\eta_{\alpha\beta}V)}{\partial x_\alpha}. \tag{996}$$

Integrate the left side by parts:

$$\frac{\partial\Phi}{\partial x^\beta}\frac{\partial^2\Phi}{\partial x^\alpha \partial x_\alpha} = \frac{\partial}{\partial x_\alpha}\left(\frac{\partial\Phi}{\partial x^\alpha}\frac{\partial\Phi}{\partial x^\beta}\right) - \frac{\partial\Phi}{\partial x^\alpha}\frac{\partial^2\Phi}{\partial x^\beta \partial x_\alpha}. \tag{997}$$

The final term of (997) is

$$-\frac{\partial\Phi}{\partial x^\alpha}\frac{\partial^2\Phi}{\partial x^\beta \partial x_\alpha} = -\frac{1}{2}\frac{\partial}{\partial x_\alpha}\left(\eta_{\alpha\beta}\frac{\partial\Phi}{\partial x^\gamma}\frac{\partial\Phi}{\partial x_\gamma}\right). \tag{998}$$

Putting the last three equations together:

$$\frac{\partial T_{\alpha\beta}}{\partial x_\alpha} = 0, \quad \text{where } T_{\alpha\beta} = -\eta_{\alpha\beta}\left[\frac{1}{2}\frac{\partial\Phi}{\partial x^\gamma}\frac{\partial\Phi}{\partial x_\gamma} + V(\Phi)\right] + \left(\frac{\partial\Phi}{\partial x^\alpha}\frac{\partial\Phi}{\partial x^\beta}\right), \tag{999}$$

with frame-independent trace $T = -(\partial_\alpha\Phi)(\partial^\alpha\Phi) - 4V$. T_{00} may now be read off directly:

$$T_{00} = \text{“}\rho c^2\text{”} = \frac{1}{2c^2}\dot{\Phi}^2 + \frac{1}{2}|\boldsymbol{\nabla}\Phi|^2 + V, \tag{1000}$$

which really does look like an energy density, but corresponds to the true rest frame ρc^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 Φ, and the final term is the *external* potential from an external driver. In textbooks on quantum fields, Chapter 1 often begins by a discussion of 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 (1000), where Φ corresponds to the mass displacement.

To extract a field rest frame energy density and pressure, begin by noting that the coefficient of $\eta_{\alpha\beta}$ is by definition the rest frame pressure associated with the Φ field:

$$P = -\left(\frac{1}{2}\partial^\gamma\Phi\partial_\gamma\Phi + V(\Phi)\right) \tag{1001}$$

To obtain the rest frame energy density ρ, note that the scalar trace $T^\alpha_{\ \alpha}$ of the stress tensor for an ideal fluid is $3P - \rho c^2$, which we identify with the trace of the field stress tensor (999):

$$T^\alpha_{\ \alpha} = 3P - \rho c^2 = -\frac{\partial\Phi}{\partial x^\alpha}\frac{\partial\Phi}{\partial x_\alpha} - 4V(\Phi) \tag{1002}$$

With P taken from (1001), it is now a simple matter to solve for ρc^2 and $\rho + P/c^2$:

$$\rho c^2 = -\left(\frac{1}{2}\partial^\alpha\Phi\partial_\alpha\Phi - V(\Phi)\right), \quad \rho + \frac{P}{c^2} = -\frac{1}{c^2}\,\partial^\alpha\Phi\partial_\alpha\Phi \tag{1003}$$

The latter equality then allows us to directly read the 4-velocity tensor $U_\alpha U_\beta$ of the field from the final term in (999):

$$U_\alpha U_\beta = -c^2\frac{\partial_\alpha\Phi\,\partial_\beta\Phi}{\partial^\gamma\Phi\,\partial_\gamma\Phi} \tag{1004}$$

Does this make sense in terms of field momenta divided by a mass-like quantity?

In thinking about these interesting results, remember that the expressions for P and ρ are the rest frame pressure and the rest frame energy density expressed in a form that is valid for any frame, similar to calculating the rest energy (squared) as $E^2 - p^2c^2$ using E and p from any inertial frame. Notice, for example, the difference between T_{00} in (1000) and ρc^2 in (1003). 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 proceed semi-classically: use this quantum field form of $T_{\mu\nu}$ in the classical spacetime cosmological equations during the earliest phase of the Universe. But why did we go through the trouble of explicitly identifying ρ and P? Why not just work directly with the stress tensor itself? The reason for explicitly identifying ρ and P is that it becomes a relatively easy matter to analyse the "quantum cosmological field equations" by direct analogy to the classical case.

