

Gravity: Newtonian, Post-Newtonian, and General Relativistic

Clifford M. Will

Abstract We present a pedagogical introduction to gravitational theory, with the main focus on weak gravitational fields. We begin with a thorough survey of Newtonian gravitational theory. After a brief introduction to general relativity, we develop the post-Minkowskian formulation of the field equations, which is ideally suited to studying weak-field gravity. We then discuss applications of this formulation, including post-Newtonian theory, the parametrized post-Newtonian framework, and gravitational radiation.

1 Introduction

General relativity is a theory of “gravity as geometry.” It provides a remarkable new conception of the nature of gravity, totally at odds with the idea of gravity as conceived by Isaac Newton. It makes predictions of remarkable new and exotic phenomena: black holes, space-time singularities, the expanding universe.

But, as we approach the centenary of the 1915 publication of the full theory, we have learned that there is a wealth of important physical and astrophysical phenomena in which general relativity is important, but for which the full, exact theory is not needed. These phenomena range from the dynamics of the solar system, where many important experimental tests of general relativity have been carried out, to binary pulsar systems, to the dynamics of stars in galactic cores containing supermassive black holes, to the inspiral of binary systems of neutron stars or black holes, which will be important sources of detectable gravitational waves. These are situations where the gravitational potentials are small (in units of c^2) and speeds are low (in units of c). These cases lend themselves to methods that forego any attempt to find exact solutions of Einstein’s equations, but instead attempt to develop systematic approximations to solutions of the equations.

Among these methods are the “post-Minkowskian” and “post-Newtonian” approximations that will be the main topic of this chapter. In post-Minkowskian theory, the strength of the gravitational field is measured by the gravitational constant G , and the Einstein field equations are formally expanded in powers of G .

C. M. Will (✉)
University of Florida, Gainesville, USA
e-mail: cmw@physics.ufl.edu

At zeroth post-Minkowskian order there is no field, and one deals with Minkowski space-time. At first post-Minkowskian order, the gravitational field appears as a correction of order G to the Minkowski metric, and the (linearized) field equations are integrated to obtain this correction. The correction is refined by terms of order G^2 in the second post-Minkowskian approximation, and the process is continued until the desired degree of accuracy is achieved. When augmented by an assumption of slow motions, the method is known as post-Newtonian theory. A slow-motion assumption is natural, as the virial theorem requires that kinetic energies be comparable to gravitational potentials, so that weak fields, characterized by $(\text{potentials})/c^2 \ll 1$, go hand in hand with slow motions, $(v/c)^2 \ll 1$.

These methods have proven in recent years to be very powerful, and unreasonably effective in describing important relativistic phenomena. Because the methods can be carried out to higher orders in the approximation, they can handle situations where the fields are not so weak or the motions not so low. For example, high-order calculations in post-Newtonian theory of the late inspiral of a compact binary system have proven to agree remarkably well with results from numerical relativity, in regimes where the two techniques overlap.

Accordingly, this chapter is designed as a pedagogical introduction to these methods. Because they yield Newtonian gravity as the first-order approximation, we begin with a fairly lengthy exposition of Newtonian theory (Sect. 2), one that goes beyond the simplest undergraduate level of “inverse square law” – “elliptical orbit” concepts. After a brief introduction to general relativity (Sect. 3), we describe the post-Minkowskian approach (Sect. 4). This is an exact reformulation of Einstein’s equations that nicely lends itself to a systematic sequence of approximations. Including slow motions yields post-Newtonian theory, along with the parametrized post-Newtonian (PPN) formalism, used to discuss alternative theories of gravity and their experimental tests (Sect. 5). Finally, we discuss the use of these methods to analyze gravitational radiation (Sect. 6).

Because this chapter is meant to be pedagogical, we shall not provide extensive references to the primary literature, but will instead give a bibliography and suggestions for further reading at the end. This chapter is based in part on a forthcoming book, *Gravity: Newtonian, post-Newtonian, Relativistic*, coauthored by Eric Poisson. I am grateful to my coauthor and to Cambridge University Press for permission to use portions of the book draft in this chapter.

2 Newtonian Gravity

The gravitational theory of Newton is an extremely good representation of gravity for a host of situations of practical and astronomical interest. It accurately describes the structure of the Earth and the tides raised on it by the Moon and Sun. It gives a detailed account of the orbital motion of the Moon around the Earth, and of the planets around the Sun. Of course, it is now well-established that Newtonian theory is not an exact description of the laws of gravitation. But apart from specialized

situations requiring very high precision, such as the orbit of Mercury, timing in the Global Positioning System, or very long baseline radio interferometry (VLBI), Newtonian gravity rules the solar system.

For the overwhelming majority of stars in the universe, Newtonian gravity also rules. The structure and evolution of the Sun and other “main sequence” stars can be completely and accurately treated using Newtonian gravity. Only for extremely compact stellar objects, such as neutron stars and, of course, black holes, is general relativity important. Newtonian gravity is also perfectly capable of handling the structure and evolution of galaxies and clusters of galaxies, and can even be used in the context of an expanding universe, as long as the relevant distance scales are shorter than the Hubble scale.

Generally speaking, the criterion that we use to decide whether to employ Newtonian gravity or general relativity is the magnitude of the quantity

$$\varepsilon \sim \frac{GM}{c^2 r} \sim \frac{v^2}{c^2}, \quad (1)$$

where G is the Newtonian gravitational constant, c is the speed of light, and where M , r , and v represent the characteristic mass, separation or size, and velocity within the system under consideration. The smaller this factor, the better is Newtonian gravity as an approximation.

2.1 The Equations of Newtonian Gravity

Most undergraduate textbooks begin their treatment of Newtonian gravity with Newton’s second law and the inverse-square law of gravitation:

$$\begin{aligned} m_I \mathbf{a} &= \mathbf{F}, \\ \mathbf{F} &= -\frac{Gm_G M}{r^2} \mathbf{n}. \end{aligned} \quad (2)$$

In the first equation, \mathbf{F} is the force acting on a body of inertial mass m_I situated at position \mathbf{r} , and $\mathbf{a} = d^2\mathbf{r}/dt^2$ is its acceleration. In the second equation, the force is assumed to be gravitational in nature, and to originate from a gravitating mass M situated at the origin of the coordinate system. The force law involves m_G , the passive gravitational mass of the first body at \mathbf{r} , while M is the active gravitational mass of the second body. The force is attractive, it varies inversely with the square of the distance $r := |\mathbf{r}| = (x^2 + y^2 + z^2)^{1/2}$, and it points in the direction opposite to the unit vector $\mathbf{n} := \mathbf{r}/r$.

An alternative form of the force law is obtained by writing it as the gradient of a potential $U = GM/r$, so that

$$\mathbf{F} = m_G \nabla U. \quad (3)$$

If the inertial and passive gravitational masses of the body are equal to each other, $m_I = m_G$, then the acceleration of the body is given by $\mathbf{a} = \nabla U$, and its magnitude is $a = GM/r^2$. Under this circumstance, the acceleration is independent of the mass of the body. This statement is known as the *weak equivalence principle* (WEP), and it was a central element in Einstein's thinking en route to the concepts of curved space-time and general relativity. Newton himself deemed this principle so important that he opened his 1687 magnum opus *The Principia* with a discussion of it and the experiments he had performed to test it. Twentieth-century experiments have shown that the two types of mass are equal (or more precisely that the ratio m_G/m_I is independent of the composition of the material) to a few parts in 10^{13} for a wide variety of materials. The chapter, "The Newtonian Gravity and Some of Its Classical Tests (Equivalence Principle, Measure of G , $1/r^2$)" by V. Iafolla, discusses these experiments in more detail).

In order to consider the motion of extended bodies made up of continuous matter (solid, fluid, or gas), allowing the bodies to be of arbitrary size, shape, and constitution, and possibly to evolve in time according to their own internal dynamics, it is necessary to generalize the primitive Eq. (2) to a form that applies to a continuous distribution of matter.

To do this, we characterize the matter distribution by a mass-density field $\rho(t, \mathbf{x})$, a pressure field $p(t, \mathbf{x})$, and a velocity field $\mathbf{v}(t, \mathbf{x})$; these quantities depend on time t and position \mathbf{x} within the fluid. The equations that govern the behavior of the matter are the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (4)$$

which expresses conservation of mass, and *Euler's equation*,

$$\rho \frac{d\mathbf{v}}{dt} = \rho \nabla U - \nabla p, \quad (5)$$

which is the generalization of Eq. (2) to continuous matter; here

$$\frac{d}{dt} := \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla, \quad (6)$$

is the convective time derivative associated with the motion of fluid elements. The equation that governs the behavior of the gravitational field is *Poisson's equation*

$$\nabla^2 U = -4\pi G \rho, \quad (7)$$

where

$$\nabla^2 := \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}, \quad (8)$$

is the familiar Laplacian operator. Using the Green function for the Laplacian, we can convert this to an integral

$$U(t, \mathbf{x}) = G \int \frac{\rho(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3 x'. \quad (9)$$

The integral can be evaluated as soon as the density field $\rho(t, \mathbf{x}')$ is specified, regardless of whether ρ is a proper solution to the remaining fluid equations. As such, Eq. (9) gives U as a *functional* of any arbitrary function ρ . The potential, however, will be physically meaningful only when ρ itself is physically meaningful, which means that it must be a proper solution to the continuity and Euler equations.

To complete the formulation of the theory we must impose a relationship between the pressure and the density of the fluid. This relationship, known as the *equation of state*, takes the general form of

$$p = p(\rho, T, \dots), \quad (10)$$

in which the pressure is expressed as a function of the density, temperature, and possibly other relevant variables such as chemical composition. The equation of state encodes information about the microphysics that characterizes the fluid, and this information must be provided as an input in most applications of the theory.

A complete description of a physical situation involving gravity and a distribution of matter will be obtained by integrating Eqs. (4), (5), and (7) simultaneously and self-consistently. The solutions must be subjected to suitable boundary conditions, which will be part of the specification of the problem. All of Newtonian gravity is contained in these equations, and all associated phenomena follow as consequences of these equations.

The equations of hydrodynamics give rise to a number of important *global conservation laws*. These refer to global quantities, defined by integrals over the entire fluid system, that are constant in time whenever the system is *isolated*, that is, whenever the system is not affected by forces external to it or by flows of matter out of it. The equation of continuity (4) leads to some useful properties of such integrals. For an arbitrary function $f(t, \mathbf{x})$, use of the equation of continuity and an integration by parts leads to

$$\frac{d}{dt} \int \rho(t, \mathbf{x}) f(t, \mathbf{x}) d^3x = \int \rho \frac{df}{dt} d^3x, \quad (11)$$

where, as before, d/dt is the convective (or Lagrangian) time derivative and where the volume of integration is fixed in space and contains all the matter.

For the integral $F(t, \mathbf{x}) := \int \rho(t, \mathbf{x}') f(t, \mathbf{x}, \mathbf{x}') d^3x'$,

$$\frac{\partial F}{\partial t} = \int \rho' \left(\frac{\partial f}{\partial t} + \mathbf{v}' \cdot \nabla' f \right) d^3x'. \quad (12)$$

The *Lagrangian time derivative* acting on F is $dF/dt = \partial F/\partial t + \mathbf{v} \cdot \nabla F$, and from Eq. (12) and the definition of $F(t, \mathbf{x})$ we find that this can be expressed as

$$\frac{dF}{dt} = \int \rho' \frac{df}{dt} d^3x', \quad (13)$$

with

$$\frac{df}{dt} := \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \mathbf{v}' \cdot \nabla' f, \quad (14)$$

denoting a generalized Lagrangian derivative.

With these results it is straightforward, using the equation of continuity and the Euler equation, to show that the following are constant in time:

$$\text{Mass : } M := \int \rho(t, \mathbf{x}) d^3x, \quad (15a)$$

$$\text{Momentum : } \mathbf{P} := \int \rho(t, \mathbf{x}) \mathbf{v}(t, \mathbf{x}) d^3x, \quad (15b)$$

$$\text{Angular Momentum : } \mathbf{J} := \int \rho \mathbf{x} \times \mathbf{v} d^3x, \quad (15c)$$

$$\text{Energy : } E := \mathcal{T}(t) + \Omega(t) + E_{\text{int}}(t). \quad (15d)$$

In addition, the center of mass, defined by

$$\mathbf{R} := \frac{1}{M} \int \rho(t, \mathbf{x}) \mathbf{x} d^3x, \quad (16)$$

is a linear function of time. Finally, one can prove the virial theorems

$$\frac{1}{2} \frac{d^2 I^{jk}}{dt^2} = 2\mathcal{T}^{jk} + \Omega^{jk} + P\delta^{jk}, \quad (17a)$$

$$\frac{1}{2} \frac{d^2 I}{dt^2} = 2\mathcal{T} + \Omega + 3P, \quad (17b)$$

where

$$\mathcal{T}(t) := \frac{1}{2} \int \rho v^2 d^3x, \quad (18a)$$

$$\mathcal{T}^{jk}(t) := \frac{1}{2} \int \rho v^j v^k d^3x, \quad (18b)$$

$$\Omega(t) := -\frac{1}{2} \int \rho U d^3x = -\frac{1}{2} G \int \frac{\rho \rho'}{|\mathbf{x} - \mathbf{x}'|} d^3x' d^3x, \quad (18c)$$

$$\Omega^{jk}(t) := -\frac{1}{2} G \int \rho \rho' \frac{(x - x')^j (x - x')^k}{|\mathbf{x} - \mathbf{x}'|^3} d^3x' d^3x, \quad (18d)$$

$$E_{\text{int}}(t) := \int \varepsilon d^3x, \quad (18e)$$

$$P(t) := \int p d^3x, \quad (18f)$$

$$I^{jk}(t) := \int \rho x^j x^k d^3x, \quad (18g)$$

$$I(t) := \int \rho r^2 d^3x, \quad (18h)$$

where ε is the energy density, subject to the first law of thermodynamics for perfect or isentropic fluids, $d(\varepsilon \mathcal{V}) + p d\mathcal{V} = 0$, where the element of volume evolves with the fluid flow according to $\mathcal{V}^{-1} d\mathcal{V}/dt = \nabla \cdot \mathbf{v}$.

2.2 Spherical and Nonspherical Bodies

In spherical symmetry, Poisson's equation (7) takes the form

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial U}{\partial r} \right) = -4\pi G \rho(t, r). \quad (19)$$

Defining the mass contained inside a sphere of radius r as

$$m(t, r) := \int_0^r 4\pi \rho(t, r') r'^2 dr', \quad (20)$$

we rewrite Poisson's equation in the form

$$\frac{\partial U}{\partial r} = -\frac{Gm(t, r)}{r^2}. \quad (21)$$

If R is the radius at which $\rho = 0$, then, defining $M = m(R)$ (a constant by virtue of the continuity equation), we can write the solution of Eq. (21) that is continuous at $r = R$ and vanishes at infinity in the form

$$U(t, r) = \frac{Gm(t, r)}{r} + 4\pi G \int_r^R \rho(t, r') r' dr'. \quad (22)$$

Outside the body, $U = GM/r$, a constant in time.

To describe the potential of nonspherical bodies, we return to Eq. (9), and, for a field point outside the body ($|\mathbf{x}'| < |\mathbf{x}|$), we carry out a Taylor expansion of $|\mathbf{x} - \mathbf{x}'|^{-1}$:

$$\begin{aligned} \frac{1}{|\mathbf{x} - \mathbf{x}'|} &= \frac{1}{r} - x'^j \partial_j \left(\frac{1}{r} \right) + \frac{1}{2} x'^{jk} \partial_{jk} \left(\frac{1}{r} \right) - \dots \\ &= \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{\ell!} x'^L \partial_L \left(\frac{1}{r} \right), \end{aligned} \quad (23)$$

where we introduce a condensed notation in which an expression like x'^{jkn} stands for the product $x'^j x'^k x'^n$, and ∂_{jkn} stands for $\partial_j \partial_k \partial_n$. We adopt the Einstein summation convention and sum over repeated indices in an expression like $x'^{jk} \partial_{jk} r^{-1}$. In the second line, we introduce a *multi-index* notation, in which an uppercase index such as L represents a collection of ℓ individual indices. Thus, x'^L stands for $x'^{j_1 j_2 \dots j_\ell}$, ∂_L stands for $\partial_{j_1 j_2 \dots j_\ell}$, and $x'^L \partial_L$ involves a summation over all ℓ pairs of repeated indices. Substituting Eq. (23) into Eq. (9) gives

$$U_{\text{ext}}(t, \mathbf{x}) = G \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{\ell!} I^{(L)} \partial_{(L)} \left(\frac{1}{r} \right), \quad (24)$$

with

$$I^{(L)}(t) := \int \rho(t, \mathbf{x}') x'^{(L)} d^3 x', \quad (25)$$

defining a set of *STF multipole moments* for the mass distribution. The STF label denotes *symmetric, trace-free*: an object $A^{(ijkl\dots)}$ is symmetric on all its indices and, for any pair of indices, $\delta_{ij}A^{(ijkl\dots)} = 0$. The gradients of r^{-1} generate STF tensors. Using the fact that

$$\partial_j r = n_j, \quad (26a)$$

$$\partial_j n_k = \partial_k n_j = \frac{1}{r}(\delta_{jk} - n_j n_k), \quad (26b)$$

it is straightforward to show that

$$\partial_j r^{-1} = -n_j r^{-2}, \quad (27a)$$

$$\partial_{jk} r^{-1} = (3n_j n_k - \delta_{jk}) r^{-3}, \quad (27b)$$

$$\partial_{jkn} r^{-1} = -\left[15n_j n_k n_n - 3(n_j \delta_{kn} + n_k \delta_{jn} + n_n \delta_{jk})\right] r^{-4}. \quad (27c)$$

We define the STF tensors

$$n^{(jk)} := n^j n^k - \frac{1}{3} \delta^{jk}, \quad (28a)$$

$$n^{(jkn)} := n^j n^k n^n - \frac{1}{5}(\delta^{jk} n^n + \delta^{jn} n^k + \delta^{kn} n^j), \quad (28b)$$

and more generally,

$$n^{(L)} := \sum_{p=0}^{[\ell/2]} (-1)^p \frac{(2\ell - 2p - 1)!!}{(2\ell - 1)!!} \left[\delta^{2p} n^{L-2p} + \text{sym}(q) \right], \quad (29)$$

where δ^{2p} stands for a product of p Kronecker deltas (with indices running from j_1 to j_{2p}), n^{L-2p} stands for a product of $\ell - 2p$ unit vectors (with indices running from j_{2p+1} to j_ℓ), and “sym(q)” denotes all distinct terms arising from permuting indices; the total number of terms within the square brackets is equal to $q := \ell! / [(\ell - 2p)!(2p)!!]$. Then, we have

$$\partial_L r^{-1} = \partial_{(L)} r^{-1} = (-1)^\ell (2\ell - 1)!! \frac{n^{(L)}}{r^{\ell+1}}. \quad (30)$$

The fact that $\partial_L r^{-1}$ is trace-free follows trivially from the fact that $\nabla^2 r^{-1} = 0$.

There is a close connection between $n^{(L)}$ and the spherical harmonics. For a chosen unit vector \mathbf{e} ,

$$e_{(L)} n^{(L)} = \frac{\ell!}{(2\ell - 1)!!} P_\ell(\cos \theta), \quad (31)$$

where $\cos \theta = \mathbf{e} \cdot \mathbf{n}$. In addition,

$$n^{(L)} := N_\ell \sum_{m=-\ell}^{\ell} \mathcal{Y}_{\ell m}^{(L)} Y_{\ell m}(\theta, \phi), \quad N_\ell = \frac{4\pi \ell!}{(2\ell + 1)!!}, \quad (32)$$

where $\mathcal{Y}_{\ell m}^{(L)}$ is a set of constant, complex STF tensors that satisfy $\mathcal{Y}_{\ell, -m}^{(L)} = (-1)^m \mathcal{Y}_{\ell m}^{*(L)}$. For $\ell = 1$, for example, Eq. (32) embodies the well-known relation between $(x \pm iy)/r$ and z/r and the $\ell=1$ spherical harmonics.

There will be occasions when we need to calculate the average of a quantity $\psi(\theta, \phi)$ over the surface of a sphere:

$$\langle\langle\psi\rangle\rangle := \frac{1}{4\pi} \int \psi(\theta, \phi) d\Omega. \quad (33)$$

Of particular interest are the spherical average of products $n^j n^k n^n \dots$ of radial vectors. These are easily computed using the fact that the average of the STF tensor $n^{(jkn\dots)}$ must be zero; this property follows directly from Eq. (32) and the identity $\int Y_{\ell m}(\theta, \phi) d\Omega = 0$ (unless $\ell = 0$). In this way we obtain, for example,

$$\langle\langle n^j n^k \rangle\rangle = \frac{1}{3} \delta^{jk}, \quad (34a)$$

$$\langle\langle n^j n^k n^n n^p \rangle\rangle = \frac{1}{15} (\delta^{jk} \delta^{np} + \delta^{jn} \delta^{kp} + \delta^{jp} \delta^{kn}), \quad (34b)$$

and so on. The general expression for such angular averages can be shown to be given by

$$\langle\langle n^L \rangle\rangle = \frac{1}{(\ell + 1)!!} [\delta^L + \text{sym}(q)], \quad (35)$$

when ℓ is an even number, and $\langle\langle n^L \rangle\rangle = 0$ when ℓ is odd; we use the same notation as in Eq. (29), in which δ^L stands for a product of $\ell/2$ Kronecker deltas, and $\text{sym}(q)$ denotes all distinct terms obtained by permuting indices; the total number of terms within the square brackets is equal to $q = (\ell - 1)!!$.