The fundamental dynamical equation of the Universe, when curvature may be neglected, is

$$H^2 = \frac{8\pi G\rho}{3}, \tag{1005}$$

where as usual $H = \dot{R}/R$. We also need a second equation to know how ρ and P depend upon R. This is the energy conservation equation, (761):

$$\dot{\rho} + 3H\left(\rho + \frac{P}{c^2}\right) = 0. \tag{1006}$$

This equation comes from $\partial^\mu T_{\mu\nu} = 0$. This, however, amounts to solving the Klein-Gordon equation itself, since the way we formed our stress tensor (999) was precisely by contracting the K-G equation with a 4-gradient of Φ. So all the relevant equations are embodied in (1005) and (1006), with (1003) for Φ. We assume that Φ depends only upon time, so that in common with classical homogeneous cosmologies, the 4-velocity has only a U^0 component.

What is this $V(\Phi)$? It helps to have a concrete mechanical model. If I have a one-dimensional line of masses on springs (we can even make the masses out of concrete), and ϕ_n is the displacment of mass n, the equation of motion is:

$$\ddot{\phi}_n = -k(\phi_n - \phi_{n-1}) - k(\phi_n - \phi_{n+1}), \tag{1007}$$

where k is the spring constant. When n is very large and the masses are closely spaced with a small separation Δx, one may take the limit

$$\phi_n - \phi_{n-1} \simeq \Delta x \phi'_{n-1/2},$$

so that

$$-k(\phi_n - \phi_{n-1}) - k(\phi_n - \phi_{n+1}) \to -k\Delta x(\phi'_{n-1/2} - \phi'_{n+1/2}), \tag{1008}$$

where the spatial x derivative ϕ' is formally assigned a label halfway between the integer n's. A second use of this limit brings us to

$$-k\Delta x(\phi'_{n-1/2} - \phi'_{n+1/2}) = k(\Delta x)^2 \phi''_n \tag{1009}$$

or

$$\frac{\partial^2 \phi}{\partial t^2} = c^2 \frac{\partial^2 \phi}{\partial x^2} \tag{1010}$$

where $c^2 = k(\Delta x)^2$ becomes the velocity of a propagating ϕ disturbance as k gets large and Δx 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 K-G 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 [29]. Then, our final partial differential equation becomes

$$\frac{\partial^2 \phi}{\partial t^2} = c^2 \frac{\partial^2 \phi}{\partial x^2} - \frac{g}{l}\phi \tag{1011}$$

This is the Klein-Gordon equation with our potential V term in the form of $g\phi/l$. This term arises from an external interaction involving the displacement ϕ (or the field Φ). Note the emphasis on *interaction*. In the K-G equation, the leading coefficient of the term embodying this type of interaction is called "mass." Hence, Richard Feynman's famous quotation: *"All mass is interaction."* In our example, the mass load, which is properly inertial in character (as opposed to disspative), is the price to be paid for being able to still jiggle the springs (or the field) when there is some sort of coupling present. The mass of an electron arises, at least in part, from the fact that when you try to accelerate it, you also produce disturbances in the photon field: the charge radiates. That is certainly one form of external coupling.

We could imagine exposing our spring masses to a more complicated external force. Perhaps $g\phi/l$ 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 important point here is that as long as the null displacement ($\phi = 0$) equilibrium state is is stable, any small deviation from $\phi = 0$ will be generically linear in the interaction, while $V(\phi)$ (or $V(\Phi)$)

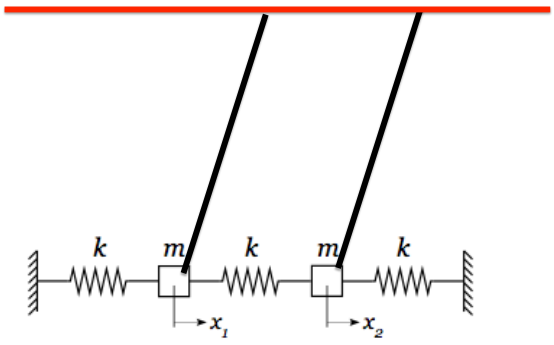


Figure 29: Mechanical analogue of the Klein-Gordon equation. A spring-like coupling between adjacent masses gives rise to the wave equation, whilst the pendula produce an acceleration directly proportional to the displacement itself, rather than the ∂_x derivative thereof. This g/l force is exactly analogous to the mass term in the KG equation, which evidently arises from similar "external couplings" of the scalar field Φ. (Only two masses are shown instead of an infinite continuum.)

is *quadratic.* This is why we may choose to think of V as a sort of potential function. If we speak casually of the field being a "ball sitting at the bottom of a potential well," the rollback oscillation frequency becomes the mass of the field particle!

The question is, in our example, what might be the actual cause of this so-called interaction? In quantum field theory, what is the external pendulum-like force that influences the Φ field? The answer is that the interaction arises from the vacuum itself. Seriously.

11.7.3 Effective Potentials in Quantum Field Theory

One thing you have to understand about the quantum vacuum: *it is a jungle out there.*

The vacuum, yes even the vacuum, is full of ground state fluctuations in the motley collection of harmonic oscillators that we are pleased to call particles. There are circumstances in which a simple K-G equation, which would represent free, massive scalar particles, may not capture the dynamics of the field in question. The K-G scalar field may interact with all these other fields, and since this interaction gives rise to what we call "mass," the interaction gives rise to a new mass coefficient. Imagine, in our mass-and-springs problem, a generic nonlinear pendulum of some sort coupling to the oscillating fields. (Let us say the length of the string depends upon the displacement.) Then, the orginal mass coefficient μ^2 is not a constant, but actually depends on the field strength Φ. In the simplest case, μ^2 would depend additively upon Φ^2, so that only the magnitude of Φ, not its sign, affects the distortion of μ^2. The resulting modified potential $V(\Phi)$ would then take the form

$$V(\Phi) = \mu^2\Phi^2/2 + \beta\Phi^4. \tag{1012}$$

This μ^2 is once again just a constant, as is β. Decades before the notion of the cosmological

inflation became popular, precisely this generic model for a modified field potential had been in wide use amongst particle physicists, for completely different reasons. The key idea is that if, despite appearances, μ^2 is for some reason *negative* in our new model, and β positive, then $\Phi = 0$ is not a stable vacuum solution to our equation: a small departure from $\Phi = 0$ does not oscillate in time, it grows exponentially! The actual stable vacuum state would have to be the global minimum, $\Phi^2 = |\mu^2/4\beta|$. The $\Phi = 0$ symmetry is *broken.* The potential looks nicely quadratic when locally perturbed about this new, stable minimum. We are then back to the K-G Equation, at least to leading order. But now the mass coefficient of our newly perturbed field has changed! This is, in essence, the formal mechanism by which the Higgs boson acquires its mass. It is all about interactions and "spontaneous symmetry breaking."

For understanding the dynamics of the early Universe, it is not the value of Φ of the new equilibrium that is important, it is the wild, exponential journey we take to reach it: down the hill of an extended, rolling potential. It is this scenario upon which cosmological inflation was orginally based. It assumes, of course, that we start off somewhere near the top of the slide!

Sound a bit woolly? Try not to be put off. This is forefront physics, and we must grope to some extent. Physicists continue to argue about inflation scenario amongst themselves. My own view borrows from Mark Twain's famous comment upon Wagnerian music: it is better than it sounds. The $V(\Phi)$'s of the early universe were likely to be complicated functions (embodying lots of interactions at high energies), but it would not have taken long before the fields sorted themselves out, landing in their respective stable local parabolic V-minima. Perhaps one of these fields may have taken a long time (in quantum terms) to get there. So to modern physicists, we denizens of the displaced vacuum, local field disturbances 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 standard theories, and then see what the most robust predictions might be. Our current theories are incomplete, we have to grope! (Einstein in 1912 was groping: what if Minkowski space was only a first approximation and could be distorted?) All within reason, of course. As long as certain ground rules like Lorentz invariance are respected, there is considerable freedom in choosing the form of an interaction potential $V(\phi)$. As noted, the type of quartic potential seen in (1012) was already well-known to particle physicists, and very much in play at the time it was culturally-appropriated by the cosmologists. The particle physicists had, for their part, already borrowed it from condensed matter physicists, who had been using it (in the form of a *thermodynamic* potential) to describe ferromagnets. The "broken symmetry" approach has been enormously fruitful in particle physics. Physicists groped, struggled, argued, predicted, and then ultimately found the Higgs boson. Just as a ferromagnet can spontaneously magnetise itself, so too can certain types of particles spontaneously acquire mass. Physically different processes often share a very similar mathematical heritage.