2.3 Motion of Extended Fluid Bodies

We consider a fluid system that is broken up into a number N of separated and isolated bodies, with the size of each body R assumed to be small compared to the interbody distance r . Each body is assigned a label $A = 1, 2, \dots, N$, and each body occupies a volume V_A bounded by a closed surface S_A . The mass density ρ is assumed to be equal to ρ_A inside V_A , and zero in the vacuum region between bodies. The fluid dynamics inside each body is governed by the continuity and Euler equations (4) and (5) and the gravitational potential U is given everywhere by Eq. (9).

We define the total mass of body A along with its center-of-mass position, velocity, and acceleration by

$$m_A := \int_A \rho(t, \mathbf{x}) d^3x, \quad (36a)$$

$$\mathbf{r}_A(t) := \frac{1}{m_A} \int_A \rho(t, \mathbf{x}) \mathbf{x} d^3x, \quad (36b)$$

$$\mathbf{v}_A(t) := \frac{1}{m_A} \int_A \rho(t, \mathbf{x}) \mathbf{v} d^3x, \quad (36c)$$

$$\mathbf{a}_A(t) := \frac{1}{m_A} \int_A \rho(t, \mathbf{x}) \frac{d\mathbf{v}}{dt} d^3x, \quad (36d)$$

where the domain of integration is a region of space that extends slightly beyond the volume V_A ; it is sufficiently small that it contains no other body, but sufficiently large that it continues to contain body A as it moves about in a small interval of time dt . It is easy to show that m_A is time-independent: $dm_A/dt = 0$ and that

$$\mathbf{v}_A = \frac{d\mathbf{r}_A}{dt}, \quad \mathbf{a}_A = \frac{d\mathbf{v}_A}{dt}. \quad (37)$$

In addition to these variables, we introduce

$$I_A^{(L)}(t) := \int_A \rho(t, \mathbf{x}) (x - r_A)^{(L)} d^3x, \quad (38)$$

the STF multipole moments of body A , which refer to its center-of-mass position $\mathbf{r}_A(t)$; notice that the dipole moment $I_A^j = \int_A \rho(x - r_A)^j d^3x$ vanishes automatically.

Integrating the Euler equation (5) over body A , and noting that $\int_A \nabla p d^3x = 0$ and $\int_A \rho \nabla U_A d^3x = 0$, we obtain

$$m_A \mathbf{a}_A = \sum_{B \neq A} \int_A \rho(t, \mathbf{x}) \int_B \frac{\rho(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x d^3x'. \quad (39)$$

Defining

$$\begin{aligned} \mathbf{x} &= \mathbf{r}_A + \bar{\mathbf{x}}, \\ \mathbf{x}' &= \mathbf{r}_B + \bar{\mathbf{x}}', \\ \mathbf{r}_{AB} &= \mathbf{r}_A - \mathbf{r}_B, \end{aligned} \quad (40)$$

and carrying out a Taylor expansion in both $\bar{\mathbf{x}}$ and $\bar{\mathbf{x}}'$, assuming both are small compared to r_{AB} , we obtain, after some rearrangement of terms,

$$\begin{aligned} a_A^j &= G \sum_{B \neq A} \left\{ -\frac{m_B n_{AB}^j}{r_{AB}^2} \right. \\ &\quad + \sum_{\ell=2}^{\infty} \frac{1}{\ell!} \left[(-1)^\ell I_B^{(L)} + \frac{m_B}{m_A} I_A^{(L)} \right] \partial_{jL}^A \left(\frac{1}{r_{AB}} \right) \\ &\quad \left. + \frac{1}{m_A} \sum_{\ell=2}^{\infty} \sum_{\ell'=2}^{\infty} \frac{(-1)^{\ell'}}{\ell! \ell'!} I_A^{(L)} I_B^{(L')} \partial_{jLL'}^A \left(\frac{1}{r_{AB}} \right) \right\}, \end{aligned} \quad (41)$$

where $\mathbf{n}_{AB} := \mathbf{r}_{AB}/r_{AB}$ is a unit vector that points from body B to body A . This expression implies that $\sum_A m_A \mathbf{a}_A = \mathbf{0}$, a statement that reflects Newton's third law or the conservation of total momentum.

The multipole moments of a perfectly spherical body vanish, $I_A^{(L)} = 0$ for $\ell \neq 0$, and when all the bodies are spherical, Eq. (41) reduces to the familiar set of “point-mass” equations of motion,

$$a_A^j = - \sum_{B \neq A} \frac{G m_B n_{AB}^j}{r_{AB}^2}. \quad (42)$$

When the bodies are not spherical, the motion of body A is affected by the distortion of the gravitational potential caused by the deformation of the other bodies; this influence is described by the terms in Eq. (41) that involve $I_B^{(L)}$. It is affected also by the nonspherical coupling of its own mass distribution to the monopole field of each external body; this influence is described by the terms in Eq. (41) that are linear in $I_A^{(L)}$. And finally, it is affected by the coupling between its own multipole moments and those of the remaining bodies, as described by the last line in Eq. (41). This last effect is analogous to the dipole–dipole coupling in electrodynamics, except for the fact that there is no dipole moment in gravitation; the leading effect comes from a quadrupole–quadrupole interaction. The presence of terms involving $I_A^{(L)}$ in the equations of motion imply that the motion of a body can depend on its internal structure, by virtue of its finite size and the nonspherical coupling of its mass distribution to the external gravitational field. This observation does not constitute a violation of the WEP; such a violation would imply a dependence on internal structure that would be present even when the bodies have a negligible size.

Because each multipole moment $I_A^{(L)}$ scales as $m_A R_A^\ell$, each term in the sum scales as

$$\frac{G m_A m_B}{r_{AB}^2} \left(\frac{R_A}{r_{AB}} \right)^\ell \left(\frac{R_B}{r_{AB}} \right)^{\ell'},$$

and the assumption that $R_A \ll r_{AB}$ implies that each term gets progressively smaller; the equation is exact, but it is most useful as a starting point for an approximation scheme. For many applications involving a small ratio R_A/r_{AB} , the sums can be safely truncated after just a few terms. For other applications, however, a large number of terms may be required. An example is the motion of a satellite in a low Earth orbit, which is sensitive to many of Earth's multipole moments; in geodesy projects such as the Gravity Recovery and Climate Experiment (GRACE) and the CHALLENGING Mini-satellite Payload (CHAMP), Earth's multipole moments up to $\ell \sim 360$ have been measured.

We now specialize to a system of two bodies. The total mass of the binary system is $m := m_1 + m_2$. We define the *relative separation* $\mathbf{r} := \mathbf{r}_1 - \mathbf{r}_2$, and work in *barycentric* coordinates where

$$\begin{aligned} \mathbf{R} &:= (m_1/m) \mathbf{r}_1 + (m_2/m) \mathbf{r}_2 = \mathbf{0}, \\ \mathbf{r}_1 &= (m_2/m) \mathbf{r}, \end{aligned}$$

$$\mathbf{r}_2 = -(m_1/m)\mathbf{r}. \quad (43)$$

Defining $r := |\mathbf{r}|$, $\mathbf{n} := \mathbf{r}/r$, $\mathbf{v} := d\mathbf{r}/dt = \mathbf{v}_1 - \mathbf{v}_2$, and $\mathbf{a} := d\mathbf{v}/dt = \mathbf{a}_1 - \mathbf{a}_2$, we obtain the effective one-body equation of motion

$$\begin{aligned} a^j = & -Gm \frac{n^j}{r^2} + Gm \sum_{\ell=2}^{\infty} \frac{1}{\ell!} \left[\frac{I_1^{(L)}}{m_1} + (-1)^\ell \frac{I_2^{(L)}}{m_2} \right] \partial_{jL} \left(\frac{1}{r} \right) \\ & + Gm \sum_{\ell=2}^{\infty} \sum_{\ell'=2}^{\infty} \frac{(-1)^{\ell'}}{\ell! \ell'!} \frac{I_1^{(L)}}{m_1} \frac{I_2^{(L')}}{m_2} \partial_{jLL'} \left(\frac{1}{r} \right). \end{aligned} \quad (44)$$

Specializing further, we assume that only body 2 has significant multipole moments, and that it is symmetric about an axis aligned with the unit vector \mathbf{e} . It is straightforward to show that

$$I_2^{(L)} = -m_2 R_2^\ell (J_\ell)_2 e_2^{(L)}, \quad (45)$$

where J_ℓ is a dimensionless multipole moment, defined by

$$J_\ell := -\sqrt{\frac{4\pi}{2\ell+1}} \frac{1}{MR^\ell} \mathcal{Y}_{\ell 0}^{(L)} I_{(L)}. \quad (46)$$

We inserted the label “2” on all quantities (such as mass, radius, symmetry axis, and multipole moments) to indicate that they belong to body 2. Equation (44) can then be written as

$$\mathbf{a} = \frac{Gm}{r^2} \left[-\mathbf{n} + \sum_{\ell=2}^{\infty} (J_\ell)_2 \left(\frac{R_2}{r} \right)^\ell \left(\mathbf{n} \frac{dP_{\ell+1}}{d\mu} - \mathbf{e}_2 \frac{dP_\ell}{d\mu} \right) \right], \quad (47)$$

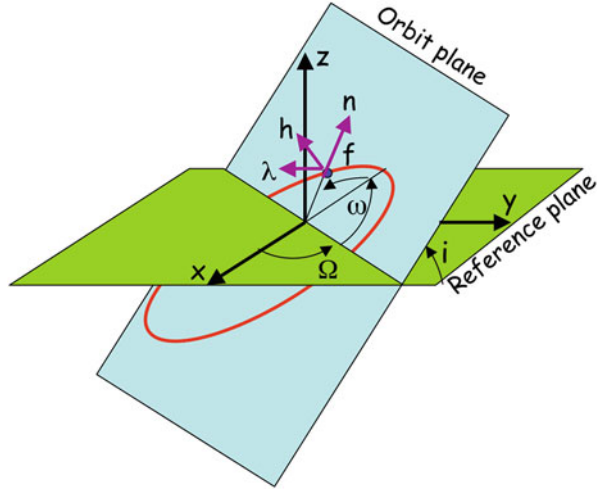
where $P_\ell(\mu)$ is the Legendre polynomial of $\mu = \mathbf{n} \cdot \mathbf{e}_2$. Note that the acceleration has a radial piece $\propto \mathbf{n}$ and a piece parallel to the companion’s symmetry axis \mathbf{e}_2 .

2.4 Newtonian Orbital Dynamics

Kepler’s problem is the determination of the motion of two bodies subjected to their mutual gravitational attraction, under the assumption that each body can be taken to be spherically symmetric. This is the simplest problem of celestial mechanics, but also one of the most relevant, because the motion of any planet around the Sun can, to a good first approximation, be calculated while ignoring the effects of the other planets. It is also a problem that can be solved exactly and completely, in terms of simple equations. In this case, dropping all terms involving multipole moments in Eq. (44), we obtain

$$\mathbf{a} = -\frac{Gm}{r^2} \mathbf{n}. \quad (48)$$

Fig. 1 Keplerian orbit in space



Because the potential is static and spherically symmetric, the equation of motion admits a conserved energy and angular momentum, given by

$$E = \frac{1}{2}\mu v^2 - \frac{G\mu m}{r},$$

$$\mathbf{L} = \mu \mathbf{r} \times \mathbf{v}, \quad (49)$$

where $\mu := m_1 m_2 / (m_1 + m_2)$ is the *reduced mass*. The constancy of \mathbf{L} implies that the motion lies in a plane perpendicular to \mathbf{L} .

In order to write down the general solution of the Kepler problem, we introduce coordinates (X, Y, Z) , with the $X - Y$ plane called the reference plane. This is arbitrary, and is often dictated by convention or convenience. For example, in the description of planetary motion in the solar system, the reference plane is chosen to coincide with Earth's own orbital plane (called the *ecliptic*); in the description of satellites orbiting the Earth, the reference plane is Earth's equatorial plane; for binary star systems, the reference plane is the plane of the sky. In each case, the direction of the X -axis is selected by convention.

The situation is represented in Fig. 1, which shows the orbit plane crossing the reference plane at an angle i called the *inclination*; this is the angle between the z -direction of the orbit frame and the Z -direction of the reference frame. The line of intersection between the two planes is known as the *line of nodes*, and the point at which the orbit cuts the reference plane from below is the *ascending node*. The angle Ω between the X -direction and the line of nodes is the *longitude of the ascending node*. The diagram also shows ω , the *longitude of pericenter*, defined as the angle between the line of nodes and the direction to the pericenter, the point of minimum r , as measured in the orbit plane.

Elementary methods show that the solution to the Keplerian two-body problem is given by

$$\mathbf{r} = r \mathbf{n}, \quad (50)$$

where

$$r = \frac{p}{1 + e \cos f}, \quad (51)$$

where e is the eccentricity, and p is the semi-latus rectum, related to the angular momentum per unit reduced mass h by $h := |\mathbf{r} \times \mathbf{v}| = \sqrt{Gmp}$. The radial unit vector \mathbf{n} , along with the unit vector $\hat{\mathbf{h}} := \mathbf{h}/h$ normal to the orbit plane, and the third vector λ perpendicular to them both, are given by

$$\begin{aligned} \mathbf{n} = & [\cos \Omega \cos(\omega + f) - \cos \iota \sin \Omega \sin(\omega + f)] \mathbf{e}_X \\ & + [\sin \Omega \cos(\omega + f) + \cos \iota \cos \Omega \sin(\omega + f)] \mathbf{e}_Y \\ & + \sin \iota \sin(\omega + f) \mathbf{e}_Z, \end{aligned} \quad (52a)$$

$$\begin{aligned} \lambda = & [-\cos \Omega \sin(\omega + f) - \cos \iota \sin \Omega \cos(\omega + f)] \mathbf{e}_X \\ & + [-\sin \Omega \sin(\omega + f) + \cos \iota \cos \Omega \cos(\omega + f)] \mathbf{e}_Y \\ & + \sin \iota \cos(\omega + f) \mathbf{e}_Z, \end{aligned} \quad (52b)$$

$$\mathbf{h} = \sin \iota \sin \Omega \mathbf{e}_X - \sin \iota \cos \Omega \mathbf{e}_Y + \cos \iota \mathbf{e}_Z. \quad (52c)$$

The relation between the angle f , known as the true anomaly, and time is given by integrating the equation

$$\frac{df}{dt} = \frac{h}{r^2}. \quad (53)$$

The total listing of orbital elements therefore consists of the size and shape elements p and e , the orientational elements ι , Ω , and ω , and the time element T arising from the constant of integration from Eq. (53); the total number of elements is six, corresponding precisely to the number of initial conditions $(\mathbf{x}(0), \mathbf{v}(0))$ required to select a unique solution to Kepler's problem.

The semimajor axis a is defined by

$$a := \frac{1}{2}(r_{\min} + r_{\max}) = \frac{p}{1 - e^2}, \quad (54)$$

and the orbital period P is given by

$$P = 2\pi a^{3/2} (Gm)^{-1/2}. \quad (55)$$

Other useful orbital quantities expressed in this language include

$$\begin{aligned} \mathbf{v} &= \dot{r} \mathbf{n} + (h/r) \lambda, \\ \dot{r} &= (he/p) \sin f, \\ E &= -G\mu m/2a. \end{aligned} \quad (56)$$

The constancy of p is directly tied to conservation of angular momentum h , and the constancy of $a := p/(1 - e^2)$ is tied to conservation of energy (so that constancy of e is also assured). In addition, the appearance of T as an integration constant was expected from the fact that the gravitational potential Gm/r does not depend

explicitly on time. The constancy of ι and Ω is related to the spherical symmetry of the problem, which fixed the direction of \mathbf{L} and the orbital plane.

The constancy of ω is a very important property of a Keplerian orbit: it ensures that the orientation of the orbit within its plane stays fixed, that the position of the pericenter does not move, and when the orbit is bound, that the orbit retraces itself after each orbital cycle. The constancy of ω is the result of a hidden, dynamical symmetry of Kepler's problem, associated with the specific $1/r$ nature of the gravitational potential; the symmetry does not exist for other potentials. The symmetry gives rise to a conservation statement for the *Runge–Lenz vector*, defined by

$$\mathbf{A} := \frac{\mathbf{v} \times \mathbf{h}}{Gm} - \mathbf{n}. \quad (57)$$

A short computation using $\mathbf{a} = -Gm\mathbf{n}/r^2$ shows that

$$\frac{d\mathbf{A}}{dt} = \mathbf{0}, \quad (58)$$

and the manipulations do indeed reveal that constancy of \mathbf{A} relies on the specific form of the gravitational acceleration. The Runge–Lenz vector can be evaluated explicitly by making use of the Keplerian results obtained previously. The result is

$$\mathbf{A} = e(\cos \omega \mathbf{e}_x + \sin \omega \mathbf{e}_y) := e\mathbf{e}_P, \quad (59)$$

where \mathbf{e}_x points along the line of nodes, and \mathbf{e}_y and $\hat{\mathbf{h}}$ complete the basis vectors for the orbital plane; it reveals that the vector \mathbf{A} points in the fixed direction of the pericenter \mathbf{e}_P . The vector has a length e , and constancy of \mathbf{A} implies that ω is a constant of the motion. This same dynamical symmetry is also responsible for the high degree of degeneracy in the quantum-mechanical energy levels of the nonrelativistic hydrogen atom. However, any deviation from a pure $1/r$ potential will generically cause a variation in the angle ω and a splitting of energy levels in hydrogen. There is also a similar dynamical symmetry for the spherical harmonic oscillator potential $\propto r^2$, both classical and quantum.

Five of the orbital elements that we have described can be expressed directly in terms of the variables \mathbf{x} and \mathbf{v} of the orbit. A simple calculation shows

$$p = \frac{h^2}{Gm}, \quad (60a)$$

$$e = |\mathbf{A}|, \quad (60b)$$

$$\cos \iota = \frac{\mathbf{h} \cdot \mathbf{e}_Z}{h}, \quad (60c)$$

$$\sin \iota \sin \Omega = \frac{\mathbf{h} \cdot \mathbf{e}_X}{h}, \quad (60d)$$

$$\sin \iota \sin \omega = \frac{\mathbf{A} \cdot \mathbf{e}_Z}{e}. \quad (60e)$$

2.5 Osculating Orbit Elements and the Perturbed Kepler Problem

In the real world, the Kepler two-body problem does not apply because one or more of the bodies may not be spherically symmetric, because of the presence of other bodies or mass in the system, or because of relativistic effects. Let us assume that the relative acceleration $\mathbf{a} := \mathbf{a}_1 - \mathbf{a}_2$ between two bodies is given by

$$\mathbf{a} = -\frac{Gm}{r^2}\mathbf{n} + \mathbf{f}, \quad (61)$$

where \mathbf{f} is a perturbing force per unit mass, which may depend on \mathbf{r} , \mathbf{v} , and time. The solution of this equation is no longer a conic section of the Kepler problem. However, whatever the solution is, at any given time t_0 , for $\mathbf{r}(t_0)$, $\mathbf{v}(t_0)$, there exists a Keplerian orbit with orbit elements e_0 , a_0 , ω_0 , Ω_0 , ι_0 , and T_0 that corresponds to those values, as we constructed in the previous section. In other words, there is a Keplerian orbit that is *tangent* to the orbit in question at the time t_0 , commonly called the *osculating orbit*.

However, because of the perturbing acceleration, at a later time, the orbit will not be the same Keplerian orbit, but will be tangent to a new osculating orbit, with new elements e' , a' , and so on. The idea then is to study a general orbit with the perturbing acceleration \mathbf{f} by finding the sequence of osculating orbits parametrized by $e(t)$, $a(t)$, and so on. If the perturbing acceleration is small in a suitable sense, then since the orbit elements of the original Kepler motion are constants, we might hope that the osculating orbit elements will vary slowly with time.

Mathematically, this approach is identical to the method of “variation of parameters” used to solve differential equations, such as the harmonic oscillator with a slowly varying frequency.

In this case, we replace our Keplerian solution for the motion with the following *definitions*:

$$\mathbf{r} := r\mathbf{n}, \quad (62a)$$

$$r := \frac{p}{1 + e \cos f}, \quad (62b)$$

$$\mathbf{v} := \frac{he \sin f}{p}\mathbf{n} + \frac{h}{r}\boldsymbol{\lambda}, \quad (62c)$$

$$p := a(1 - e^2), \quad (62d)$$

$$h^2 := Gmp, \quad (62e)$$

where the unit vectors \mathbf{n} and $\boldsymbol{\lambda}$ are given by Eq. (52).

We first decompose \mathbf{f} as

$$\mathbf{f} = \mathcal{R}\mathbf{n} + \mathcal{S}\boldsymbol{\lambda} + \mathcal{W}\hat{\mathbf{h}}, \quad (63)$$

in terms of components \mathcal{R} , \mathcal{S} , and \mathcal{W} .

The effect of the perturbing acceleration \mathbf{f} on the vectors \mathbf{h} and \mathbf{A} can be calculated by appealing directly to their definitions. We find that

$$\frac{d\mathbf{h}}{dt} = \mathbf{r} \times \mathbf{f} = -r\mathcal{W}\lambda + r\mathcal{S}\hat{\mathbf{h}}, \quad (64)$$

and

$$Gm\frac{d\mathbf{A}}{dt} = \mathbf{f} \times \mathbf{h} + \mathbf{v} \times (\mathbf{r} \times \mathbf{f}) = 2h\mathcal{S}\mathbf{n} - (h\mathcal{R} + r\dot{r}\mathcal{S})\lambda - r\dot{r}\mathcal{W}\hat{\mathbf{h}}. \quad (65)$$

These equations imply that

$$\frac{dh}{dt} = r\mathcal{S}, \quad (66)$$

and therefore, that \mathcal{S} produces a change in the magnitude of the angular-momentum vector, while \mathcal{W} produces a change in its direction. Similarly, both \mathcal{R} and \mathcal{S} produce a change in \mathbf{A} 's magnitude, as well as a change of direction orthogonal to $\hat{\mathbf{h}}$.