A few digressive words on the ferromagnet problem may be useful. In a ferromagnet, the equilibrium state minimises the Gibbs Free Energy G, a requirement for thermodynamic systems at fixed temperature and pressure. G is a function of the magnetisation M, and typically is modelled to follow the form of (1012):

$$G = \alpha M^2 + \beta M^4. \tag{1013}$$

A ferromagnet is spontaneously magnetised at low temeperatures, minimising G, with dipole moments all alligned in a minimum energy state. On the other hand, at high temperatures, G is minimised at maximum entropy, with unalligned dipoles. This is captured in equation (1013) by demanding that above some critical temperature T_c, both α and β are positive. Below T_c, α passes through zero and changes sign. It then becomes more favourable — G attains a smaller value—when M^2 is finite and equal to $-\alpha/2\beta$. The system, in other words, is spontaneously magnetised. Yet another, very similar, mathematical arrangement of interactions also produces the phenomenon of superconductivity. We have seen that in

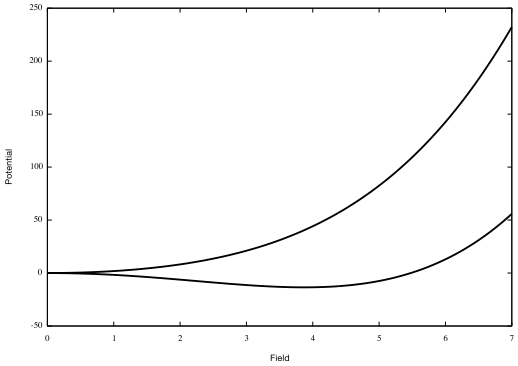


Figure 30: $V(\Phi)$ potential functions of the form (1012). The upper curve corresponds to $\mu^2 > 0$, the lower to $\mu^2 < 0$. If there is a scalar field described by the $\mu^2 < 0$ potential in the early Universe, it could trigger an episode of inflation. Before cosmological applications, these potential functions were used to describe ferromagnets, and to explain how symmetry is broken in fundamental particle physics.

particle physics this is the core of the argument for how a class of fields acquires mass. The cosmologists, a little late to the party, have at last embraced spontaneous symmetry breaking as a key concept in their quest to explain the gross features of our Universe.

11.7.4 The slow roll inflationary scenario

The essence of the slow roll inflationary model is now easily grasped. Start by using (1003) in (1005) and (1006). This leads to the equations

$$H^2 = \frac{8\pi G}{3c^2}\left(V + \frac{\dot{\Phi}^2}{2c^2}\right) \tag{1014}$$

and

$$\ddot{\Phi} + 3H\dot{\Phi} + c^2\frac{dV}{d\Phi} = 0 \tag{1015}$$

The idea is that the gross form of the governing interaction potential $V(\Phi)$ changes as the Universe expands and cools, evolving from the top form of figure [30] 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 (1015) and there is an extended period when (1014) is simply

$$H = \frac{\dot{R}}{R} \simeq \sqrt{\frac{8\pi G V_0}{3c^2}} \tag{1016}$$

where V_0 is the (approximately) constant of $V(\Phi)$. R grows exponentially,

$$R \propto \exp\left(\sqrt{\frac{8\pi G V_0}{3c^2}}t\right) \tag{1017}$$

and the Universe enters its inflationary phase. Φ 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, motivated by earlier lectures of Robert Dicke and Steven Weinberg, put together this picture in 1981 and brought to the fore the concept of inflation as a phase of the early history of the Universe. In particular, Guth explicitly argued that certain 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 (1014) and (1015), are

$$H^2 = \left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G V}{3c^2}, \quad \dot{\Phi} = -\left(\frac{c^6}{24\pi G V}\right)^{1/2} \frac{dV}{d\Phi}.$$

Under what conditions are these equations valid? Next, consider a potential of the form

$$V(\Phi) = V_0(1 - \epsilon^2\Phi^2)^2,$$

where ϵ is small and V_0 constant. Plot V as a function of Φ. 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 (yet) based on fundamental theory. They are phenomenological. The task is to carry through a detailed mathematical exploration of a candidate theory to see how naturally it leads to the behaviour in question. Perhaps the theory will explain something it wasn't specifically designed to do. That is when you know you are on the right track. Inflation models have this property, which is why there are considered attractive by many.