We can now systematically develop equations for the variations with time of the osculating orbit elements. For example, since $\mathbf{h} \cdot \mathbf{e}_Z = h \cos \iota$, then $\dot{\mathbf{h}} \cdot \mathbf{e}_Z = \dot{h} \cos \iota - \sin \iota (d\iota/dt) = r\mathcal{S} \cos \iota - r\mathcal{W} \cos(\omega + f) \sin \iota$, with the result that $d\iota/dt = (r\mathcal{W}/h) \cos(\omega + f)$. Similarly, since $\mathbf{h} \cdot \mathbf{e}_Y = -h \sin \iota \cos \Omega$, then taking the derivative of both sides and substituting our previous results for $\dot{\mathbf{h}}$, \dot{h} and $d\iota/dt$, we obtain $\sin \iota \dot{\Omega} = (r\mathcal{W}/h) \sin(\omega + f)$. To obtain \dot{e} , we note that $e\dot{e} = \mathbf{A} \cdot \dot{\mathbf{A}}$, and use the fact that $\mathbf{A} = \mathbf{n} \cos f - \lambda \sin f$. For \dot{a} , we use the definition $h^2 = Gma(1 - e^2)$, from which $\dot{a}/a = 2\dot{h}/h + 2e\dot{e}/(1 - e^2)$. For $\dot{\omega}$, we use the fact that $\mathbf{A} \cdot \mathbf{e}_Z = e \sin \iota \sin \omega$, combined with previous results for \dot{e} and $d\iota/dt$. The final equations for the osculating orbit elements are

$$\frac{da}{dt} = \frac{2a^2}{h} \left(\mathcal{S} \frac{p}{r} + \mathcal{R} e \sin f \right), \quad (67a)$$

$$\frac{de}{dt} = \frac{1 - e^2}{h} \left(\mathcal{R} a \sin f + \frac{\mathcal{S}}{er} (ap - r^2) \right), \quad (67b)$$

$$\frac{d\omega}{dt} = -\mathcal{R} \frac{p}{eh} \cos f + \mathcal{S} \frac{p + r}{eh} \sin f - \mathcal{W} \frac{r}{h} \cot \iota \sin(\omega + f), \quad (67c)$$

$$\sin \iota \frac{d\Omega}{dt} = \mathcal{W} \frac{r}{h} \sin(\omega + f), \quad (67d)$$

$$\frac{d\iota}{dt} = \mathcal{W} \frac{r}{h} \cos(\omega + f). \quad (67e)$$

Notice that the orbit elements a and e are affected only by components of \mathbf{f} in the plane of the orbit, while the elements Ω and ι are affected only by the component out of the plane. The pericenter change has both, but this is because of the combination of intrinsic, in-plane perturbations (the first two terms) with the perturbation of the line of nodes from which ω is measured (the third term).

We are missing an equation for the variation of T , the time of pericenter passage, which determines the true anomaly f via Eq. (53). If one is engaged in real-life orbit

determinations, such as navigating a spacecraft, timing a binary pulsar's orbit, or developing ephemerides, then the variation of T is an important quantity. However, if one is interested in elucidating physical effects on an orbit, it is more practical to close the system of equations by providing an expression for df/dt , from which the true anomaly can be obtained directly. Because f is the angle between the (varying) pericenter and the position vector \mathbf{n} , we have that $\cos f = \mathbf{n} \cdot \mathbf{e}_P$, and this can immediately be differentiated with respect to time. After some simplification, we find

$$\frac{df}{dt} = \frac{h}{r^2} - \left(\frac{d\omega}{dt} + \cos \iota \frac{d\Omega}{dt} \right), \quad (68)$$

which shows that df/dt differs from the usual Keplerian expression by a term $d\omega/dt + \cos \iota d\Omega/dt$ which possesses a direct geometrical meaning. We recall that ω is the angle from the (varying) pericenter to the (varying) line of nodes, as measured in the orbital plane, while Ω is the angle from the line of nodes to the (fixed) X -direction, as measured in the reference X - Y plane. The combination $d\omega + \cos \iota d\Omega$ can then be seen to describe the change in the direction to the pericenter relative to the X -direction, as measured entirely in the orbital plane. The non-Keplerian terms in Eq. (68) therefore appear because the true anomaly f is now measured relative to a varying set of directions. In fact, it is common to discuss the variations in the combined orbit “element”

$$d\varpi := d\omega + \cos \iota d\Omega. \quad (69)$$

The formalism of osculating orbital elements, in the formulation of Eqs. (67) and (68), is *exactly equivalent* to the original formulation of the equations of motion in Eq. (61); no approximations have been introduced in the transcription. The usefulness of the formalism, however, is most immediate when the perturbing force is small, so that the changes in the orbital elements are small. In such a context, one can achieve a very good approximation of the orbital dynamics by inserting the constant, zeroth-order values on the right-hand side of the equations, and integrating with respect to t to get the first-order changes. In such applications, it is usually convenient to use f as an independent variable instead of t , and in this approximate context, one can neglect the non-Keplerian terms on the right-hand side of Eq. (68). The system of osculating equations become

$$\frac{dp}{df} \simeq 2 \frac{p^3}{Gm} \frac{1}{(1 + e \cos f)^3} \mathcal{S}, \quad (70a)$$

$$\frac{de}{df} \simeq \frac{p^2}{Gm} \left[\frac{\sin f}{(1 + e \cos f)^2} \mathcal{R} + \frac{2 \cos f + e(1 + \cos^2 f)}{(1 + e \cos f)^3} \mathcal{S} \right], \quad (70b)$$

$$\frac{d\iota}{df} \simeq \frac{p^2}{Gm} \frac{\cos(\omega + f)}{(1 + e \cos f)^3} \mathcal{W}, \quad (70c)$$

$$\sin \iota \frac{d\Omega}{df} \simeq \frac{p^2}{Gm} \frac{\sin(\omega + f)}{(1 + e \cos f)^3} \mathcal{W}, \quad (70d)$$

$$\frac{d\omega}{df} \simeq \frac{1}{e} \frac{p^2}{Gm} \left[-\frac{\cos f}{(1 + e \cos f)^2} \mathcal{R} + \frac{2 + e \cos f}{(1 + e \cos f)^3} \sin f \mathcal{S} - e \cot \iota \frac{\sin(\omega + f)}{(1 + e \cos f)^3} \mathcal{W} \right], \quad (70e)$$

with

$$\frac{dt}{df} \simeq \sqrt{\frac{p^3}{Gm}} \frac{1}{(1 + e \cos f)^2} \quad (71)$$

providing the temporal information.

2.6 Non-Keplerian Behavior: Worked Examples

In this section, we work out some real-world examples to illustrate the use of osculating orbit elements and other non-Keplerian effects.

Perturbations by a Third Body

Consider a two-body system in the presence of a distant third body. We assume that the effect of the third body can be treated as a perturbation of the Keplerian two-body orbit of the primary bodies. From the N-body equation of motion (42), we write down the equation of motion of the two primary bodies,

$$\begin{aligned} \mathbf{a}_1 &= -G \frac{m_2 \mathbf{r}_{12}}{r_{12}^3} - G \frac{m_3 \mathbf{r}_{13}}{r_{13}^3}, \\ \mathbf{a}_2 &= G \frac{m_1 \mathbf{r}_{12}}{r_{12}^3} - G \frac{m_3 \mathbf{r}_{23}}{r_{23}^3}. \end{aligned} \quad (72)$$

We now assume that $r_{12} \ll r_{23}$, so that we can expand

$$\frac{r_{13}^i}{r_{13}^3} \approx \frac{r_{23}^i}{r_{23}^3} - \sum_{\ell=1}^{\infty} \frac{1}{\ell!} r_{12}^L \partial^{(iL)} \left(\frac{1}{r_{23}} \right). \quad (73)$$

The relative acceleration $\mathbf{a} := \mathbf{a}_1 - \mathbf{a}_2$ can then be written

$$\mathbf{a} = -G \frac{m \mathbf{r}}{r^3} - G \frac{m_3}{R^3} [\mathbf{r} - 3\mathbf{N}(\mathbf{N} \cdot \mathbf{r})] + O(m_3 r^2 / R^4), \quad (74)$$

where we define $R := |\mathbf{r}_{23}|$, and $\mathbf{N} := \mathbf{r}_{23}/|\mathbf{r}_{23}|$, and where we have kept only the $\ell = 1$ term in the sequence. The first term on the right-hand side of Eq. (74) is the normal Newtonian acceleration for the effective one-body problem. As long as $(m_3/m)(r/R)^3 \ll 1$, we can treat the second term in Eq. (74) as a disturbing

acceleration of a Keplerian orbit. From Eq. (74), we obtain the components of the disturbing acceleration

$$\begin{aligned}\mathcal{R} &= -A[1 - 3(\mathbf{N} \cdot \mathbf{n})^2], \\ \mathcal{S} &= 3A(\mathbf{N} \cdot \mathbf{n})(\mathbf{N} \cdot \boldsymbol{\lambda}), \\ \mathcal{W} &= 3A(\mathbf{N} \cdot \mathbf{n})(\mathbf{N} \cdot \hat{\mathbf{h}}),\end{aligned}\tag{75}$$

where $A = Gm_3r/R^3$.

We now focus on a specific example: the orbit of a planet like Mercury, perturbed by an outer planet, such as Jupiter. Let Mercury be in an orbit with osculating orbit elements a , e , and ω ; we will assume that the orbit lies on the reference plane (the ecliptic in this case), so that $i = 0$ and Ω is not defined and can be put equal to zero. The orbit is then described by Eqs. (62a–62e). For simplicity, assume that the perturbing planet is in a circular orbit in the same reference plane, with $R = \text{constant}$, and

$$\mathbf{N} = \mathbf{e}_x \cos F + \mathbf{e}_y \sin F,\tag{76}$$

where F is its eccentric anomaly, with constant $dF/dt = (Gm/R^3)^{1/2}$. Using Eqs. (52a) and (52b), we show that $\mathbf{N} \cdot \mathbf{n} = \cos(f - F + \omega)$, $\mathbf{N} \cdot \boldsymbol{\lambda} = -\sin(f - F + \omega)$, and $\mathbf{N} \cdot \hat{\mathbf{h}} = 0$. We are interested in the perihelion advance of our planet, so working to first order in small perturbations, we combine Eqs. (70d), (70e), and (69) to obtain

$$\begin{aligned}\frac{d\varpi}{df} &= -\frac{r^2}{h^2e} [p\mathcal{R} \cos f - (p+r)\mathcal{S} \sin f] \\ &= \frac{Gm_3p^4}{h^2R^3e} \left\{ \frac{\cos f}{(1+e \cos f)^3} [1 - 3 \cos^2(f - F + \omega)] \right. \\ &\quad \left. - 3 \frac{(2 + e \cos f) \sin f}{(1 + e \cos f)^4} \cos(f - F + \omega) \sin(f - F + \omega) \right\}.\end{aligned}\tag{77}$$

In first-order perturbation theory, we can set the orbit elements equal to constants on the right-hand side and then integrate with respect to f to obtain $\varpi(f)$. But to keep things simple, we will further assume that the perturbing planet is far enough away that it barely moves during one orbit of our Mercury, that is, that $dF/dt \ll df/dt$, which follows from $(m_3/m)(r/R)^3 \ll 1$. Hence we set $F = \text{constant}$ in (77). We will also expand in powers of the eccentricity e . The solution for $\varpi(f)$ will therefore be proportional to sines and cosines of multiples of f , indicating that the osculating perihelion angle moves back and forth during the orbit as a result of the external perturbation.

Notice that, in the expansion of Eq. (77) in eccentricity, the leading term will be proportional to $1/e$, which might be worrisome in the limit $e \rightarrow 0$. However, it can be shown that the actual orbital displacements resulting from a change in ϖ are proportional to $e\delta\varpi$, so those observable displacements are well-behaved in the small-eccentricity limit.

Table 1 Planetary contributions to Mercury's perihelion advance (in arcseconds per century)

Planet	Advance
Venus	277.8
Earth	90.0
Mars	2.5
Jupiter	153.6
Saturn	7.3
Total	531.2
Discrepancy	42.9
Modern measured value	42.98 ± 0.04
General relativity prediction	42.98

But an important question is, after one orbit of Mercury, does the osculating perihelion angle return to its original value, that is, is it purely periodic, or does it suffer a *secular* change over each orbit? To explore this, we integrate Eq. (77) over f from zero to 2π , with the result that the net change $\Delta\varpi$ over one orbit is given by

$$\Delta\varpi = 3\pi \frac{m_3}{m} \left(\frac{p}{R}\right)^3 \frac{1}{(1-e^2)^{5/2}} [5 \cos^2(F - \omega) - 2]. \quad (78)$$

The secular perihelion advance itself varies as the distant planet orbits the Sun, but if we average over all positions of the distant planet using $\langle \cos^2(F - \omega) \rangle = 1/2$, we obtain

$$\Delta\varpi = \frac{3\pi}{2} \frac{m_3}{m} \left(\frac{a}{R}\right)^3 (1-e^2)^{1/2}. \quad (79)$$

Inserting standard orbit elements for Mercury and Jupiter, we see that $\Delta\varpi = 1.79 \times 10^{-6}$. The smallness of this effect is what justifies working only at first order in perturbation theory. It is customary to express the perihelion advance as a rate in units of arcseconds (as) per century, so using the fact that 1 radian = 2.063×10^5 as, and that Mercury's orbital period is 0.24 years, we obtain $d\varpi/dt = 154$ as per century. This is very close to the accurately computed value shown in Table 1. However, for perturbations of Mercury's orbit by the Earth, we get 69.2 as per century, quite a bit off from the accurate value of 90. But in this case $r/R \approx 0.38$, so that it is necessary to include the higher-order terms from the expansion in Eq. (74). Also, whereas holding Jupiter's position fixed while integrating over Mercury's orbit and then averaging over Jupiter's position might have been a good approximation since Jupiter's period is 40 times longer than Mercury's, it is not a good approximation in the case of Earth, whose period is only four times longer. Nevertheless, the method of osculating orbit elements allows us to incorporate such details systematically in a straightforward manner, leading to the accurately calculated planetary perturbations listed in Table 1. Still, the total effect of the planetary perturbations on Mercury's orbit falls well short of the observed perihelion advance of 574 as per century. In a later section, we will return to this question once we have obtained the equations of motion including relativistic terms.

Orbits Around a Nonspherical Body

We now consider a two-body problem where one of the bodies has nonzero multipole moments, typically caused by rotation. This could mean, for example, Mercury orbiting the oblate, rotating Sun, a satellite orbiting the Earth with all its multipole moments, or a neutron star orbiting a massive, rotating normal star. In this case, we can use Eq. (47), here restricted to $\ell = 2$:

$$\mathbf{a} = -\frac{Gm}{r^2} \left[\mathbf{n} - \frac{3}{2}(J_2)_2 \frac{R_2^2}{r^2} \left(5\mathbf{n}(\mathbf{e} \cdot \mathbf{n})^2 - 2\mathbf{e}(\mathbf{e} \cdot \mathbf{n}) - \mathbf{n} \right) \right], \quad (80)$$

where R_2 , $(J_2)_2$, and \mathbf{e} are the radius, dimensionless quadrupole moment and symmetry axis direction of the central body. Hence

$$\begin{aligned} \mathcal{R} &= \frac{1}{2}A[3(\mathbf{e} \cdot \mathbf{n})^2 - 1], \\ \mathcal{S} &= -A(\mathbf{e} \cdot \boldsymbol{\lambda})(\mathbf{e} \cdot \mathbf{n}), \\ \mathcal{W} &= -A(\mathbf{e} \cdot \hat{\mathbf{h}})(\mathbf{e} \cdot \mathbf{n}), \end{aligned} \quad (81)$$

where $A = 3(Gm/r^2)(J_2)_2(R_2/r)^2$. We choose \mathbf{e} to be along the z -axis, and consider an orbit with osculating elements a , e , ω , ι , and Ω . Then from Eq. (52), we have $\mathbf{e} \cdot \mathbf{n} = \sin \iota \sin(f + \omega)$, $\mathbf{e} \cdot \boldsymbol{\lambda} = \sin \iota \cos(f + \omega)$, and $\mathbf{e} \cdot \hat{\mathbf{h}} = \cos \iota$. Working in first-order perturbation theory, we integrate Eqs. (70a–70e), over 2π and obtain the secular changes in the orbit elements, $\Delta a = \Delta e = \Delta \iota = 0$, while

$$\Delta \varpi = 3\pi(J_2)_2 \left(\frac{R_2}{p} \right)^2 \left(1 - \frac{3}{2} \sin^2 \iota \right), \quad (82a)$$

$$\Delta \Omega = -3\pi(J_2)_2 \left(\frac{R_2}{p} \right)^2 \cos \iota. \quad (82b)$$

Notice that, even in the case of zero inclination, where $\mathbf{e} \cdot \mathbf{n} = 0$, and thus $\mathcal{S} = \mathcal{W} = 0$, the quadrupole moment still induces a perihelion advance because it modifies the radial dependence of the acceleration from a pure $1/r^2$ behavior. For inclined orbits, the deviation from spherical symmetry induces advances in both the perihelion and the node. In the latter case, the effect is merely a precession of the orbital angular momentum $\mathbf{h} = \mathbf{r} \times \mathbf{v}$ around the symmetry axis \mathbf{e} , as can be obtained directly from Eq. (80),

$$\frac{d\mathbf{h}}{dt} = -3\frac{Gm}{r^2}(J_2)_2 \frac{R_2^2}{r^2} (\mathbf{r} \times \mathbf{e})(\mathbf{e} \cdot \mathbf{n}). \quad (83)$$

From this, we see that $\mathbf{e} \cdot d\mathbf{h}/dt = 0$, so that \mathbf{h} precesses around the \mathbf{e} axis, inducing a rotation of the node.

For Mercury, the perihelion advance induced by the Sun's J_2 is negligible. Inserting the relevant orbit elements and the value $(J_2)_\odot = 2.2 \times 10^{-7}$, we obtain $d\varpi/dt = 0.03$ as/century, below the experimental uncertainties (see Table 1). It

was not always so. In fact, it is extremely difficult to measure J_2 for the Sun because at planetary distances, its effects are too small to be measured. The best way is to send a spacecraft relatively close to the Sun, well inside the orbit of Mercury and to determine precisely how J_2 affects its orbit; despite a number of proposals for such a mission, none has come to pass to date.

During the 1960s, Robert Dicke and Mark Goldenberg, attempted to infer a value of J_2 by measuring the visual shape of the Sun. They claimed a J_2 of order of 2.5×10^{-5} , over 100 times larger than the number quoted above. Later observations of the visible shape of the Sun by Henry Hill and others, along with analyses to try to understand how temperature differences across the Sun might contaminate the observations of apparent shape, did not fully resolve this controversy. The resolution came with the advance of “helioseismology.” This was the discovery that the Sun actually vibrates in a superposition of thousands of normal modes with an array of frequencies, as reflected in the behavior of Doppler-shifted solar spectral lines. The specific pattern of frequencies proves to depend on how much differential rotation is present in the Sun. Through a systematic program of ground-based and space-based observations of the Sun, it was possible to determine the angular velocity profile over much of the solar interior. The conclusion was that the interior does not rotate much faster than the surface, and the solar models built from that information resulted in the value $J_2 = 2.2 \times 10^{-7}$, which is more or less what one would infer from a Sun that rotates uniformly at its observed surface rate. The bottom line is that, as far as Mercury and general relativity are concerned, the solar quadrupole moment does not play a role.

However, J_2 for the Earth does play an important role in the motion of satellites. Using $(J_2)_\oplus = 1.08 \times 10^{-3}$, and assuming a satellite in a circular orbit with semi-major axis a , with orbital period $P = 2\pi(a^3/Gm)^{1/2} = 83.91(a/R)^{3/2}$ minutes, we can show that the secular rate of advance of the node is given by

$$\frac{d\Omega}{dt} = 3639 \left(\frac{R}{a} \right)^{7/2} \cos \iota \text{ } ^\circ\text{yr}^{-1}, \quad (84)$$

where R is the Earth’s radius. Thus, for example, the orbit of the laser geodynamics satellite LAGEOS I, with $a = 1.93R_\oplus$ and $\iota = 109^\circ.8$ will precess at a rate of -123° per year. Suppression of this enormous Newtonian effect in order to detect the possible effect of general relativistic frame-dragging (~ 30 milliarcseconds per year) is a major challenge. The chapter “Fundamental Physics with the LAGEOS Satellites” by R. Peron, addresses how this suppression has been carried out in analysing data from the LAGEOS I and II satellites.

3 General Relativity

3.1 Mathematics of Curved Space-Time

The foundation of general relativity is the space-time metric $g_{\alpha\beta}$, which connects the arbitrary coordinates $x^\alpha = (ct, x^j)$ used to label events in space-time to physical

measurements of time and length made by clocks and rods, via the invariant interval

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta. \quad (85)$$

Under a change of coordinates described by $x^\alpha = f^\alpha(x'^\mu)$, where f^α are functions of the new coordinates x'^μ , the coordinate displacements change according to

$$dx^\alpha = \frac{\partial f^\alpha}{\partial x'^\mu} dx'^\mu, \quad (86)$$

and the space-time interval becomes

$$ds^2 = g_{\alpha\beta} \frac{\partial f^\alpha}{\partial x'^\mu} \frac{\partial f^\beta}{\partial x'^\nu} dx'^\mu dx'^\nu. \quad (87)$$

This expression shows that the metric is replaced by

$$g'_{\mu\nu} = g_{\alpha\beta} \frac{\partial f^\alpha}{\partial x'^\mu} \frac{\partial f^\beta}{\partial x'^\nu} \quad (88)$$

in the new coordinate system, so that $ds^2 = g'_{\mu\nu} dx'^\mu dx'^\nu$. The coordinate displacements change, and the metric also changes, but ds^2 remains the same during a coordinate transformation. In this way, for example, the proper time $d\tau = \sqrt{-ds^2}/c$ between two events as measured by an atomic clock is the same, regardless of the coordinates used to label the space-time events.

It can be shown that, in the neighborhood of any event \mathcal{P} in space-time (and more generally along any world line), one can always find a coordinate system in which

$$g'_{\mu\nu}|_{\mathcal{P}} = \eta_{\mu\nu}, \quad (\partial'_\gamma g'_{\mu\nu})|_{\mathcal{P}} = 0, \quad (89)$$

where $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$ is the Minkowski metric of special relativity. This corresponds to the local freely falling frame, in which freely moving particles move on straight lines and the laws of nongravitational physics take on their standard special-relativistic, Lorentz invariant forms. This embodies the Einstein equivalence principle, the foundation of the geometrical formulation of gravity.

A coordinate system x^α can be used to define a set of basis vectors \mathbf{e}_α , such that a vector describing the displacement from one event in space-time to a neighboring event is given by

$$d\mathbf{x} = \mathbf{e}_\alpha dx^\alpha. \quad (90)$$

An arbitrary vector \mathbf{A} can then be expressed as

$$\mathbf{A} = A^\alpha \mathbf{e}_\alpha \quad (91)$$

in terms of the basis vectors and its components A^α .

The inner product between $d\mathbf{x}$ and itself is the space-time invariant ds^2 . We have $ds^2 = d\mathbf{x} \cdot d\mathbf{x} = (\mathbf{e}_\alpha \cdot \mathbf{e}_\beta) dx^\alpha dx^\beta$, and comparing this with Eq. (85) reveals that

$$g_{\alpha\beta} = \mathbf{e}_\alpha \cdot \mathbf{e}_\beta. \quad (92)$$

If \mathbf{A} is a vector field in space-time, its derivative with respect to coordinate x^β is

$$\partial_\beta \mathbf{A} = (\partial_\beta A^\alpha) \mathbf{e}_\alpha + A^\alpha (\partial_\beta \mathbf{e}_\alpha), \quad (93)$$

in which the first term accounts for the variation of the components, while the second accounts for the variation of the basis vectors. We define

$$\partial_\beta \mathbf{e}_\alpha := \Gamma_{\alpha\beta}^\mu \mathbf{e}_\mu, \quad (94)$$

which states the obvious fact that a change in basis vector induced by a coordinate displacement is itself a vector that can be decomposed in terms of basis vectors. Equation (94) provides a definition for the quantities $\Gamma_{\alpha\beta}^\mu$, which are known as *Christoffel symbols*. Because $\partial_\beta \mathbf{e}_\alpha = \partial^2 \mathbf{x} / \partial x^\alpha \partial x^\beta = \partial_\alpha \mathbf{e}_\beta$, we have that

$$\Gamma_{\beta\alpha}^\mu = \Gamma_{\alpha\beta}^\mu; \quad (95)$$

the Christoffel symbols are symmetric in their lower indices.

With Eq. (94), our previous expression for $\partial_\beta \mathbf{A}$ becomes $\partial_\beta \mathbf{A} = (\nabla_\beta A^\mu) \mathbf{e}_\mu$, where

$$\nabla_\beta A^\mu := \partial_\beta A^\mu + \Gamma_{\alpha\beta}^\mu A^\alpha \quad (96)$$

is known as the *covariant derivative* of the vector components A^μ .

Differentiating the relation $g_{\alpha\beta} = \mathbf{e}_\alpha \cdot \mathbf{e}_\beta$ with respect to x^γ , substituting Eq. (94) and doing some further algebra, we can show that

$$\Gamma_{\alpha\beta}^\mu = \frac{1}{2} g^{\mu\nu} (\partial_\alpha g_{\nu\beta} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta}). \quad (97)$$

The generalization of the covariant derivative to a tensor is:

$$\nabla_\beta A^{\mu\nu} = \partial_\beta A^{\mu\nu} + \Gamma_{\alpha\beta}^\mu A^{\alpha\nu} + \Gamma_{\alpha\beta}^\nu A^{\mu\alpha}. \quad (98)$$

This rule can easily be extended to tensors with an arbitrary number of indices; there is one Christoffel symbol per tensorial index.

Defining $A_\mu := g_{\mu\nu} A^\nu$, we can “lower” indices on vectors and tensors; using the inverse of the metric, identified as $g^{\mu\nu}$, we can raise indices. Then we can show that

$$\nabla_\beta A_\mu = \partial_\beta A_\mu - \Gamma_{\mu\beta}^\alpha A_\alpha, \quad (99)$$

and furthermore that

$$\nabla_\gamma g_{\alpha\beta} = 0. \quad (100)$$

We now examine a timelike world line $x^\alpha = r^\alpha(\tau)$ in a curved space-time, parameterized by proper time τ , the time measured by a clock moving on this world line. The world line’s tangent vector is $u^\alpha = dr^\alpha/d\tau$, and this satisfies the normalization condition

$$g_{\alpha\beta} u^\alpha u^\beta = -c^2. \quad (101)$$

On this world line, let there exist a vector field $\mathbf{A}(\tau)$; we wish to evaluate the derivative of this vector with respect to τ :

$$\frac{d\mathbf{A}}{d\tau} = \left(\frac{dA^\mu}{d\tau} + \Gamma_{\alpha\beta}^\mu A^\alpha u^\beta \right) \mathbf{e}_\mu. \quad (102)$$

The quantity within brackets is the covariant derivative of the component A^μ along the world line. We denote this

$$\frac{DA^\mu}{d\tau} := \frac{dA^\mu}{d\tau} + \Gamma_{\alpha\beta}^\mu A^\alpha u^\beta, \quad (103)$$

so that $d\mathbf{A}/d\tau = (DA^\mu/d\tau)\mathbf{e}_\mu$. When the vector field \mathbf{A} is defined also in a neighborhood of the world line (and not just directly on the world line), it becomes a function of all the space-time coordinates x^α (instead of just proper time τ); then $dA^\mu/d\tau = u^\beta \partial_\beta A^\mu$ and the covariant derivative can be expressed as $DA^\mu/d\tau = u^\beta \nabla_\beta A^\mu$.

The vector \mathbf{A} is *parallel-transported* along the world line when it stays constant in both direction and magnitude. The mathematical statement of this is $d\mathbf{A}/d\tau = 0$, or

$$\frac{DA^\mu}{d\tau} = 0 \quad (\text{parallel transport}). \quad (104)$$

A timelike world line $r^\alpha(\tau)$ is a *geodesic* of the curved space-time when its own tangent vector \mathbf{u} is parallel-transported along the world line. A geodesic, defined in this way, is everywhere *locally straight*. The mathematical statement of the geodesic equation is

$$\frac{Du^\mu}{d\tau} = 0, \quad (105)$$

or

$$\frac{du^\mu}{d\tau} + \Gamma_{\alpha\beta}^\mu u^\alpha u^\beta = 0, \quad (106)$$

or

$$\frac{d^2 r^\mu}{d\tau^2} + \Gamma_{\alpha\beta}^\mu \frac{dr^\alpha}{d\tau} \frac{dr^\beta}{d\tau} = 0. \quad (107)$$

This last form is a system of second-order differential equations for the functions $r^\mu(\tau)$. Given initial conditions $r^\mu(0)$ and $u^\mu(0)$ at some initial time $\tau = 0$, these equations admit a unique solution.

It is useful to note that the geodesic equation (107) can also be obtained on the basis of a variational principle, in which the action functional is the elapsed proper time $\int_1^2 d\tau$ along a parameterized curve $r^\alpha(\tau)$ linking the fixed events 1 and 2. If we write the particle action as

$$S = -mc^2 \int_1^2 d\tau = -mc \int_1^2 \sqrt{-g_{\alpha\beta} \frac{dr^\alpha}{dt} \frac{dr^\beta}{dt}} dt, \quad (108)$$

then its extremization with respect to world line variations returns the geodesic equation.

The symmetry of the Christoffel symbols in the lower indices implies that the action of two covariant derivatives on a scalar field f is independent of their order:

$$\nabla_\alpha \nabla_\beta f - \nabla_\beta \nabla_\alpha f = 0. \quad (109)$$

The same is not true, however, when the covariant derivatives act on a vector field A^μ ; in this case

$$\nabla_\alpha \nabla_\beta A^\mu - \nabla_\beta \nabla_\alpha A^\mu = R^\mu_{\nu\alpha\beta} A^\nu, \quad (110)$$

and the operations do not commute. This equation defines the *Riemann curvature tensor* $R^\mu_{\nu\alpha\beta}$. A lengthy evaluation of the left-hand side of Eq. (110) shows that the Riemann tensor is given explicitly by

$$R^\alpha_{\beta\gamma\delta} = \partial_\gamma \Gamma^\alpha_{\beta\delta} - \partial_\delta \Gamma^\alpha_{\beta\gamma} + \Gamma^\alpha_{\mu\gamma} \Gamma^\mu_{\beta\delta} - \Gamma^\alpha_{\mu\delta} \Gamma^\mu_{\beta\gamma}. \quad (111)$$

The Riemann tensor is evidently antisymmetric in its last two indices. It also possesses additional symmetries that are not immediately revealed by Eqs. (110) and (111). The symmetries are

$$\begin{aligned} R_{\alpha\beta\delta\gamma} &= -R_{\alpha\beta\gamma\delta}, \\ R_{\beta\alpha\gamma\delta} &= -R_{\alpha\beta\gamma\delta}, \\ R_{\mu\alpha\beta\gamma} + R_{\mu\gamma\alpha\beta} + R_{\mu\beta\gamma\alpha} &= 0. \end{aligned} \quad (112)$$

An additional symmetry,

$$R_{\gamma\delta\alpha\beta} = +R_{\alpha\beta\gamma\delta}, \quad (113)$$

is an immediate consequence of the preceding ones. By virtue of these symmetries, the Riemann tensor possesses 20 independent components in a four-dimensional space-time.

Another important set of identities satisfied by the Riemann tensor is

$$\nabla_\alpha R_{\mu\nu\beta\gamma} + \nabla_\gamma R_{\mu\nu\alpha\beta} + \nabla_\beta R_{\mu\nu\gamma\alpha} = 0. \quad (114)$$

These are known as the *Bianchi identities*, and play a fundamental role in Einstein's general relativity.

The Riemann tensor is important as a measure of *geodesic deviation*. Given a family of nearby geodesics parametrized by τ and a deviation vector ξ that joins neighboring geodesics, one can show that the relative acceleration between geodesics is given by

$$\frac{D^2 \xi^\alpha}{d\tau^2} = -R^\alpha_{\beta\gamma\delta} u^\beta \xi^\gamma u^\delta. \quad (115)$$

Other curvature tensors can be defined from the Riemann tensor. By contracting the first and third indices of the Riemann tensor, we obtain the *Ricci tensor*

$$R_{\alpha\beta} := R^\mu_{\alpha\mu\beta}. \quad (116)$$

The symmetries of the Riemann tensor imply that $R_{\beta\alpha} = R_{\alpha\beta}$, and the Ricci tensor possesses ten independent components. By contracting its indices, we obtain the *Ricci scalar*

$$R := R^\mu{}_\mu = g^{\alpha\beta} R_{\alpha\beta}. \quad (117)$$

Closely related to the Ricci tensor is the *Einstein tensor*

$$G_{\alpha\beta} := R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R, \quad (118)$$

which is also symmetric in its indices: $G_{\beta\alpha} = G_{\alpha\beta}$. The Einstein tensor possesses ten independent components, and its trace is given by $G := g^{\alpha\beta} G_{\alpha\beta} = -R$, because $g^{\alpha\beta} g_{\alpha\beta} = \delta^\alpha{}_\alpha = 4$.

The Bianchi identities of Eq. (114) give rise to

$$\nabla_\beta G^{\alpha\beta} = 0 \quad (119)$$

after two contractions of their indices. Equation (119) is known as the *contracted Bianchi identities*.

3.2 Physics in Curved Spacetime

The physical laws of the standard model of particles and fields are normally formulated in the Lorentz invariant language of special relativity. For example, Maxwell's equations have the form $\partial_\beta F^{\alpha\beta} = \mu_0 j_e^\alpha$, and $\partial_\alpha F_{\beta\gamma} + \partial_\gamma F_{\alpha\beta} + \partial_\beta F_{\gamma\alpha} = 0$, where $F^{\alpha\beta}$ is the antisymmetric field tensor, j_e^α is the electric current vector, and μ_0 is the permeability of vacuum. A freely moving neutral particle moves with unchanging four-velocity, in other words $du^\alpha/d\tau = u^\beta \partial_\beta u^\alpha = 0$.

Since all the fundamental laws of physics are derivable from an action, then, by virtue of the Lorentz invariance of the matter action, the laws imply a conservation statement for the energy-momentum tensor $T^{\alpha\beta} := -2(-\eta)^{-1/2} \delta \mathcal{L}_M / \delta \eta_{\alpha\beta}$ given by $\partial_\beta T^{\alpha\beta} = 0$.

The Einstein equivalence principle tells us that these laws are now to be regarded as being valid in a local freely falling frame, where the Minkowski metric $\eta_{\alpha\beta} = g_{\alpha\beta}|_{\mathcal{P}}$, is the transformed version of the space-time metric. To find the laws in a form valid in any frame or coordinate system, one only has to transform back. As a result of this transformation, $\eta_{\alpha\beta} \rightarrow g_{\alpha\beta}$, and $\partial_\alpha \rightarrow \nabla_\alpha$. Thus, in curved space-time, the laws mentioned above take the form

$$\begin{aligned} \text{Maxwell's equations : } \nabla_\beta F^{\alpha\beta} &= \mu_0 j_e^\alpha, \\ \nabla_\alpha F_{\beta\gamma} + \nabla_\gamma F_{\alpha\beta} + \nabla_\beta F_{\gamma\alpha} &= 0, \end{aligned} \quad (120a)$$

$$\text{Geodesic equation : } Du^\alpha/d\tau = u^\beta \nabla_\beta u^\alpha = 0, \quad (120b)$$

$$\text{Energy – momentum conservation : } \nabla_\beta T^{\alpha\beta} = 0, \quad (120c)$$

with $T^{\alpha\beta} := -2(-g)^{-1/2}\delta\mathcal{L}_M/\delta g_{\alpha\beta}$.

For most astrophysical applications, we will be more interested in a phenomenological description of hydrodynamics than in the underlying fundamental laws. In a manner similar to Newtonian theory, we define a perfect fluid with the proper mass density ρ , proper density of internal (thermodynamic) energy ε , proper density of total energy $\mu = \rho c^2 + \varepsilon$, and pressure p . The densities are all measured in a freely falling frame that is momentarily comoving with a selected fluid element; the four-velocity of this frame is denoted u^α , and the mass current is $j^\alpha = \rho u^\alpha$. Transforming from the local inertial frame, we can show that the fluid's energy-momentum tensor $T^{\alpha\beta}$, is given by

$$T^{\alpha\beta} = (\mu + p)u^\alpha u^\beta / c^2 + pg^{\alpha\beta}. \quad (121)$$

The statement of mass conservation now takes the form of

$$\nabla_\alpha j^\alpha = \frac{1}{\sqrt{-g}}\partial_\alpha(\sqrt{-g}j^\alpha) = 0. \quad (122)$$

Substitution of $j^\alpha = \rho u^\alpha$ into Eq. (122) yields

$$\frac{d\rho}{d\tau} + \rho\nabla_\alpha u^\alpha = 0. \quad (123)$$

Since $\nabla_\alpha u^\alpha = \mathcal{V}^{-1}d\mathcal{V}/d\tau$, this equation implies $d(\rho\mathcal{V})/d\tau = 0$ as expected. Alternatively, Eq. (122) implies that

$$\frac{\partial\rho^*}{\partial t} + \nabla \cdot (\rho^*\mathbf{v}) = 0, \quad (124)$$

where

$$\rho^* := \rho\sqrt{-g}u^0 \quad (125)$$

is sometimes called the *conserved density*. Because it satisfies a continuity equation exactly parallel to Eq. (4), the results gathered in Eqs. (11)–(14) now apply to the analogous integrals using ρ^* .

Substitution of Eq. (121) into Eq. (120c) produces

$$\frac{d\mu}{d\tau} + (\mu + p)\nabla_\beta u^\beta = 0 \quad (126)$$

and

$$(\mu + p)\frac{Du^\alpha}{d\tau} + c^2(g^{\alpha\beta} + u^\alpha u^\beta / c^2)\nabla_\beta p = 0. \quad (127)$$

The first equation gives rise to

$$\rho\frac{d\varepsilon}{d\tau} - (\varepsilon + p)\frac{d\rho}{d\tau} = 0, \quad (128)$$

which can be easily seen to be equivalent to the local first law of thermodynamics $d(\varepsilon\mathcal{V}) + pd\mathcal{V} = 0$. The second equation is the curved-space-time version of Euler's

equation. In the first term, we recognize $Du^\alpha/d\tau = u^\beta \nabla_\beta u^\alpha$ as the covariant acceleration of a selected fluid element, which would be zero if the fluid element were moving on a geodesic of the curved space-time. This, however, is prevented by the pressure forces acting within the fluid; the fluid element is not moving freely in the gravitational field.

3.3 Einstein Field Equations

The Einstein field equations relate the curvature of space-time to the distribution of matter within space-time. They read

$$G^{\alpha\beta} = \frac{8\pi G}{c^4} T^{\alpha\beta}, \quad (129)$$

with the Einstein curvature tensor of Eq. (118) on the left-hand side, and the total energy-momentum tensor of all forms of matter and non-gravitational fields on the right-hand side. Taking into account the symmetries of the Einstein and energy-momentum tensors, the Einstein field equations are a set of ten second-order, partial differential equations for the metric tensor $g_{\alpha\beta}$. The equations are all coupled, and they are highly nonlinear in the metric and its first derivatives; they are, however, linear in the second derivatives of the metric tensor.

A naive counting of the number of equations might suggest that given suitable initial and boundary conditions for the metric, the solution to the Einstein field equations should be unique. This suggestion, however, is false, as the freedom to perform coordinate transformations must be retained; two metrics $g_{\alpha\beta}$ and $g'_{\mu\nu}$ related to each other by a coordinate transformation should both be valid solutions to the field equations. This freedom is guaranteed by the contracted Bianchi identities,

$$\nabla_\beta G^{\alpha\beta} = 0, \quad (130)$$

which reveal that of the ten field equations, only six are truly independent from each other. The Bianchi identities, together with the field equations, are compatible with the local statement of energy-momentum conservation

$$\nabla_\beta T^{\alpha\beta} = 0. \quad (131)$$

The equations of general relativity can be derived from an action principle. The total action S involves a gravitational piece given by the Hilbert action

$$S_{\text{grav}} = \frac{c^3}{16\pi G} \int R dV, \quad (132)$$

where R is the Ricci scalar and $dV = \sqrt{-g} d^4x$ is the invariant volume element, as well as a matter piece given by

$$S_M = \int \mathcal{L}_M dV, \quad (133)$$

where $\mathcal{L}_{\mathcal{M}}$ is the Lagrangian density for all the matter (and field) variables. The Einstein tensor results from the functional variation of the gravitational action:

$$\frac{\delta S_{\text{grav}}}{\delta g_{\alpha\beta}} = -\frac{c^3}{16\pi G} G^{\alpha\beta}. \quad (134)$$

The energy-momentum tensor, on the other hand, is defined by

$$\frac{\delta S_M}{\delta g_{\alpha\beta}} = \frac{1}{2c} T^{\alpha\beta}. \quad (135)$$

The Einstein field equations then follow from the requirement that $\delta(S_{\text{grav}} + S_M) = 0$ under an arbitrary variation of the metric tensor.

4 Post-Minkowskian and Post-Newtonian Theory

In this section, we specialize the formalism of general relativity to a description of weak gravitational fields, such as those inside and near the Sun, those inside and near white dwarfs, and those near (but not too near) neutron stars or black holes. This approximation should, of course, reproduce the predictions of Newtonian theory, but we wish to go beyond this and formulate a method of approximation that can be pushed systematically to higher and higher order and generate increasingly accurate descriptions of a weak gravitational field. The foundation for this approach is “post-Minkowskian theory.”

4.1 Landau–Lifshitz Formulation of the Field Equations

The post-Minkowskian approach is based on the Landau and Lifshitz formulation of the Einstein equations. In this framework, the main variables are not the components of the metric tensor $g_{\alpha\beta}$, but those of the “gothic inverse metric density”

$$\mathfrak{g}^{\alpha\beta} := \sqrt{-g} g^{\alpha\beta}, \quad (136)$$

where $g^{\alpha\beta}$ is the inverse metric and g the metric determinant. Knowledge of the gothic metric is sufficient to determine the metric itself: note first that $\det[\mathfrak{g}^{\alpha\beta}] = g$, so that g can be directly obtained from the gothic metric; then Eq. (136) gives $g^{\alpha\beta}$, which can be inverted to give $g_{\alpha\beta}$.