Here is an example. During the period of inflation, small density fluctuations go through two types of behaviour in sequence. In their very earliest development, fluctuations oscillate like a sound wave. But as the Universe rapidly expands, at some point the peak and the trough of the same wave find themselves outside of one other's horizons! A wave can't possibly oscillate coherently under those conditions. The disturbance remains "frozen," stretched by the expansion.

But the rapid inflation eventually slows while the Universe is still very much a newborn baby. Then, the expansion no longer is accelerating, but decelerating, and it does not take very long before the ever-expanding horizon scale once again is able to enclose a wavelength. In the parlance of cosmology, the wave "enters" the horizon, and the oscillation behaviour can then restart. Consider the case in which 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 precisely one full contraction of the oscillation. (Or one full expansion, if the disturbance starts on its expanding phase.) This half-period of oscillation correponds to a particular wavelength disturbance, 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 in the same time period. And so on. The so-called power spectrum[32] (figure [31]) 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 their oscillations start up at a well-defined, coordinated beginning time. *This is possible only because of*

[32] 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 $|\mathcal{T}|^2$.

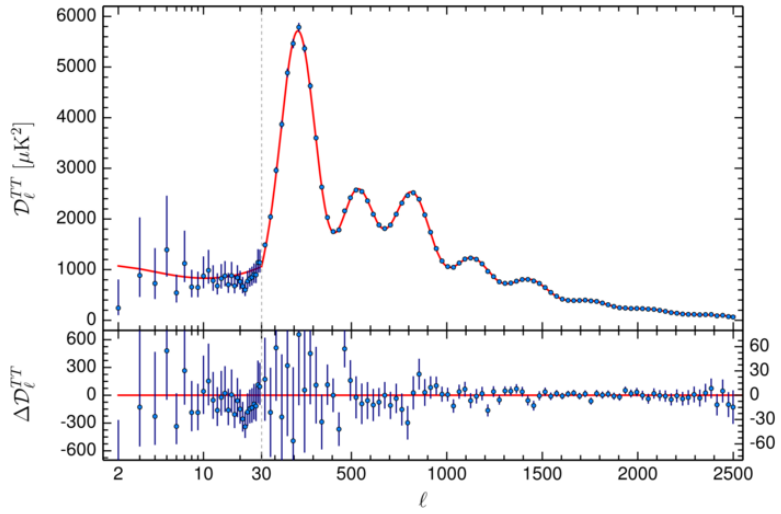


Figure 31: The temperature power spectrum of fluctuations (known as "TT spectrum" in the literature) as a function of angular scale from recent Planck data. (The bottom plot is a measure of the discrepancy between the data and a model, and does not concern us here.) The excellent model fit in the top plot is strong evidence for inflation.

inflation. During inflation, the rapid expansion prevents the random oscillations that would otherwise be present. It then allows the oscillation dance to begin again, choreographed to start from zero velocity. This moment happens very early on, once the exponential expansion has slowed and the oscillations enter the horizon. The orchestrated release at only *slightly* different (shorter wavelengths enter the horizon a tiny bit sooner!) early times is what makes well-defined peaks possible in the first place. Their very existence is powerful evidence for the concept of inflation.

11.7.5 Inflated expectations

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 predictions of inflation, including predictions of where the power spectrum should have its peaks. There is so much more that, unfortunately, 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 initial seed fluctuations. This is from highly inhomogeneous, spontaneous *quantum* fluctuations that become inflated to macroscopic scales, and grow (it is believed) into the large scale structure we see today. The power spectrum of the density fluctations that result from the fluctuating quantum inflation field can actually be predicted from theory: it is close to (but not exactly!) linear in wavenumber k. The fluctuations leave an imprint in the CMB that can be measured. What do the recent (2015) results from the Planck satellite show? A power-law power spectrum, with index a little bit less than unity

(0.96). This is remarkable agreement with theory.