In the Landau–Lifshitz formulation, the left-hand side of the field equations is built from

$$H^{\alpha\mu\beta\nu} := \mathfrak{g}^{\alpha\beta} \mathfrak{g}^{\mu\nu} - \mathfrak{g}^{\alpha\nu} \mathfrak{g}^{\beta\mu}. \quad (137)$$

This tensor density is readily seen to possess the same symmetries as the Riemann tensor,

$$H^{\mu\alpha\beta\nu} = -H^{\alpha\mu\beta\nu}, \quad H^{\alpha\mu\nu\beta} = -H^{\alpha\mu\beta\nu}, \quad H^{\beta\nu\alpha\mu} = H^{\alpha\mu\beta\nu}. \quad (138)$$

It turns out that $H^{\mu\alpha\beta\nu}$ satisfies the remarkable identity

$$\partial_{\mu\nu} H^{\alpha\mu\beta\nu} = (-g) \left(2G^{\alpha\beta} + \frac{16\pi G}{c^4} t_{\text{LL}}^{\alpha\beta} \right), \quad (139)$$

where $G^{\alpha\beta}$ is the Einstein tensor, and

$$\begin{aligned} (-g)t_{\text{LL}}^{\alpha\beta} := & \frac{c^4}{16\pi G} \left\{ \partial_\lambda g^{\alpha\beta} \partial_\mu g^{\lambda\mu} - \partial_\lambda g^{\alpha\lambda} \partial_\mu g^{\beta\mu} + \frac{1}{2} g^{\alpha\beta} g_{\lambda\mu} \partial_\rho g^{\lambda\nu} \partial_\nu g^{\mu\rho} \right. \\ & - g^{\alpha\lambda} g_{\mu\nu} \partial_\rho g^{\beta\nu} \partial_\lambda g^{\mu\rho} - g^{\beta\lambda} g_{\mu\nu} \partial_\rho g^{\alpha\nu} \partial_\lambda g^{\mu\rho} + g_{\lambda\mu} g^{\nu\rho} \partial_\nu g^{\alpha\lambda} \partial_\rho g^{\beta\mu} \\ & \left. + \frac{1}{8} (2g^{\alpha\lambda} g^{\beta\mu} - g^{\alpha\beta} g^{\lambda\mu}) (2g_{\nu\rho} g_{\sigma\tau} - g_{\rho\sigma} g_{\nu\tau}) \partial_\lambda g^{\nu\tau} \partial_\mu g^{\rho\sigma} \right\} \end{aligned} \quad (140)$$

is the *Landau–Lifshitz pseudotensor*, which (very loosely speaking) represents the distribution of gravitational-field energy. It is called a pseudotensor because it contains ordinary derivatives of the metric, and does not by itself transform as a tensor under general coordinate transformations. The full Eq. (139), of course, is generally covariant, although it is not written in a manifestly covariant form. Equation (139) is valid for any space-time.

We now impose Einstein's equations, $G^{\alpha\beta} = 8\pi G T^{\alpha\beta}/c^4$, to obtain

$$\partial_{\mu\nu} H^{\alpha\mu\beta\nu} = \frac{16\pi G}{c^4} (-g) (T^{\alpha\beta} + t_{\text{LL}}^{\alpha\beta}), \quad (141)$$

which is therefore equivalent to Einstein's equations, combining as it does Einstein's equations and an identity.

By virtue of the antisymmetry of $H^{\alpha\mu\beta\nu}$ in the last pair of indices, we have that the equation

$$\partial_{\beta\mu\nu} H^{\alpha\mu\beta\nu} = 0 \quad (142)$$

holds as a trivial identity. This, together with Eq. (141), implies that

$$\partial_\beta \left[(-g) (T^{\alpha\beta} + t_{\text{LL}}^{\alpha\beta}) \right] = 0. \quad (143)$$

These are conservation equations for the total energy-momentum pseudotensor (which includes a contribution from the matter and another contribution from the gravitational field), expressed in terms of a partial-derivative operator. These equations are strictly equivalent to the usual expression of energy-momentum conservation, $\nabla_\beta T^{\alpha\beta} = 0$, which involves only the matter's energy-momentum tensor and a covariant-derivative operator.

This is an *exact* reformulation of the standard form of the theory. No approximations are involved, and no restrictions are placed on the choice of coordinates. It has

to be acknowledged, however, that the usefulness of the formalism is largely limited to situations in which (i) the coordinates $x^\alpha = (ct, x^j)$ are modest deformations of the Lorentzian coordinates of flat space-time, and (ii) $g^{\alpha\beta}$ deviates only moderately from the Minkowski metric $\eta^{\alpha\beta}$.

Because they involve a partial-derivative operator, the differential identities of Eq. (143) can immediately be turned into integral identities. We consider a three-dimensional region V , a fixed (time-independent) domain of the spatial coordinates x^j , bounded by a two-dimensional surface S . We assume that V contains at least some of the matter (so that $T^{\alpha\beta}$ is nonzero somewhere within V), but that S does not intersect any of the matter (so that $T^{\alpha\beta} = 0$ everywhere on S).

We formally define a total momentum four-vector $P^\alpha[V]$ associated with the region V by the three-dimensional integral

$$P^\alpha[V] := \frac{1}{c} \int_V (-g)(T^{\alpha 0} + t_{\text{LL}}^{\alpha 0}) d^3x. \quad (144)$$

This total momentum includes a contribution from the matter's momentum density $c^{-1}T^{\alpha 0}$, and a contribution from the gravitational field represented by $c^{-1}t_{\text{LL}}^{\alpha 0}$.

The momentum four-vector can be broken down into a time component $P^0[V]$ and a spatial three-vector $P^j[V]$. The time component can be used to define an energy $E[V] := cP^0[V]$ associated with the region V ,

$$E[V] := \int_V (-g)(T^{00} + t_{\text{LL}}^{00}) d^3x, \quad (145)$$

and a corresponding mass $M[V] := E[V]/c^2$. The three-momentum is given by

$$P^j[V] := \frac{1}{c} \int_V (-g)(T^{j0} + t_{\text{LL}}^{j0}) d^3x. \quad (146)$$

The total angular-momentum four-tensor $J^{\alpha\beta}[V]$ is defined by

$$J^{\alpha\beta}[V] := \frac{1}{c} \int_V \left[x^\alpha (-g)(T^{\beta 0} + t_{\text{LL}}^{\beta 0}) - x^\beta (-g)(T^{\alpha 0} + t_{\text{LL}}^{\alpha 0}) \right] d^3x, \quad (147)$$

$$:= \frac{2}{c} \int_V x^{[\alpha} (-g)(T^{\beta]0} + t_{\text{LL}}^{\beta]0}) d^3x, \quad (148)$$

and we note that this tensor is antisymmetric in its indices. The interpretation of $J^{\alpha\beta}[V]$ is easier to extract once it is decomposed into time and spatial components. The antisymmetry of the tensor implies that $J^{00}[V] = 0$. The time-space components can be expressed in the form

$$c^{-1}J^{0j}[V] = P^j[V]t - M[V]R^j[V], \quad (149)$$

where

$$R^j[V] := \frac{1}{M[V]c^2} \int_V (-g)(T^{00} + t_{\text{LL}}^{00})x^j d^3x \quad (150)$$

represents the position of the center-of-mass of the region V . Equation (149) reveals that, if $c^{-1}J^{0j}[V]$ is constant, it fixes the location of the center-of-mass at $t = 0$. The spatial components of the angular-momentum tensor are

$$J^{jk}[V] = \frac{1}{c} \int_V \left[x^j (-g)(T^{k0} + t_{\text{LL}}^{k0}) - x^k (-g)(T^{j0} + t_{\text{LL}}^{j0}) \right] d^3x, \quad (151)$$

and this is best recognized in its equivalent angular-momentum vectorial form as

$$J_j[V] := \frac{1}{2} \varepsilon_{jpq} J^{pq}[V] = \frac{1}{c} \int_V \varepsilon_{jpq} x^p (-g)(T^{q0} + t_{\text{LL}}^{q0}) d^3x, \quad (152)$$

where ε_{jpq} is the completely antisymmetric permutation symbol.

To obtain the conservation statements satisfied by $P^\alpha[V]$, we differentiate its defining expression with respect to x^0 and use the local conservation identity of Eq. (143). Starting with Eq. (144), we get

$$\frac{d}{dx^0} P^\alpha[V] = \frac{1}{c} \int_V \partial_0 \left[(-g)(T^{\alpha 0} + t_{\text{LL}}^{\alpha 0}) \right] d^3x = -\frac{1}{c} \int_V \partial_k \left[(-g)(T^{\alpha k} + t_{\text{LL}}^{\alpha k}) \right] d^3x.$$

Converting this to a surface integral, and recalling our previous assumption that S does not intersect the matter distribution, so that $T^{\alpha\beta} = 0$ on S , we arrive at

$$\dot{P}^\alpha[V] = - \oint_S (-g) t_{\text{LL}}^{\alpha k} dS_k, \quad (153)$$

in which an overdot indicates differentiation with respect to $t := x^0/c$. The rate of change of $P^\alpha[V]$ is therefore expressed as a flux integral over S , and the flux is measured by the Landau–Lifshitz pseudotensor.

Equation (153) gives rise to the individual statement about the total energy,

$$\dot{E}[V] = -c \oint_S (-g) t_{\text{LL}}^{0k} dS_k. \quad (154)$$

Similar conservation statements can be derived for linear momentum, angular momentum, and the center-of-mass.

In the limit in which V is taken to include all of three-dimensional space, $P^\alpha[V]$ is known to coincide with the Arnowitt–Deser–Misner four-momentum of an asymptotically flat space-time, and its physical interpretation as a measure of total momentum is robust. This statement is true whenever the coordinates x^α coincide with a Lorentzian system at infinity; the coordinates do not have to be Lorentzian (and indeed, they could not be) at finite spatial distances.

4.2 The Relaxed Einstein Equations

It is advantageous at this stage to impose the four conditions

$$\partial_\beta g^{\alpha\beta} = 0 \quad (155)$$

on the gothic inverse metric. These are known as the *harmonic coordinate conditions*. It is also useful to introduce the potentials

$$h^{\alpha\beta} := \eta^{\alpha\beta} - \mathfrak{g}^{\alpha\beta}, \quad (156)$$

where $\eta^{\alpha\beta} := \text{diag}(-1, 1, 1, 1)$ is the Minkowski metric expressed in Lorentzian coordinates. In terms of these potentials, the harmonic coordinate conditions read

$$\partial_\beta h^{\alpha\beta} = 0, \quad (157)$$

and in this context they are usually referred to as the *harmonic gauge conditions*.

The introduction of the potentials $h^{\alpha\beta}$ and the imposition of the harmonic gauge conditions simplify the appearance of the Einstein field equations. It is easy to verify that the left-hand side becomes

$$\partial_{\mu\nu} H^{\alpha\mu\beta\nu} = -\square h^{\alpha\beta} + h^{\mu\nu} \partial_{\mu\nu} h^{\alpha\beta} - \partial_\mu h^{\alpha\nu} \partial_\nu h^{\beta\mu}, \quad (158)$$

where $\square := \eta^{\mu\nu} \partial_{\mu\nu}$ is the flat-space-time wave operator. Isolating the wave operator on the left-hand side, and putting everything else on the right-hand side gives us the formal wave equation

$$\square h^{\alpha\beta} = -\frac{16\pi G}{c^4} \tau^{\alpha\beta} \quad (159)$$

for the potentials $h^{\alpha\beta}$, where

$$\tau^{\alpha\beta} := (-g)(T^{\alpha\beta}[\mathfrak{m}, g] + t_{\text{LL}}^{\alpha\beta}[h] + t_{\text{H}}^{\alpha\beta}[h]) \quad (160)$$

is the *effective energy-momentum pseudotensor* for the wave equation. We have introduced

$$(-g)t_{\text{H}}^{\alpha\beta} := \frac{c^4}{16\pi G} \left(\partial_\mu h^{\alpha\nu} \partial_\nu h^{\beta\mu} - h^{\mu\nu} \partial_{\mu\nu} h^{\alpha\beta} \right) \quad (161)$$

as an additional (harmonic-gauge) contribution to the effective energy-momentum pseudotensor. The harmonic conditions slightly simplify the form of the Landau–Lifshitz pseudotensor, as can be seen from Eq. (140).

In our expression for $\tau^{\alpha\beta}$, we have indicated that the matter's energy-momentum tensor $T^{\alpha\beta}$ is a functional of matter variables \mathfrak{m} , in addition to being a functional of the metric tensor $g_{\alpha\beta}$ (which is obtained from the gravitational potentials). As an example, when the matter consists of a number of isolated fluid bodies, \mathfrak{m} collectively denotes fluid variables such as the mass density ρ , pressure p , and velocity field u^α . We have also indicated that the Landau–Lifshitz and harmonic pseudotensors are functionals of $h^{\alpha\beta}$.

The imposition of the gauge conditions (157) is equivalent to imposing the conservation equations

$$\partial_\beta \tau^{\alpha\beta} = 0. \quad (162)$$

It is easy to verify that $(-g)t_{\text{H}}^{\alpha\beta}$ is separately conserved, in that it satisfies the equation $\partial_\beta [(-g)t_{\text{H}}^{\alpha\beta}] = 0$ automatically.

The wave equation of Eq. (159) is the main starting point of post-Minkowskian theory. It is worth emphasizing the fact that this equation, together with Eq. (157) or (162) are an *exact formulation* of the Einstein field equations; no approximations have been introduced at this stage.

For a metric $g_{\alpha\beta}$ to satisfy the complete set of Einstein field equations, it is necessary for the potentials $h^{\alpha\beta}$ to satisfy *both* the wave equation *and* the gauge condition/conservation statement; it is the *union* of Eqs. (159) and (157) or (162) that is equivalent to the original form of the Einstein field equations, $G^{\alpha\beta} = (8\pi G/c^4)T^{\alpha\beta}$. The two sets of equations play different roles, however. The wave equation (159) determines the gravitational potentials $h^{\alpha\beta}[\mathbf{m}]$ (and therefore the metric) as functions of the harmonic coordinates x^α , in terms of the matter variables \mathbf{m} ; these, however, remain undetermined until we also invoke the conservation Eq. (162). It is this equation that determines the behavior of the matter variables in a curved space-time whose metric is extracted from $h^{\alpha\beta}[\mathbf{m}]$. Solving both sets of equations amounts to integrating the full set of Einstein field equations, and this procedure simultaneously determines the metric *and* the matter variables.

One is entirely free to solve the wave equation (159) without also enforcing the gauge condition of Eq. (157) or the conservation statement of Eq. (162). Solving the wave equation independently of the gauge condition/conservation statement amounts to integrating only a subset of the Einstein field equations, and the outcome of this procedure is ten gravitational potentials $h^{\alpha\beta}[\mathbf{m}]$ that are expressed as functionals of undetermined matter variables \mathbf{m} . The metric obtained from these potentials is also a functional of \mathbf{m} , and it is not yet a solution to the Einstein field equations; it becomes a solution only when the gauge condition/conservation statement is imposed as an additional condition on the matter variables. The wave equation (Eq. (159)), taken by itself independently of Eq. (157), is known as the *relaxed Einstein field equation*.

4.3 Solution of the Wave Equation

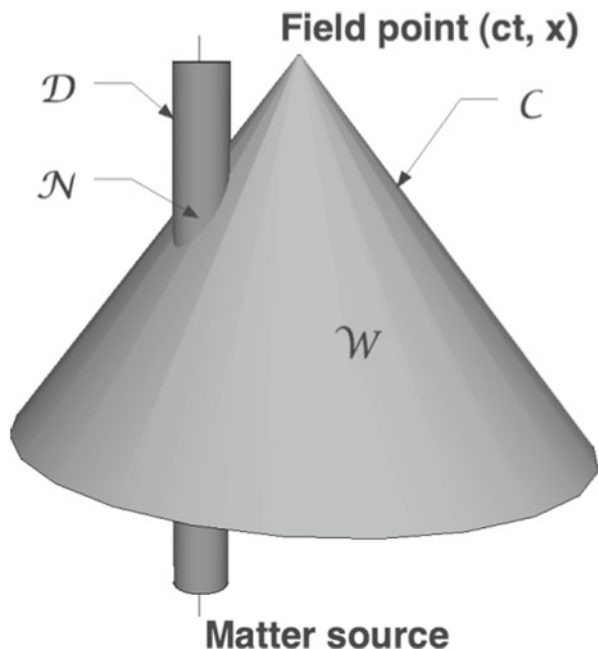
The wave equation (159) admits the immediate formal solution

$$h^{\alpha\beta}(t, \mathbf{x}) = \frac{4G}{c^4} \int \frac{\tau^{\alpha\beta}(t - |\mathbf{x} - \mathbf{x}'|/c, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'. \quad (163)$$

This is the *retarded solution* to the wave equation, and the domain of integration extends over $\mathcal{C}(x)$, the *past light cone* of the field point $x = (ct, \mathbf{x})$ (see Fig. 2).

To evaluate the integral, we partition the integration domain into two pieces, the *near-zone domain* $\mathcal{N}(x)$ and the *wave-zone domain* $\mathcal{W}(x)$. We place the boundary of the near/wave zones at an arbitrarily selected radius \mathcal{R} , with \mathcal{R} imagined to be of the same order of magnitude as $\lambda_c \sim c t_c$, the characteristic wavelength of the radiation emitted by the source, where t_c is the characteristic time scale for noticeable changes to occur within the source. For slowly moving sources, $\lambda_c \sim cr_c/v_c \gg r_c$, where r_c and v_c are the characteristic size and internal velocity of the material source.

Fig. 2 Integration domains for the retarded solution of the wave equation



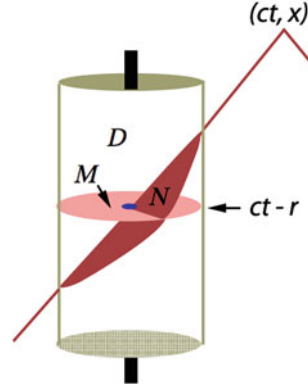
Thus, the near zone is the region of three-dimensional space in which $r := |\mathbf{x}|$ is small compared with a characteristic wavelength λ_c , while the wave zone is the region in which r is large compared with this length scale. The potential behaves very differently in the two zones: in the near zone, the difference between $t - r/c$ and t is small (the field retardation is unimportant), and derivatives with respect to $x^0 = ct$ are small compared with spatial derivatives; in the wave zone, the difference between $t - r/c$ and t is large, and x^0 -derivatives are comparable to spatial derivatives. These properties are shared by all generic solutions to the wave equation.

In space-time, this sphere of radius \mathcal{R} traces a near-zone *world tube*, \mathcal{D} . We let $\mathcal{N}(x)$ be the intersection between $\mathcal{C}(x)$ and the near zone \mathcal{D} , formally defined as the spatial region such that $r' := |\mathbf{x}'| < \mathcal{R}$. Similarly, we let $\mathcal{W}(x)$ be the intersection between $\mathcal{C}(x)$ and the wave zone, formally defined as the spatial region such that $r' > \mathcal{R}$, or $\mathcal{C}(x) - \mathcal{N}(x)$. We write Eq. (163) as

$$h^{\alpha\beta}(x) = h_{\mathcal{N}}^{\alpha\beta}(x) + h_{\mathcal{W}}^{\alpha\beta}(x). \quad (164)$$

In electromagnetism, there is generally no contribution from the wave zone integration, because the source—the charges and currents—are confined to the near zone. But because of the nonlinearity of gravity, gravity itself acts as a source, so that, even though $T^{\alpha\beta}$ may be confined to the near zone, $t_{\text{LL}}^{\alpha\beta}$ and $t_{\text{H}}^{\alpha\beta}$ are not. Special integration techniques must be used to carry out the integrals over \mathcal{W} , but for most physical systems of interest, the wave-zone integrals generate higher-order corrections to the

Fig. 3 Near-zone integration, far-zone field point



dominant terms, which come from the near-zone integrals. In this chapter, we will henceforth concern ourselves only with calculating $h_{\mathcal{N}}^{\alpha\beta}(x)$.

We first evaluate $h_{\mathcal{N}}^{\alpha\beta}(x)$ when the field point x is situated in the wave zone, that is, when $r \gg \mathcal{R}$. Knowing that \mathbf{x}' lies within the near zone, we treat it as a small vector, and we Taylor-expand the variable $|\mathbf{x} - \mathbf{x}'|$ in $h_{\mathcal{N}}^{\alpha\beta}(x)$ about $\mathbf{x}' = \mathbf{0}$, with the result

$$h_{\mathcal{N}}^{\alpha\beta}(t, \mathbf{x}) = \frac{4G}{c^4} \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{\ell!} \partial_L \left[\frac{1}{r} \int_{\mathcal{M}} \tau^{\alpha\beta}(\tau, \mathbf{x}') x'^L d^3 x' \right], \quad (165)$$

where

$$\tau := t - r/c \quad (166)$$

is a retarded-time variable. Notice that the temporal dependence of the source function no longer involves \mathbf{x}' , the variable of integration. The integration domain has therefore become a surface of constant time (the constant being equal to $\tau = t - r/c$) bounded externally by the sphere $r' = \mathcal{R}$. This domain is denoted by \mathcal{M} in Eq. (165), and is illustrated in Fig. 3. The partial derivatives ∂_L operate both on $1/r$ and on the r embedded in τ .

Equation (165) is valid everywhere within the wave zone. It simplifies when $r \rightarrow \infty$, that is, when $h_{\mathcal{N}}^{\alpha\beta}$ is evaluated in the *far-away wave zone*. In this limit, we retain only the dominant, r^{-1} term in $h_{\mathcal{N}}^{\alpha\beta}$, and we approximate Eq. (165) by

$$h_{\mathcal{N}}^{\alpha\beta} = \frac{4G}{rc^4} \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{\ell!} \int_{\mathcal{M}} \partial_L \tau^{\alpha\beta}(\tau, \mathbf{x}') x'^L d^3 x' + O(r^{-2}). \quad (167)$$

The dependence of $\tau^{\alpha\beta}$ on x^j is contained in $\tau = t - r/c$, so, using the fact for any function $f(\tau)$ that $\partial_j f = -c^{-1} \partial_j r (df/d\tau) = -c^{-1} n_j (df/d\tau)$, we find that Eq. (167) becomes

$$h_{\mathcal{N}}^{\alpha\beta}(t, \mathbf{x}) = \frac{4G}{rc^4} \sum_{\ell=0}^{\infty} \frac{1}{\ell! c^\ell} n_L \left(\frac{d}{d\tau} \right)^\ell \int_{\mathcal{M}} \tau^{\alpha\beta}(\tau, \mathbf{x}') x'^L d^3 x' + O(r^{-2}). \quad (168)$$

This is a *multipole expansion* for the potential $h_{\mathcal{N}}^{\alpha\beta}$, in which each ℓ -pole moment $\int_{\mathcal{M}} \tau^{\alpha\beta} x^L d^3x$ is differentiated ℓ -times with respect to τ . Notice that $n_L x'^L = n_{j_1} n_{j_2} \cdots n_{j_\ell} x'^{j_1} x'^{j_2} \cdots x'^{j_\ell} = (\mathbf{n} \cdot \mathbf{x}')^\ell$.

We next evaluate $h_{\mathcal{N}}^{\alpha\beta}(x)$ when x is situated in the near zone, that is, when $r = |\mathbf{x}| < \mathcal{R}$. In this situation, both \mathbf{x} and \mathbf{x}' lie within the near zone, and $|\mathbf{x} - \mathbf{x}'|$ can be treated as a small quantity. To evaluate the integral, we simply Taylor-expand the time-dependence of the source with the result

$$h_{\mathcal{N}}^{\alpha\beta}(t, \mathbf{x}) = \frac{4G}{c^4} \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{\ell! c^\ell} \left(\frac{\partial}{\partial t} \right)^\ell \int_{\mathcal{M}} \tau^{\alpha\beta}(t, \mathbf{x}') |\mathbf{x} - \mathbf{x}'|^{\ell-1} d^3x', \quad (169)$$

which is valid everywhere within the near zone. Notice that once more, the domain of integration is \mathcal{M} , a surface of constant time bounded externally by the sphere $r' = \mathcal{R}$, but here evaluated at the unretarded time t .

4.4 Iteration of the Relaxed Field Equations

We have cast the solution of the wave equation (Eq. (159)) as a retarded integral, but we really cannot call it a “solution”, because $h^{\alpha\beta}$ appears inside the integral. In order to construct solutions for a particular choice of matter variables, we proceed by successive approximations, or iterations.

In the *zeroth iteration* of the relaxed field equations, we set $h_0^{\alpha\beta} = 0$ and immediately get $g_{\alpha\beta}^0 = \eta_{\alpha\beta}$, the metric of Minkowski space-time. From this, we construct $T^{\alpha\beta}[\mathbf{m}, g] = T^{\alpha\beta}[\mathbf{m}, \eta]$, $t_{\text{LL}}^{\alpha\beta}[h] = t_{\text{LL}}^{\alpha\beta}[h_0] = 0$, and $t_{\text{H}}^{\alpha\beta}[h] = t_{\text{H}}^{\alpha\beta}[h_0] = 0$. From all this, we obtain $\tau_0^{\alpha\beta} = T^{\alpha\beta}[\mathbf{m}, \eta]$; this is the energy-momentum tensor of the matter variables \mathbf{m} , and in the zeroth iteration these live in Minkowski space-time.

In the *first iteration* of the relaxed field equations, we solve the wave equation $\square h^{\alpha\beta} = -(16\pi G/c^4) \tau_0^{\alpha\beta}$ for $h_1^{\alpha\beta}$. Because the source $\tau_0^{\alpha\beta}$ is known from the zeroth iteration, the wave equation can be integrated without difficulty (at least in principle), and this returns the potentials $h_1^{\alpha\beta}$ as functionals of the matter variables \mathbf{m} , which are yet to be determined. From the potentials, we form the metric $g_{\alpha\beta}^1$ and construct $\tau_1^{\alpha\beta}$, an improved version of the effective energy-momentum pseudotensor. This involves the material contribution $T^{\alpha\beta}[\mathbf{m}, g_1]$, as well as the field contributions $t_{\text{LL}}^{\alpha\beta}[h_1]$ and $t_{\text{H}}^{\alpha\beta}[h_1]$.

In the *second iteration* of the relaxed field equations, we solve the wave equation $\square h^{\alpha\beta} = -(16\pi G/c^4) \tau_1^{\alpha\beta}$ for $h_2^{\alpha\beta}$, an improved version of the gravitational potentials. Because the source $\tau_1^{\alpha\beta}$ is known from the first iteration, the wave equation can once more be integrated, and $h_2^{\alpha\beta}$ are again functionals of the undetermined matter variables \mathbf{m} . From the new potentials, we form the metric $g_{\alpha\beta}^2$ and construct $\tau_2^{\alpha\beta}$, the latest version of the effective energy-momentum pseudotensor. The stage is ready

for the next iteration. One continues iterating until the desired precision is reached for determining the equations of motion for the matter and the resulting space-time metric.

5 Post-Newtonian Theory

Post-Newtonian theory can be considered as a special case of post-Minkowski theory in which one (a) imposes a slow-motion condition on top of the weak-field condition and (b) focusses on the near zone. Because the main physical applications of interest will be gravitationally self-interacting systems, such as binary systems, we know from Newton's equations that $U \sim v^2$, and from the equations of structure that $p/\rho \sim U$. In units of c^2 , these are all small quantities for weak fields, and thus we assign a dimensionless parameter ε to characterize this smallness, so that

$$\varepsilon \sim \frac{U}{c^2} \sim \frac{v^2}{c^2} \sim \frac{p}{\rho c^2}. \quad (170)$$

In addition, since, in the near zone, quantities vary by virtue of their motion, one can also argue that a partial time derivative $\partial/\partial t$ is of order v compared to a spatial gradient, and thus that $\partial/\partial x^0 \sim (v/c)\nabla$, or

$$\frac{\partial/\partial x^0}{\partial/\partial x^j} \sim \varepsilon^{1/2}. \quad (171)$$

This approximation modifies to a certain extent how one treats the different iterations of the relaxed Einstein equations. For example, in the first iteration, with $\tau^{00} = T^{00}[\mathbf{m}, \eta]$, we find

$$h_1^{00} = \frac{4G}{c^4} \int \frac{T^{00}[\mathbf{m}, \eta]}{|\mathbf{x} - \mathbf{x}'|} d^3x'. \quad (172)$$

But the special-relativistic $T^{00} \sim \rho c^2[1 + O(v/c)^2]$, and hence h_1^{00} has the schematic form

$$h_1^{00} \sim 4U/c^2 + O(v^2U/c^4) \sim \varepsilon + \varepsilon^2. \quad (173)$$

At the second iteration, contributions of $h_1^{\alpha\beta}$ to τ^{00} will induce contributions of order

$$h_2^{00} \sim \frac{4G}{c^4} \int \frac{T^{00}[\mathbf{m}, \eta] h_1^{00}}{|\mathbf{x} - \mathbf{x}'|} d^3x' \sim \frac{G^2}{c^4} \left(\frac{m}{r}\right)^2 \sim \frac{U^2}{c^4} \sim \varepsilon^2, \quad (174)$$

which are of the same order as the $O(v/c)^2$ corrections to the first iteration. Thus, to obtain the metric to the first *post-Newtonian* order, one must iterate the relaxed equations twice, as well as retain suitable $(v/c)^2$ correction terms. We will make the requirements more explicit shortly.

5.1 General Structure of the Fields

It is instructive to examine the general structure of the fields $h_{\mathcal{N}}^{\alpha\beta}$ in the near zone. The first few terms in the expansion look like

$$h_{\mathcal{N}}^{\alpha\beta}(t, \mathbf{x}) = \frac{4G}{c^4} \left[\int_{\mathcal{M}} \frac{\tau^{\alpha\beta}(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' - \frac{1}{c} \frac{\partial}{\partial t} \int_{\mathcal{M}} \tau^{\alpha\beta}(t, \mathbf{x}') d^3x' \right. \\ \left. + \frac{1}{2c^2} \frac{\partial^2}{\partial t^2} \int_{\mathcal{M}} \tau^{\alpha\beta}(t, \mathbf{x}') |\mathbf{x} - \mathbf{x}'| d^3x' - \dots \right]. \quad (175)$$

Taking into account the conservation equation $\partial_\beta \tau^{\alpha\beta} = 0$, we can convert certain terms into surface integrals that either vanish in the absence of gravitational radiation, or that can be shown to contribute very small, higher-order corrections. To simplify the notation somewhat, it is useful to define $N := h^{00}$, $K^i := h^{0i}$, $B^{ij} := h^{ij}$, and $B := \delta_{ij} h^{ij}$. Then

$$N_{\mathcal{N}} = \frac{4G}{c^2} \left[\int_{\mathcal{M}} \frac{c^{-2} \tau^{00}(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \frac{1}{2c^2} \frac{\partial^2}{\partial t^2} \int_{\mathcal{M}} c^{-2} \tau^{00}(t, \mathbf{x}') |\mathbf{x} - \mathbf{x}'| d^3x' \right. \\ - \frac{1}{6c^3} \mathcal{I}^{kk(3)}(t) + \frac{1}{24c^4} \frac{\partial^4}{\partial t^4} \int_{\mathcal{M}} c^{-2} \tau^{00}(t, \mathbf{x}') |\mathbf{x} - \mathbf{x}'|^3 d^3x' \\ - \frac{1}{120c^5} \left\{ (4x^{kl} + 2r^2 \delta^{kl}) \mathcal{I}^{kl(5)}(t) - 4x^k \mathcal{I}^{kll(5)}(t) + \mathcal{I}^{kkll(5)}(t) \right\} \\ \left. + h_{\partial\mathcal{M}}^{00} + O(c^{-6}) \right], \quad (176a)$$

$$K_{\mathcal{N}}^i = \frac{4G}{c^3} \left[\int_{\mathcal{M}} \frac{c^{-1} \tau^{0i}(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \frac{1}{2c^2} \frac{\partial^2}{\partial t^2} \int_{\mathcal{M}} c^{-1} \tau^{0i}(t, \mathbf{x}') |\mathbf{x} - \mathbf{x}'| d^3x' \right. \\ + \frac{1}{18c^3} \left\{ 3x^k \mathcal{I}^{ik(4)}(t) - \mathcal{I}^{ikk(4)}(t) + 2\epsilon^{mik} \mathcal{J}^{mk(3)}(t) \right\} \\ \left. + h_{\partial\mathcal{M}}^{0i} + O(c^{-4}) \right], \quad (176b)$$

$$B_{\mathcal{N}}^{ij} = \frac{4G}{c^4} \left[\int_{\mathcal{M}} \frac{\tau^{ij}(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' - \frac{1}{2c} \mathcal{I}^{ij(3)}(t) + \frac{1}{2c^2} \frac{\partial^2}{\partial t^2} \int_{\mathcal{M}} \tau^{ij}(t, \mathbf{x}') |\mathbf{x} - \mathbf{x}'| d^3x' \right. \\ - \frac{1}{36c^3} \left\{ 3r^2 \mathcal{I}^{ij(5)}(t) - 2x^k \mathcal{I}^{ijk(5)}(t) - 8x^k \epsilon^{mk(i} \mathcal{J}^{m|j)(4)}(t) + 6 M^{ijkk(3)}(t) \right\} \\ \left. + h_{\partial\mathcal{M}}^{ij} + O(c^{-4}) \right], \quad (176c)$$

where

$$\mathcal{I}^Q \equiv \int_{\mathcal{M}} c^{-2} \tau^{00} x^Q d^3x, \quad (177a)$$

$$\mathcal{J}^i Q \equiv \varepsilon^{iab} \int_{\mathcal{M}} c^{-1} \tau^{0b} x^a Q d^3x, \quad (177b)$$

$$M^{ijQ} \equiv \int_{\mathcal{M}} \tau^{ij} x^Q d^3x, \quad (177c)$$

and where $h_{\partial\mathcal{M}}^{\alpha\beta}$ denote the surface terms, which will not be displayed here. The notation (n) over quantities such as \mathcal{I}^{jk} denotes the number of time derivatives. In obtaining these forms, we have made good use of a number of identities that follow from the conservation statement $\partial_\beta \tau^{\alpha\beta} = 0$, namely,

$$\tau^{0j} = \partial_0(\tau^{00} x^j) + \partial_k(\tau^{0k} x^j), \quad (178a)$$

$$\tau^{jk} = \frac{1}{2} \partial_{00}(\tau^{00} x^j x^k) + \frac{1}{2} \partial_p(\tau^{jp} x^k + \tau^{kp} x^j - \partial_q \tau^{pq} x^j x^k), \quad (178b)$$

$$\begin{aligned} \tau^{jk} x^n &= \frac{1}{2} \partial_0(\tau^{0j} x^k x^n + \tau^{0k} x^j x^n - \tau^{0n} x^j x^k) \\ &\quad + \frac{1}{2} \partial_p(\tau^{jp} x^k x^n + \tau^{kp} x^j x^n - \tau^{np} x^j x^k). \end{aligned} \quad (178c)$$

The first term in N clearly leads off with the Newtonian potential/ c^2 , of order ε . Embedded within that term are corrections of order ε and higher, both from $(v/c)^2$ corrections and from contributions from higher iterations of the relaxed equations. The second term in N is known as a superpotential; because of the two time derivatives, it clearly leads off at order $(v/c)^2$ compared to the Newtonian term, and is thus already a post-Newtonian correction (and higher). (The superpotential itself is of order $mc^2 r$, but since $\partial/\partial t \sim v/r$, the result is of order $(v/c)^2$ compared to the first term.) Notice that there was no term involving a single time derivative. By virtue of the fact that $\partial_0 \tau^{00} = -\partial_k \tau^{0k}$, that term in fact is converted purely to a surface term within $h_{\partial\mathcal{M}}^{00}$, and contributes only at a very high post-Newtonian order via a flux of gravitational radiation.

The third term in N contains three time derivatives, and one might be tempted to give it the name of a (post) $^{3/2}$ -Newtonian correction. However, it can be shown that, because this term depends only on time, it can always be absorbed to lowest order into a redefinition of time, and hence removed by a coordinate transformation. Alternatively, since the leading contribution of N to the equations of motion arises from a gradient, this term will make no contribution to leading order.

The fourth term is called a superduperpotential term, and because of the four time derivatives, contributes at second post-Newtonian (2PN) order and higher.

The fifth term looks schematically like $mc^2(\partial/\partial ct)^5 r^4 \sim c^2(v/c)^5(m/r) \sim \varepsilon^5 \times$ Term 1. This is a 2.5PN correction, and represents the leading effects of gravitational

radiation reaction. Because it involves an odd number of time derivatives, it is asymmetric under the transformation $t \rightarrow -t$; as such, it represents the effects of an irreversible, dissipative process, the loss of energy into gravitational waves.

The other field components have a similar structure: K^i leads at order $\varepsilon^{3/2}$, while B^{ij} and B lead at order ε^2 .

We now need to establish the order in ε to which each of these components must be calculated to reach a desired post-Newtonian order. To do this, we return to the expression for the action for a test particle moving in the space-time of a metric $g_{\alpha\beta}$, Eq. (108), and express it in terms of separated time and space components, namely

$$S = -mc^2 \int_1^2 \sqrt{-g_{00} - 2g_{0j} \frac{v^j}{c} - g_{ij} \frac{v^j v^k}{c^2}} dt. \quad (179)$$

Newtonian gravity is clearly given by the approximations $g_{00} \approx -1 + 2U/c^2$, $g_{0i} \approx 0$, and $g_{ij} \approx \delta_{ij}$, in other words, by evaluating the action to order ε . Post-Newtonian gravity is then given by evaluating the action to order ε^2 , which requires evaluating the metric coefficients to the following orders of approximation:

$$\begin{aligned} g_{00} &\text{ to } O(\varepsilon^2), \\ g_{0i} &\text{ to } O(\varepsilon^{3/2}), \\ g_{ij} &\text{ to } O(\varepsilon). \end{aligned} \quad (180)$$

The orders needed to descend because of the additional factors of v/c in the action. From this, we can generalize to the requirement that to work to nPN order, we must determine the metric components to the orders

$$\begin{aligned} g_{00} &\text{ to } O(\varepsilon^{n+1}), \\ g_{0i} &\text{ to } O(\varepsilon^{n+\frac{1}{2}}), \\ g_{ij} &\text{ to } O(\varepsilon^n). \end{aligned} \quad (181)$$

To determine the order to which we must then determine the components of the field, we must expand in powers of ε the relation between the physical metric $g^{\alpha\beta}$ and the fields using Eq. (136). The result is

$$\begin{aligned} g_{00} = & -(1 - \frac{1}{2}N + \frac{3}{8}N^2 - \frac{5}{16}N^3) + \frac{1}{2}B(1 - \frac{1}{2}N) + \frac{1}{2}K^j K^j \\ & + O(\varepsilon^4), \end{aligned} \quad (182a)$$

$$g_{0i} = -K^i(1 - \frac{1}{2}N) + O(\varepsilon^{7/2}), \quad (182b)$$

$$g_{ij} = \delta^{ij}(1 + \frac{1}{2}N - \frac{1}{8}N^2) + B^{ij} - \frac{1}{2}B\delta^{ij} + O(\varepsilon^3), \quad (182c)$$

$$(-g) = 1 + N - B + O(\varepsilon^3). \quad (182d)$$

This expansion is sufficient to determine the metric through 2.5PN order.

5.2 The Post-Newtonian Limit of General Relativity

To illustrate the use of the post-Minkowskian framework, we will derive the post-Newtonian limit of general relativity. In going to post-Newtonian order, we wish to express the density in terms of ρ^* instead of ρ , because the former density satisfies a continuity equation to all orders. This makes a number of calculations simpler.

The Newtonian limit is trivial: inserting the zeroth iteration of the relaxed equations, $h_0^{\alpha\beta} = 0$ in the source terms, it is clear that $\tau_0^{00} = \rho^* c^2$, and thus that the first iteration of the equations yields $N_1 = 4U/c^2$, and $g_{00} = -1 + 2U/c^2$, $g_{0i} = 0$, and $g_{ij} = \delta_{ij}$, where U is the “Newtonian” potential derived from $\nabla^2 U = -4\pi G \rho^*$. From the first iteration, we also obtain expressions for K^i and B^{ij} , but these are not needed for the Newtonian limit of $g_{\alpha\beta}$; we will return to these shortly. Notice that in determining N , we also get the bonus that $g_{ij} = \delta_{ij}(1 + 2U/c^2)$; this also is not needed for the Newtonian limit, but will be needed in order to obtain $(-g)$ for the next iteration.

Using the definition (125), together with the approximations $u^0/c \approx 1 + v^2/2c^2 + U/c^2$, and $-g \approx 1 + 4U/c^2$ we find to the order needed for the next iteration that

$$\rho^* = \rho \left(1 + \frac{v^2}{2c^2} + \frac{3U}{c^2} \right). \quad (183)$$

Substituting these results into Eq. (121), and carefully evaluating the contributions of $t_{LL}^{\alpha\beta}$ and $t_H^{\alpha\beta}$ needed to this order, we obtain the first iterated version of $\tau^{\alpha\beta}$, given to the required post-Newtonian order by

$$\tau_1^{00} = \rho^* c^2 \left(1 + \frac{v^2}{2c^2} + \frac{3U}{c^2} + \frac{\Pi}{c^2} \right) - \frac{7}{8\pi G} \nabla U \cdot \nabla U, \quad (184a)$$

$$\tau_1^{0i} = \rho^* c v^i, \quad (184b)$$

$$\tau_1^{ii} = \rho^* v^2 + 3p - \frac{1}{8\pi G} \nabla U \cdot \nabla U, \quad (184c)$$

where $\Pi := \varepsilon/\rho^*$ is the internal energy per unit mass within the fluid.

Recalling the identity

$$\nabla U \cdot \nabla U = \frac{1}{2} \nabla^2 U^2 + 4\pi G \rho^* U, \quad (185)$$

we substitute Eqs. (184) into (176) and keep all contributions needed for the metric through 1.5PN order. The result is

$$N_2 = \frac{4U}{c^2} + \frac{1}{c^4} \left(7U^2 + 2\Phi_1 - 2\Phi_2 + 4\Phi_3 + 2\ddot{X} \right) - \frac{2G}{3c^5} \mathcal{I}^{kk(3)}(t) + O(\varepsilon^3), \quad (186a)$$

$$K_2^i = \frac{4U^i}{c^3} + O(\varepsilon^{5/2}), \quad (186b)$$

$$B_2 = \frac{1}{c^4} \left(U^2 + 4\Phi_1 - 2\Phi_2 + 12\Phi_4 \right) - \frac{2G}{c^5} \mathcal{I}^{kk(3)}(t) + O(\varepsilon^3), \quad (186c)$$

where we define the auxiliary gravitational potentials

$$\begin{aligned} U^i &:= G \int \frac{\rho^*(t, \mathbf{x}') v^i(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x', \\ \Phi_1 &:= G \int \frac{\rho^*(t, \mathbf{x}') v'^2}{|\mathbf{x} - \mathbf{x}'|} d^3x', \\ \Phi_2 &:= G \int \frac{\rho^*(t, \mathbf{x}') U(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x', \\ \Phi_3 &:= G \int \frac{\rho^*(t, \mathbf{x}') \Pi(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x', \\ \Phi_4 &:= G \int \frac{p(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x', \\ X &:= G \int \rho^*(t, \mathbf{x}') |\mathbf{x} - \mathbf{x}'| d^3x'. \end{aligned} \quad (187)$$

Substituting these results into Eq. (182), we obtain finally the metric for general relativity through 1.5PN order,

$$g_{00} = -1 + \frac{2}{c^2} U + \frac{2}{c^4} (\Psi - U^2) - \frac{4G}{3c^5} \mathcal{I}^{kk(3)}(t) + O(\varepsilon^3), \quad (188a)$$

$$g_{0i} = -\frac{4}{c^3} U^i + O(\varepsilon^{5/2}), \quad (188b)$$

$$g_{ij} = \delta_{ij} \left(1 + \frac{2}{c^2} U \right) + O(\varepsilon^2), \quad (188c)$$

where

$$\Psi = \frac{3}{2} \Phi_1 - \Phi_2 + \Phi_3 + 3\Phi_4 + \frac{1}{2} \ddot{X}. \quad (189)$$

By making a coordinate transformation

$$t = t' - \frac{2G}{3c^5} \ddot{X}^{kk}(t), \quad (190)$$

we can eliminate the $O(G/c^5)$ term from g_{00} to lowest order, indicating that the term is purely a gauge or coordinate artifact. The leading nontrivial odd-order terms in g_{00} will occur at $O(\varepsilon^{7/2})$, corresponding to 2.5PN order.