To form real galaxies and real clusters, it is necessary for what is known as *Cold Dark Matter* to take part in the gravitational response to the initial seeds laid down by the quantum field. Cold Dark Matter (CDM) is cold in the sense that it responds readily to gravitational perturbations, forming in the process the bulk of the mass distribution in galaxies and clusters. These are extended dark matter haloes. Though we don't know very much about CDM (another theory dominated by phenemonology), there are reasons to think that it may be some kind of weakly interacting massive particle that pervades the Universe. There are searches underway to try to detect such particles via their more easily found decay products[33]. But it is hard to escape the notion that it all must have begun 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 it is difficult to 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? Is it inevitable? What is the underlying imperative to inflate, and then how do we exit gracefully from inflation? While there are some promising ideas afoot, we still don't really know why inflation occurred. Maybe somebody reading these notes will settle the matter.[34]

[33]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.

[34]I couldn't be more serious: given the historical track record, when this problem is solved, and it will be, it would not be at all surprising if the solver spent their formative years as an Oxford undergraduate.

11.7.6 A final word

Astrophysics can be a very messy and speculative business. But every once and a while, something truly outstanding is accomplished. We began this course by talking about the triumphs of theoretical physics. The development of a theory for 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 new ones ourselves, a stunning achievement. Another milestone is the coming of age of the theory of black holes, brilliantly confirmed in the last few years by the LIGO detections of gravitational radiation from a number of merging black hole binaries. (A real "two birds with one stone" triumph for general relativity: both black holes and gravitational radiation exist beyond a shadow of doubt.) Surely the development of precision cosmology, the discovery and construction of a rigorous model of the Universe, must also rank as one of the great historical 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 an important technical advance, but a transformative understanding. We have taken the measure of the entire Universe, and you know what? The vacuum is on board, a full paying client. What we reckoned a generation ago as the stuff of the Universe is just 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. It is not about atomic ensembles like us. It is about the vacuum. The time when matter and radiation have much of a say in telling the Universe what to do will be a fleeting instant in its history, one from which we are only just now emerging. We ensembles of ordinary matter are merely soot from the fumes of 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, we 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, something of 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.

End of notes January 16, 2022.

Bibilography

Binney, J. & Tremaine, S. 2008, *Galactic Dynamics,* 2nd ed. (Princeton Univ.: Princeton)

Birrell, N. D. & Davies, P. C. W. 1982, *Quantum Fields in Curved Space,* (Cambridge Univ.: Cambridge)

Carroll, S. M. 2019, *Spacetime and Geometry. An Introduction to General Relativity,* (Cambridge Univ.: Cambridge)

Einstein, A. 1952, *The Principle of Relativity: A Collection of Original Papers on the Special and General Theory of Relativity* (Dover: New York)

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

Hobson, M. P., Efstathiou, G. & Lasenby, A. N. 2006, *General Relativity: An Introduction for Physicists,* (Cambridge Univ.: Cambridge)

Jackson, J. D. 1999, *Classical Electrodynamics,* 3rd ed. (John Wiley & Sons: Hoboken)

Landau, L. D. & Lifshitz, E. M. 1975, *Classical Theory of Fields,* 4th English ed. (Elsevier: Oxford)

Longair, M., 2006, *The Cosmic Century,* (Cambridge Univ.: Cambridge)

Misner, C. W., Thorne, K. S., & Wheeler, J. A. 1973, *Gravitation,* (Freeman: New York)

Pais, A. 1982, *'Subtle is the Lord...'*, (Oxford Univ.: Oxford)

Peebles, P. J. E. 1993, *Principles of Physical Cosmology,* (Princeton Univ.: Princeton)

Ryden, B. 2003, *Introduction to Cosmology,* (Cambridge Univ.: Cambridge)

Schwartz, M. D. 2014, *Quantum Field Theory and the Standard Model,* (Cambridge Univ.: Cambridge)

Shapiro, S. L. & Teukolsky, S. A. 1983, *Black Holes, White Dwarfs and Neutron Stars* (Wiley–Interscience: New York)

Thorne, K. 1994, *Black Holes and Time Warps: Einstein's Outrageous Legacy,* (Norton: New York)

Tolman, R. C. 1987, *Relativity, Thermodynamics and Cosmology,* (Dover: New York)

Weinberg, S. 1972, *Gravitation and Cosmology,* (John Wiley & Sons: New York)

Weinberg, S. 2008, *Cosmology,* (John Wiley & Sons: New York) (Oxford Univ.: Oxford)

Zeldovich, Ya. B. & Novikov, I. D. 1971, *Stars and Relativity,* (Dover: Mineola)