Calculating the Christoffel symbols from the metric (188), inserting them into the energy-momentum conservation equation $\nabla_\beta T^{\alpha\beta} = 0$, one obtains the post-Newtonian Euler equation of hydrodynamics,

$$\rho^* \frac{dv^j}{dt} = \rho^* \partial_j U - \partial_j p$$

$$\begin{aligned}
& + \frac{1}{c^2} \left[\left(\frac{1}{2} v^2 + U + \Pi + \frac{p}{\rho^*} \right) \partial_j p - v^j \partial_t p \right] \\
& + \frac{1}{c^2} \rho^* \left[(v^2 - 4U) \partial_j U - v^j (3 \partial_t U + 4 v^k \partial_k U) \right. \\
& \quad \left. + 4 \partial_t U_j + 4 v^k (\partial_k U_j - \partial_j U_k) + \partial_j \Psi \right] \\
& + O(c^{-4}).
\end{aligned} \tag{191}$$

The post-Newtonian Euler equation, together with an equation of state $p(\rho)$ can be solved consistently, leading to a solution for the space-time metric, valid to post-Newtonian order.

5.3 Non-Newtonian Behavior: Worked Examples

Using the post-Newtonian metric derived above, we can work out some important examples of the effects of relativistic gravity. For simplicity, we will specialize to a single, spherically symmetric body of mass M at rest in our chosen coordinate system. This is a good approximation for studying the motion of a planet or a ray of light in the solar system. In this case, the metric of Eq. (188) reduces to

$$g_{00} = -1 + 2 \frac{GM}{c^2 r} - 2 \left(\frac{GM}{c^2 r} \right)^2, \tag{192a}$$

$$g_{0i} = 0, \tag{192b}$$

$$g_{ij} = \delta_{ij} \left(1 + 2 \frac{GM}{c^2 r} \right), \tag{192c}$$

where

$$\begin{aligned}
M &:= \int \rho^* \left[1 + \frac{1}{c^2} \left(\frac{3}{2} v^2 - U + \Pi + 3 \frac{p}{\rho^*} \right) \right] d^3x \\
&= m + \frac{1}{c^2} [3\mathcal{T} + 2\Omega + E_{\text{int}} + 3P] \\
&= m + \frac{1}{c^2} [\mathcal{T} + \Omega + E_{\text{int}}],
\end{aligned} \tag{193}$$

where $m = \int \rho^* d^3x$ is the total conserved or baryonic mass of the body, and \mathcal{T} , Ω and E_{int} are the internal kinetic, gravitational, and thermal energy of the body, defined as in Eq. (18) (but using ρ^* instead of the Newtonian ρ). We have used the Newtonian virial theorem (17b) in the static limit to simplify the post-Newtonian corrections to the mass. Note that the final answer, not surprisingly, is the total of rest mass and the mass equivalent of the internal energies.

The geodesic equation (107) can be rewritten using coordinate time as the parameter; this is a useful step because the motion of planets or light is normally referred

to an external “time” coordinate, such as atomic time measured on Earth, not proper time along the particle’s world line. The result is

$$\frac{d^2 r^i}{dt^2} + \left(\Gamma_{\alpha\beta}^i - \Gamma_{\alpha\beta}^0 \frac{v^i}{c} \right) v^\alpha v^\beta = 0, \quad (194)$$

where $v^\alpha := (c, v^i)$. Evaluating the Christoffel symbols to post-Newtonian order using Eq. (192), we obtain the equation of motion for a test body

$$\frac{d^2 \mathbf{r}}{dt^2} = -\frac{GM}{r^2} \mathbf{n} + \frac{GM}{c^2 r^2} \left\{ \mathbf{n} \left[4 \frac{GM}{r} - v^2 \right] + 4 \dot{\mathbf{r}} \right\}, \quad (195)$$

where $\dot{\mathbf{r}} = \mathbf{n} \cdot \mathbf{v}$.

This equation of motion is ready-to-use for analysis of perturbed orbit elements, since it is in the form of the Newtonian two-body acceleration plus a disturbing acceleration. The relevant components of the disturbing acceleration (Eq. (63)) are given by

$$\begin{aligned} \mathcal{R} &= \frac{GM}{c^2 r^2} \left[4 \frac{GM}{r} - v^2 + 4 \dot{\mathbf{r}}^2 \right], \\ \mathcal{S} &= 4 \frac{GM}{c^2 r^2} \dot{\mathbf{r}} \mathbf{v} \cdot \boldsymbol{\lambda}, \\ \mathcal{W} &= 0. \end{aligned} \quad (196)$$

From the orbit element perturbation equations (67) we immediately conclude that the inclination ι and angle of nodes Ω are constants, consistent with the motion staying in a fixed plane. We can choose this plane to be the $X - Y$ plane to simplify matters. We now substitute the osculating orbit formulae for the variables that appear in \mathcal{R} and \mathcal{S} , such as $r = p/(1 + e \cos f)$, obtained from Eq. (62). In the perturbation equations (67), we convert from d/dt to d/df using the relation $df/dt = (h/r^2)$, valid to first order in perturbation theory (see Eq. (68)). Integrating over one orbit, we find the net change in the orbit elements,

$$\Delta a = 0, \quad (197a)$$

$$\Delta e = 0, \quad (197b)$$

$$\Delta \varpi = \frac{6\pi GM}{c^2 p}. \quad (197c)$$

The first two results express the fact that energy and angular momentum are conserved to post-Newtonian order. The second is the famous advance of the pericenter.

Einstein derived this for a test particle in a geodesic around a point object of mass M , using an approximate solution of his equations equivalent to Eq. (192). The same result, to 1PN order is obtained from the Schwarzschild geometry. It turns out that it also applies to an arbitrary binary system of masses m_1 and m_2 ; in this case, $M = m_1 + m_2$. Since the mass of Mercury is only one 6 millionth of that of

the Sun, its contribution is negligible, and it can be treated as a massless test particle moving on a geodesic of a fixed space-time. This cannot be done for a binary star system because both masses move about the common center of mass; nevertheless, the answer for $\Delta\varpi$ is the same, and depends only on the total mass.

The advance per orbit can be converted to a rate of advance by dividing by the orbital period. We can also eliminate the semimajor axis appearing in p by using Kepler's third law, $a = (GM)^{1/3}(P/2\pi)^{2/3}$, where P is the orbital period, valid here because the effect is already of 1PN order. The result is

$$\begin{aligned} \frac{d\varpi}{dt} &= \frac{3}{1-e^2} \frac{2\pi}{P} \left(\frac{GM/c^3}{P/2\pi} \right)^{2/3} \\ &= 716.25 \frac{1}{1-e^2} \left(\frac{M}{M_\odot} \right)^{2/3} \left(\frac{P}{1 \text{ day}} \right)^{-5/3} \text{ as yr}^{-1}. \end{aligned} \quad (198)$$

Note that GM/c^3 is the light-crossing time of half the Schwarzschild radius corresponding to mass M , which is generally much smaller than the orbital period. Substituting the values for Mercury, $e = 0.2056$, $P = 87.97$ days, we obtain 42.98 as per century; as we saw in Table 1, the modern difference between the measured advance and that predicted by Newtonian N -body perturbations is 42.98 ± 0.04 as per century.

The discovery of binary pulsar systems with total masses of order 2–3 solar masses, and with orbital periods as small as fractions of a day resulted in the observation of huge periastron advances of several degrees per year. But in these systems, the relativistic periastron advance played an interesting role. In the solar system, GM for the Sun is known to high precision from the measured orbital period and orbital radius of the Earth, combined with Kepler's third law. By contrast, the masses of the neutron stars in these binary systems are not known, apart from the general expectation that they should be around the Chandrasekhar limit, $1.4 M_\odot$, based on models of how such systems might have formed. In the famous Hulse–Taylor binary pulsar 1913+16, the first such system to be discovered, it was possible to measure the pericenter advance, the eccentricity and the orbital period very accurately; the current values are $\dot{\omega} = 4.226595(5) \text{ deg yr}^{-1}$, $P = 0.322997448930(4) \text{ day}$, $e = 0.6171338(4)$, where the number in parentheses denotes the error in the final digit. Assuming that there is no other source of periastron advance, we can turn Eq. (198) around and use it to measure the total mass of the system. The result is $M = 2.828296(5) M_\odot$. This is a remarkably accurate measurement of an astrophysical quantity, and general relativity played a central role in the analysis. The recently discovered “double pulsar” J0737-3039A/B, in which both stars were observed as pulsars, has $e = 0.0877775(9)$, $P = 0.10225156248(5) \text{ day}$, and $\dot{\omega} = 16.8995(7) \text{ deg yr}^{-1}$, so the total mass is $2.5871(2) M_\odot$. The relativistic advance is so large in these systems that we could easily see the evolving orientation of the orbit month by month with the naked eye, that is if we could see pulsars with the naked eye! For further details on binary pulsars, see the chapter, “The Role of Binary Pulsars in Testing Gravity Theories” by A. Possenti and M. Burgay.

The motion of light near the sun can also be analyzed, either by considering the equation for a null geodesic, or equivalently by taking the $v^i \rightarrow cN^i$ limit of Eq. (195). The result is

$$\frac{d^2 \mathbf{r}}{dt^2} = -\frac{GM}{r^2} [\mathbf{n} - 2\mathbf{N}(\mathbf{N} \cdot \mathbf{n})] . \quad (199)$$

This equation can be integrated by setting \mathbf{N} equal to the unperturbed direction of the light ray to obtain the deflection of a light ray that passes close to the sun, given by

$$\delta\theta = \frac{4GM_\odot}{c^2 d} = 1.7504 \frac{R_\odot}{d} \text{ as} , \quad (200)$$

where d is the distance of closest approach of the ray.

The first successful measurement of the bending of light by the Sun was carried out by British astronomer Arthur Stanley Eddington and colleagues during the total solar eclipse of May 29, 1919. It was a differential measurement: photographs of the stars near the Sun taken during the eclipse were compared with photographs of the same stars taken at night from the same locations later in the year, and the *changes* in angles between pairs of stars were carefully measured. Eddington's announcement in November 1919 of the measurement of the bending helped make Einstein an international celebrity. However, the experiments of Eddington and his coworkers had only 30 % accuracy, and succeeding eclipse measurements were not much better. The results were scattered between one half and twice the Einstein prediction, and the accuracies were low.

However, the advent of radio astronomy, and in particular the development of radio interferometry, and later of very-long-baseline radio interferometry (VLBI), produced greatly improved determinations of the deflection. These techniques now have the capability of measuring angular separations and changes in angles to accuracies of 10–100 microarcseconds. Early measurements took advantage of the fact that a few strong quasistellar radio sources (quasars) pass very close to the Sun (as seen from the Earth), including the group 3C273, 3C279, and 3C48, and the group 0111+02, 0119+11, and 0116+08. As the Earth moves in its orbit, changing the lines of sight of the quasars relative to the Sun, the angular separation $\delta\theta$ between pairs of quasars varies. A number of measurements of this kind over the period 1969–1975 yielded results in agreement with general relativity at a few parts in 10^3 . In recent years, transcontinental and intercontinental VLBI observations of quasars and radio galaxies have been made primarily to monitor the Earth's rotation and to establish a highly accurate “reference” frame for astronomy and navigation. These measurements are sensitive to the deflection of light over almost the entire celestial sphere. A 2004 analysis of almost 2 million VLBI observations of 541 radio sources, made by 87 VLBI sites over a 20-year period verified Einstein's prediction to a few parts in 10^4 . Analysis of observations made by the Hipparcos optical astrometry satellite yielded a test at the level of 0.3 %, and a future orbiting observatory called GAIA has the capability of testing the deflection to parts per million.

The deflection of light has become a cornerstone of the empirical edifice that supports general relativity. But in 1979, the phenomenon became much more than

that. That year, astronomers Dennis Walsh, Robert Carswell, and Ray Weymann discovered the “double quasar” Q0957+561, which consisted of two quasar images about 6 as apart, with almost the same redshift ($z = 1.41$) and very similar spectra. Given that quasars are thought to be among the most distant objects in the universe, the probability of finding two so close together was low. It was soon realized that there was just one quasar but that intervening matter in the form of a galaxy or cluster of galaxies was bending the light from the quasar and producing two separate images.

Since then, over 60 lensed quasars have been discovered. But more importantly, gravitational lensing has become a major tool in efforts to map the distribution of mass around galaxies and clusters, in the search for dark matter, dark energy, and compact objects, and in the search for extrasolar planets. Major subtopics in lensing have been developed to cover different astronomical realms: microlensing for the search for dim compact objects and extrasolar planets, the use of luminous arcs in the effort to map the distribution of mass and dark matter, and weak/statistical lensing in the effort to measure the properties of dark energy. Lensing has to be taken into account in interpreting certain aspects of the cosmic microwave background radiation, and in extracting information from the gravitational-wave signal from sources at large redshift.

Equation (199) can also be used to derive the Shapiro time delay, a retardation of the propagation of light past a massive object. For light emitted at time t_e at \mathbf{x}_e and received by an observer at t_o at \mathbf{x}_o , the propagation time is given by

$$c(t_o - t_e) = |\mathbf{x}_o - \mathbf{x}_e| + \frac{2GM}{c^2} \ln \left[\frac{(r_o + \mathbf{x}_o \cdot \mathbf{n})(r_e - \mathbf{x}_e \cdot \mathbf{n})}{d^2} \right]. \quad (201)$$

For a ray that propagates in a round-trip from the Earth to a planet or spacecraft on the far side of the Sun, the total delay can be as large as 200 μs .

Radio astronomer Irwin I. Shapiro discovered that this effect was a prediction of general relativity in 1964, and called it the “fourth” test of relativity. Shapiro and colleagues then carried out the first measurement of the time delay in 1967, by bouncing radar signals off the surface of Mercury. Later experiments involved radar echos from Venus and tracking of the Mars exploration spacecraft, Mariners 6, 7, and 9, and the Viking landers and orbiters.

The most recent measurement of the Shapiro delay involved tracking the Cassini spacecraft while it was on its way to Saturn. Several circumstances made this mission particularly favorable. One was the ability to do tracking measurements using both X-band (7175 MHz) and Ka-band (34316 MHz) radar, thereby significantly reducing the dispersive effects of the solar corona. In addition, during the 2002 superior conjunction of Cassini, the spacecraft was at 8.43 astronomical units from the Sun, and the distance of closest approach of the radar signals to the Sun was only $1.6 R_\odot$. The result was a test in agreement with general relativity to two parts in 10^5 . This and other solar system tests are discussed in the chapter, “Space-based Tests of Relativistic Gravitation” by V. Turyshev.

The Shapiro delay now figures in range of astrophysical phenomena. It has been measured in a number of binary pulsar systems, most notably in the double pulsar

J0737-3039A/B, where the orbit is seen almost edge-on, and the pulsed radio signals from each pulsar pass close to the other once per orbit. Analyses of the spectra and time variations of high-energy emissions (X-rays) from accretion disks around black holes must take into account both the strong bending of light and the Shapiro delay, but now using full general relativity, not just the post-Newtonian limit.

5.4 The Parametrized Post-Newtonian Formalism

An important application of the post-Newtonian limit of general relativity has been its extension to encompass a range of alternative metric theories of gravity. In many cases, such as the scalar-tensor theory of Jordan, Brans, and Dicke and generalizations thereof, the post-Newtonian limit has a form quite similar to that of general relativity, except that the coefficients in front of some terms are different, and a few additional post-Newtonian potentials may be present.

This is a very fortunate circumstance, because the post-Newtonian limit is sufficient to describe the solar system and the experimental tests one can perform there, and to a limited degree can also describe binary pulsar systems. Therefore, if we simply replace the numerical coefficients in the post-Newtonian limit of general relativity with arbitrary parameters, and add a few new potentials with their own parameters, we will have a framework that encompasses a wide range of alternative theories, and that can be used to calculate a wide range of testable phenomena.

This framework is called the parametrized post-Newtonian (PPN) framework. Eddington was the first to consider such a framework. In his classic 1922 textbook, *The Mathematical Theory of Relativity*, he parametrized the post-Newtonian limit of the Schwarzschild metric (in isotropic coordinates) in the form

$$ds^2 = - \left[1 - 2\alpha \frac{GM}{c^2 r} + 2\beta \left(\frac{GM}{c^2 r} \right)^2 \right] c^2 dt^2 + \left[1 + 2\gamma \frac{GM}{c^2 r} \right] (dx^2 + dy^2 + dz^2), \quad (202)$$

where $\alpha = \beta = \gamma = 1$ in general relativity. Actually, the α parameter is redundant, since it can always be absorbed into GM , which is by definition the Kepler-measured mass, or into Newton's constant G as measured in a Cavendish-type experiment. The modern version of the PPN framework was developed by Kenneth Nordtvedt, Jr. in 1968, using a collection of N point masses as the physical system rather than a single mass; later, Will generalized the framework to post-Newtonian hydrodynamics, and in 1972 Nordtvedt and Will unified the two approaches into the PPN framework used today.

In a gauge that parallels the harmonic gauge of general relativity, the PPN metric has the form

$$g_{00} = -1 + \frac{2}{c^2} U + \frac{2}{c^4} (\Psi - \beta U^2) + O(\varepsilon^3), \quad (203a)$$

$$g_{0i} = -\frac{1}{c^3} \left[2(1 + \gamma) + \frac{1}{2}\alpha_1 \right] U^i + O(\varepsilon^{5/2}), \quad (203b)$$

$$g_{ij} = \delta_{ij} \left(1 + \frac{2}{c^2} \gamma U \right) + O(\varepsilon^2), \quad (203c)$$

where

$$\begin{aligned} \Psi := & \frac{1}{2}(2\gamma + 1 + \alpha_3)\Phi_1 - (2\beta - 1 - \zeta_2 - \xi)\Phi_2 + (1 + \zeta_3)\Phi_3 \\ & + [3(\gamma + \zeta_4) - \zeta_1]\Phi_4 + \frac{1}{2}(\zeta_1 - 2\xi)\Phi_5 - \xi\Phi_W + \frac{1}{2}(1 + \alpha_2)\ddot{X}. \end{aligned} \quad (204)$$

The potentials

$$\begin{aligned} \Phi_5 &:= G \int \rho^* \nabla' U(t, \mathbf{x}') \cdot \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|} d^3x', \\ \Phi_W &:= G^2 \int \int \rho^* \rho^{*''} \frac{(x - x')_j}{|\mathbf{x} - \mathbf{x}'|^3} \left(\frac{(x' - x'')^j}{|\mathbf{x} - \mathbf{x}''|} - \frac{(x - x'')^j}{|\mathbf{x}' - \mathbf{x}''|} \right) d^3x' d^3x'', \end{aligned} \quad (205)$$

do not appear in the post-Newtonian limit of general relativity in this gauge.

The choice of parameters is made in part so that the values in general relativity are particularly simple, but more importantly so that specific meanings or implications can be attached to them. These heuristic meanings are summarized in Table 2. The parameters γ and β are directly related to Eddington's original parameters. Roughly speaking, γ measures the amount of curvature of space generated by a body, compared to what general relativity would predict; specifically, a calculation of the Riemann curvature tensor of the three-dimensional subspace defined by $dt = 0$ gives a result proportional to $\gamma GM/c^2 r^3$. Roughly speaking, β measures how nonlinear gravity is, in that it multiplies the U^2 term in g_{00} . This is *very* rough, because it is a coordinate-dependent statement. In general relativity for example, the Schwarzschild metric has no $(GM/c^2 r)^2$ term in Schwarzschild coordinates; the post-Newtonian limit of the Schwarzschild geometry takes the Eddington form of Eq. (202) with $\gamma = \beta = 1$ *only* in isotropic coordinates. The lesson is that these interpretations of the PPN parameters must be considered heuristic at best.

The parameters α_i are linked to violations of local Lorentz invariance in gravitational physics, which some theories predict. Suppose that all three parameters vanish in a given theory. Then the PPN metric in the above form is valid in any chosen coordinate system, with all velocities defined relative to that system. However, if any of them is nonzero, then the theory in question singles out a preferred universal rest frame (in most cases, it is the frame in which the cosmic background radiation is isotropic), and the PPN metric above is valid *only* when written in that rest coordinate system. To obtain the metric in a different frame, such as the barycentric frame of the solar system, one must perform what is called a post-Galilean transformation of the metric: this is a low-velocity Lorentz transformation evaluated to the appropriate post-Newtonian order, together with a post-Newtonian gauge transformation to put the final metric in a simple form. Such a transformation generates additional terms in the metric that depend *explicitly* on the velocity of the coordinate system relative to the preferred frame, and on the PPN parameters α_i . Theories of

Table 2 The PPN parameters and their significance (note that α_3 has been shown twice to indicate that it is a measure of two effects)

Parameter	What it measures relative to GR	GR value	Semiconservative theories	Fully conservative theories
γ	How much space-curvature produced by unit rest mass?	1	γ	γ
β	How much “nonlinearity” in the superposition of gravity?	1	β	β
ξ	Preferred-location effects?	0	ξ	ξ
α_1	Preferred-frame effects?	0	α_1	0
α_2		0	α_2	0
α_3		0	0	0
α_3	Is total momentum conserved?	0	0	0
ζ_1		0	0	0
ζ_2		0	0	0
ζ_3		0	0	0
ζ_4		0	0	0

PPN parametrized post-Newtonian, GR general relativity

this type predict *preferred-frame* effects, many of which have been tested, leading to strong upper bounds on these parameters. The ζ_i parameters indicate whether the theory admits conservation laws for energy and momentum of the kind defined by Eq. (144); any theory that is based on an action predicts that these parameters vanish.

A number of well-known relativistic effects can now be expressed in terms of these PPN parameters (compare with Eqs. (198), (200), and (201)):

Perihelion advance:

$$\begin{aligned}
 \frac{d\varpi}{dt} &= \left(\frac{2 + 2\gamma - \beta}{3} \right) \frac{3}{1 - e^2} \frac{2\pi}{P} \left(\frac{GM/c^3}{P/2\pi} \right)^{2/3} \\
 &= \left(\frac{2 + 2\gamma - \beta}{3} \right) \times 42.98 \text{ arcsec}/100 \text{ yr} ,
 \end{aligned} \tag{206}$$

where P and e are the period and eccentricity of the orbit; the second line is the value for Mercury.

Deflection of light:

$$\begin{aligned}
 \Delta\theta &= \left(\frac{1 + \gamma}{2} \right) \frac{4GM}{c^2 d} \\
 &= \left(\frac{1 + \gamma}{2} \right) \times 1.7504 \frac{R_\odot}{d} \text{ as} .
 \end{aligned} \tag{207}$$

Shapiro time delay:

$$c(t_O - t_e) = |\mathbf{x}_O - \mathbf{x}_e| + \left(\frac{1 + \gamma}{2} \right) \frac{2GM}{c^2} \ln \left[\frac{(r_O + \mathbf{x}_O \cdot \mathbf{n})(r_e - \mathbf{x}_e \cdot \mathbf{n})}{d^2} \right]. \quad (208)$$

Nordtvedt effect:

$$\frac{m_G - m_I}{m_I} = \left(4\beta - \gamma - 3 - \frac{10}{3}\xi - \alpha_1 - \frac{2}{3}\alpha_2 - \frac{2}{3}\zeta_1 - \frac{1}{3}\zeta_2 \right) \frac{|E_g|}{m_I c^2}, \quad (209)$$

where m_G and m_I are the gravitational and inertial masses of a body such as the Earth or Moon, and E_g is its gravitational binding energy. A nonzero Nordtvedt effect would cause the Earth and Moon to fall with a different acceleration toward the Sun. This effect does not occur in general relativity, a result of the *Strong Equivalence Principle*, satisfied by that theory. The chapter “Probing Gravity with Next Generation Lunar Laser Ranging” by S. Dell’Agnello et al., discusses how lunar laser ranging has set a strong bound on this effect.

Precession of a gyroscope:

$$\begin{aligned} \Omega_{FD} &= -\frac{1}{2} \left(1 + \gamma + \frac{\alpha_1}{4} \right) \frac{G}{r^3 c^2} [\mathbf{J} - 3\mathbf{n}(\mathbf{n} \cdot \mathbf{J})] \\ &= \frac{1}{2} \left(1 + \gamma + \frac{\alpha_1}{4} \right) \times 0.041 \text{ arcsec yr}^{-1}, \\ \Omega_{Geo} &= -\frac{1}{2} (1 + 2\gamma) \mathbf{v} \times \frac{Gm\mathbf{n}}{r^2 c^2} \\ &= \frac{1}{3} (1 + 2\gamma) \times 6.6 \text{ arcsec yr}^{-1}, \end{aligned} \quad (210)$$

where Ω_{FD} and Ω_{Geo} are the precession angular velocities caused by the dragging of inertial frames (Lense–Thirring effect), and by the geodetic effect, a combination of Thomas precession and precession induced by spatial curvature; \mathbf{J} is the angular momentum of the Earth, and \mathbf{v} , \mathbf{n} , and r are the velocity, direction, and distance of the gyroscope. The second line in each case is the corresponding value for a gyroscope in polar Earth orbit at about 650 km altitude.

A wide range of experiments have led to tight bounds on the PPN parameters, all consistent with general relativity. The current values are listed in Table 3.

Gravity Probe B

On May 4, 2011, NASA announced the long-awaited results of Gravity Probe B (GP-B). Over 47 years and 750 million \$ in the making, GP-B was an orbiting physics experiment, designed to measure the frame-dragging and geodetic effects of Eq. (210). Josef Lense and Hans Thirring first pointed out the existence of the frame-dragging phenomenon in 1918, but it was not until the 1960s that George Pugh in the Defense Department and Leonard Schiff at Stanford independently pursued the idea of measuring it with gyroscopes.

Table 3 Current limits on the PPN parameters

Parameter	Effect	Limit	Remarks
$\gamma - 1$	(i) Time delay	2.3×10^{-5}	Cassini tracking
	(ii) Light deflection	4×10^{-4}	VLBI
$\beta - 1$	(i) Perihelion shift	3×10^{-3}	Assumes $J_2 = 10^{-7}$ from helioseismology
	(ii) Nordtvedt effect	2.3×10^{-4}	$\eta = 4\beta - \gamma - 3$ assumed
ξ	Earth tides	10^{-3}	Gravimeter data
α_1	Orbital polarization	10^{-4}	Lunar laser ranging
		4×10^{-5}	PSR J1738+0333
α_2	Solar spin precession	4×10^{-7}	Alignment of Sun and ecliptic
α_3	Pulsar acceleration	2×10^{-20}	Pulsar \dot{P} statistics
η^a	Nordtvedt effect	9×10^{-4}	Lunar laser ranging
ζ_1	–	2×10^{-2}	Combined PPN bounds
ζ_2	Binary motion	4×10^{-5}	\ddot{P}_p for PSR 1913+16
ζ_3	Newton’s 3rd law	10^{-8}	Lunar acceleration
ζ_4	–	–	Not independent

^aHere $\eta = 4\beta - \gamma - 3 - 10\xi/3 - \alpha_1 - 2\alpha_2/3 - 2\zeta_1/3 - \zeta_2/3$

PPN parametrized post-Newtonian, VLBI very-long-baseline radio interferometry

GP-B started officially in late 1963 when NASA funded the initial research and development (R&D) work that identified the new technologies needed to make such a difficult measurement possible. Francis Everitt became Principal Investigator of GP-B in 1981, and the project moved to the mission design phase in 1984. Following a major review of the program by a National Academy of Sciences committee in 1994, GP-B was approved for flight development, and began to collaborate with the Lockheed-Martin and Marshall Space Flight Center. The satellite launched on April 20, 2004 for a planned 16-month mission, but another five years of data analysis were needed to tease out the effects of relativity from a background of other disturbances of the gyros¹.

There were four gyroscopes aboard the GP-B satellite, launched into a polar orbit with an altitude of 640 km above the Earth’s surface. Each gyroscope was a fused silica rotor, about the size of a ping-pong ball, machined to be spherical and homogeneous to tolerances better than a part per million, and coated with a thin film of niobium. The gyroscope assembly, which sat in a dewar of 2440 l of superfluid helium, was held at 1.8 K. At this temperature, niobium is a superconductor, and the supercurrents in the niobium of each spinning rotor produce a “London” magnetic moment parallel to its spin axis. Extremely sensitive magnetometers (superconducting quantum interference devices, or SQUIDs) attached to the gyroscope housing

¹ From 1998 to 2011, CMW chaired the NASA Science Advisory Committee for GP-B.

were capable of detecting minute changes in the orientation of the gyros' magnetic moments and hence the precessions predicted by general relativity.

At the start of the mission, the four gyros were aligned to spin along the symmetry axis of the spacecraft. This was also the optical axis of a telescope directly mounted on the end of the structure housing the rotors. Spacecraft thrusters oriented the telescope to point precisely toward the star IM Pegasi (HR 8703) in our galaxy (except when the Earth intervened, once per orbit). In order to average out numerous unwanted torques on the gyros, the spacecraft rotated about its axis once every 78 s.

Almost every aspect of the spacecraft, its subsystems, and the science instrumentation performed extremely well, some far better than expected. Still, success relied on figuring out the sources of error. In particular, having an accurate calibration of the electronic readout from the SQUID magnetometers with respect to the tilt of the gyros was essential. The plan for calibrating the SQUIDs was to exploit the aberration of starlight, which causes a precisely calculable misalignment between the rotors and the telescope as the latter shifts its pointing toward the guide star by up to 20 as to compensate for the orbital motion of the spacecraft and the Earth. However, three important, but unexpected, phenomena were discovered during the experiment that affected the accuracy of the results.

First, because each rotor is not exactly spherical, its principal axis rotates around its spin axis with a period of several hours, with a fixed angle between the two axes. This is the familiar "polhode" motion of a spinning top. In fact, this polhoding was essential in the calibration process because it led to modulations of the SQUID output via the residual trapped magnetic flux on each rotor (about 1 % of the London moment). But the polhode period and angle of each rotor were observed to decrease monotonically with time, implying the presence of some damping mechanism, and this significantly complicated the calibration analysis. In addition, each rotor was found to make occasional, seemingly random "jumps" in its orientation—some as large as 100 milliarcseconds. Some rotors displayed more frequent jumps than others. Without being able to continuously monitor the rotors' orientation, the GP-B team could not fully exploit the calibrating effect of the stellar aberration in their analysis. Finally, during a planned 40-day, end-of-mission calibration phase, the team discovered that when the spacecraft was deliberately pointed away from the guide star by a large angle, the misalignment induced much larger torques on the rotors than expected. From this, they inferred that even the very small misalignments that occurred during the science phase of the mission induced torques that were probably several hundred times larger than the designers had estimated.

What ensued during the data analysis phase following the mission was worthy of a detective novel. The critical clue came from the calibration tests. Here, they took advantage of the residual trapped magnetic flux on the gyroscope. (The designers used superconducting lead shielding to suppress stray fields before they cooled the niobium coated gyroscopes, but no shielding is ever perfect.) This flux adds a periodic modulation to the SQUID output, which the team used to figure out the phase and polhode angle of each rotor throughout the mission. This helped them to conclude that interactions between random patches of electrostatic potential fixed to the

Table 4 Final results of Gravity Probe B

Effect (mas/yr)	Measured	Predicted
Geodetic precession	6602 ± 18	6606
Frame-dragging	37.2 ± 7.2	39.2

surface of each rotor, and similar patches on the inner surface of its spherical housing, were causing the extraneous torques. In principle, the rolling spacecraft should have suppressed these effects, but they were larger than expected.

Fortunately, the patches are fixed on the various surfaces, and so it was possible to build a parametrized model of the patches on both surfaces using multipole expansions, and to calculate the torques induced by those interactions when the spin and spacecraft axes are misaligned, as a function of the parameters. One prediction of the model was that the induced torque should be perpendicular to the plane formed by the two axes, and this was clearly seen in the data. Another prediction of the model was that, when the slowing decreasing polhode period crosses an integer multiple of the spacecraft roll period, the torques fail to average over the roll period, whereupon the spin axis precesses about its initial direction in an opening Cornu spiral, then migrates to a new direction along a closing Cornu spiral. This is known as a loxodromic path, familiar to navigators as a path of fixed bearing on the Earth’s surface. Detailed observation of the orientation of the rotors during such “resonant jumps” showed just such loxodromic behavior. In the end, every jump of every rotor could be identified by its “mode number,” the integer relating its polhode period to the spacecraft roll period.

The original goal of GP-B was to measure the frame-dragging precession with an accuracy of 1 %, but the problems discovered over the course of the mission dashed the initial optimism that this was possible. Although the GP-B team was able to model the effects of the patches, they had to pay the price of the increase in error that comes from using a model with so many parameters. The experiment uncertainty quoted in the final result—roughly 20 % for frame dragging—is almost totally dominated by those errors (see Table 4). Nevertheless, after the model was applied to each rotor, all four gyros showed consistent relativistic precessions. Gyro 2 was particularly “unlucky”—it had the largest uncertainties because it suffered the most resonant jumps. Numerous cross-checks were carried out, including estimating the relativity effect during different segments of the 12-month science phase (various events, including computer reboots and a massive solar storm in January 2005, caused brief interruptions in data-taking), increasing and decreasing the number of parameters in the torque model, and so on.

The most serious competition for the results from GP-B comes from the LAGEOS experiment, in which laser ranging accurately tracked the paths of two laser geodynamics satellites orbiting the Earth. Relativistic frame dragging induces a small precession (around 30 milliarcseconds per year) of the orbital plane of each satellite in the direction of the Earth’s rotation. However, the competing Newtonian effect of the Earth’s nonspherical shape (Eq. (84)) had to be subtracted to very high precision using a model of the Earth’s gravity field. The first published result

from LAGEOS in 1996 quoted an error for the frame-dragging measurement of 20–30 %, though this result was likely too optimistic given the quality of the gravity models available at the time. Later, the GRACE geodesy mission offered dramatically improved Earth gravity models, and the analysis of the LAGEOS satellites has finally yielded tests at a quoted level of approximately 10 %. In the chapter “Fundamental Physics with the LAGEOS Satellites” by R. Peron, the LAGEOS measurements are described in more detail.

5.5 Gravitational Radiation Reaction

We return to general relativity, and focus our attention on the odd-post-Newtonian-order terms in $N_{\mathcal{N}}$, $K_{\mathcal{N}}^i$, and $B_{\mathcal{N}}^{ij}$. We have already argued that the 1.5PN term proportional to $d^3\mathcal{I}^{kk}(t)/dt^3$ in $N_{\mathcal{N}}$ is a pure coordinate artifact to lowest order. However, the next odd-order terms in these potentials are of 2.5PN order and are nontrivial. They are responsible for the loss of energy, momentum, and angular momentum in a dynamical system as the result of the radiation of gravitational waves. After some work, it can be shown that these terms generate an additional radiation-reaction force to the post-Newtonian Euler equation of hydrodynamics (191), given by

$$\begin{aligned} \rho^* \frac{dv^j}{dt} = & \rho^* \partial_j U - \partial_j p + O(c^{-2}) + O(c^{-4}) \\ & + c^{-5} f_{\text{RR}}^j + O(c^{-6}), \end{aligned} \quad (211)$$

where

$$\begin{aligned} f_{\text{RR}}^j = & \frac{G}{c^5} \left[-\rho^* \mathcal{I}^{pq(3)} \partial_{jpq} X + 2(\rho^* \partial_k U - \partial_k p) \mathcal{I}^{jk(3)} + \frac{2}{3} (2\rho^* \partial_j U - \partial_j p) \mathcal{I}^{pp(3)} \right. \\ & \left. + 2\rho^* \mathcal{I}^{jk(4)} v_k + \frac{3}{5} \rho^* \mathcal{I}^{(jk)(5)} x^k - \frac{2}{15} \rho^* \mathcal{I}^{jpp(5)} + \frac{2}{3} \rho^* \varepsilon^{pj q} \mathcal{J}^{pq(4)} \right]. \end{aligned} \quad (212)$$

The $O(c^{-2})$ symbol represents the post-Newtonian terms shown in Eq. (191), and the $O(c^{-4})$ symbol represents 2PN terms not shown here. The radiation reaction force is itself coordinate dependent; one can change the coordinates by 2.5PN-order corrections so that the 1PN and 2PN terms are unaffected, but the radiation-reaction force is transformed into the simple form

$$f_{\text{RR}}^j = -\frac{2G}{5c^5} \rho^* \mathcal{I}^{(jk)(5)} x^k. \quad (213)$$

This is known as the Burke–Thorne gauge.

The Newtonian, 1PN, and 2PN contributions to the equations of motion are *conservative*, in that they enforce the conservation of total energy, momentum, and angular momentum for an isolated system. The energy is given by

$$E = \mathcal{T} + \mathcal{\Omega} + E_{\text{int}} + O(c^{-2}) + O(c^{-4}), \quad (214)$$

where \mathcal{T} , Ω , and E_{int} are defined just as in Eq. (18), but using ρ^* . Now, however, including the radiation reaction force, one finds that

$$\frac{dE}{dt} = \int \rho^* v^j f_{\text{RR}}^j d^3x = -\frac{2G}{5c^5} \mathcal{I}^{(5)(jk)} \int \rho^* x^k v^j d^3x = -\frac{G}{5c^5} \mathcal{I}^{(5)(jk)} \dot{\mathcal{I}}^{jk}. \quad (215)$$

However this last expression can be written as

$$\frac{dE}{dt} = -\frac{G}{5c^5} \mathcal{I}^{(3)(jk)} \mathcal{I}^{(3)(jk)} + \frac{d}{dt} \frac{G}{c^5} (\dots). \quad (216)$$

But because E is only well-defined to even post-Newtonian orders, the time-derivative term on the right-hand side of Eq. (216) can be moved to the left-hand side and absorbed into a modified energy that has been corrected by meaningless 2.5PN order terms. The final result is

$$\frac{dE}{dt} = -\frac{G}{5c^5} \mathcal{I}^{(3)(jk)} \mathcal{I}^{(3)(jk)}. \quad (217)$$

We will see this exactly equals the energy loss as measured in the far zone. Similar arguments can be made for the effects of radiation reaction on linear momentum and angular momentum.

6 Far-zone Fields and Gravitational Radiation

6.1 The Quadrupole Formula

We now turn to the faraway wave-zone regime of our formal solution $h_{\mathcal{N}}^{\alpha\beta}(t, \mathbf{x})$ of the relaxed Einstein's equation, given by Eq. (168). We first notice that this has the general form of $h \sim r^{-1} f(\tau)$, where $\tau = t - r/c$. As a consequence, it is simple to deduce that $\partial_j h = -(n_j/c) \dot{h} + O(r^{-2})$. Combining this fact with the harmonic gauge condition $\partial_\beta h^{\alpha\beta} = 0$, we can show that

$$\begin{aligned} \dot{h}^{0j} &= n_k \dot{h}^{jk} + O(r^{-2}), \\ \dot{h}^{00} &= n_j n_k \dot{h}^{jk} + O(r^{-2}), \end{aligned} \quad (218)$$

and consequently, in the faraway wave zone, a knowledge of h^{jk} is enough to determine the other components of the field, modulo terms that are constant in time. For gravitational waves, of course, this is all one needs.

The leading $\ell = 0$ term in the expression (168) for h^{jk} is given by

$$h_{\mathcal{N}}^{jk} = \frac{4G}{c^4 r} \int_{\mathcal{M}} \tau^{jk}(\tau, \mathbf{x}') d^3x' + O(r^{-2}). \quad (219)$$

Making use of the identity (178b), and discarding the surface term, we can rewrite this in the form

$$h_{\mathcal{N}}^{jk} = \frac{2G}{c^4 r} \ddot{\mathcal{I}}^{jk}(\tau) + O(r^{-2}), \quad (220)$$

where \mathcal{I}^{jk} is given by Eq. (177a), evaluated at the retarded time $\tau = t - r/c$.

This is the famous “quadrupole formula” of general relativity, that gives the gravitational waveform in terms of two time derivatives of the source quadrupole moment. During the middle 1970s, there was considerable debate over the validity of the derivation of this formula, largely because of the fact that the effective source $\tau^{\alpha\beta}$ extends over all space-time. The questions raised included the possibility that higher moments of the effective source $\tau^{\alpha\beta}$ could diverge if the integrals were taken to infinity, or that various surface integrals, such as the one involved in going from Eqs. (219) to (220), could actually fail to vanish. Ultimately, these concerns were laid to rest. One approach involved recalling that the formal solution is actually given by Eq. (164). Since $h_{\mathcal{N}}^{\alpha\beta}$ is an integral over a finite spatial volume—the near zone—it is by definition finite. Furthermore, it could be shown that the integration over the far zone $h_{\mathcal{W}}^{\alpha\beta}$ was likewise finite. In addition, any dependence of each integral on the arbitrarily chosen radius \mathcal{R} of the boundary between the two zones was shown to cancel term-by-term, order-by-order, in the sum $h_{\mathcal{N}}^{\alpha\beta} + h_{\mathcal{W}}^{\alpha\beta}$. Similar arguments were applied to the near-zone expressions for the fields, used in the equations of motion.

The field h^{jk} in the faraway wave zone is a central ingredient in determining the response of a gravitational wave detector. From the equation of geodesic deviation (115), one can show that the relative acceleration of a pair of freely moving masses that are nearly at rest in a local freely falling frame is given by

$$\frac{d^2 \xi^j}{dt^2} = \frac{1}{2} \ddot{h}_{TT}^{jk} \xi_k, \quad (221)$$

where ξ^j is the physically measured displacement between the two bodies, and t is proper time as measured at the origin of the local frame. The subscript “TT” denotes the *transverse-traceless* part, given by

$$h_{TT}^{jk} := h^{lm} \left(P_l^j P_m^k - \frac{1}{2} P^{jk} P_{lm} \right), \quad (222)$$

where P_j^i is a projection tensor, given by

$$P_j^i := \delta_j^i - n^i n_j. \quad (223)$$

Gravitational-wave detection is covered in the chapter “The Detection of Gravitational Waves” by S. Braccini and F. Fidecaro.

The field h^{jk} also determines the energy flux radiated to infinity, as expressed by Eq. (154). Examining the expression for $\tau^{\alpha\beta}$ (Eq. (160)) in the faraway zone, and making use of Eq. (218), it is straightforward to show that

$$\tau_{LL}^{0k} = n^k \tau_{LL}^{00} + O(r^{-3}). \quad (224)$$

This is not surprising, since for a field propagating in a null direction, the energy flux should be the energy density times a unit vector. Substituting this into Eq. (154), with $dS_k = r^2 n_k d\Omega$ and taking the surface to infinity, we obtain

$$\dot{E} = -c \oint_{\infty} (-g) \tau_{LL}^{00} r^2 d\Omega. \quad (225)$$

After some algebra using Eq. (218), it is also possible to show that

$$(-g) \tau_{LL}^{00} = \frac{c^2}{32\pi G} \dot{h}_{TT}^{jk} \dot{h}_{TT}^{jk} + O(r^{-3}). \quad (226)$$

Using the fact that $P_k^i P_j^k = P_j^i$, and $P_i^i = 2$, together with Eq. (220), we obtain

$$\dot{E} = -\frac{G}{2c^5} \mathcal{I}^{lm(3)} \mathcal{I}^{pq(3)} \langle \langle P_{lp} P_{mq} - \frac{1}{2} P_{lm} P_{pq} \rangle \rangle, \quad (227)$$

where $\langle \langle \dots \rangle \rangle$ denote the angular average, defined in Eq. (33). Making use of Eq. (34), we arrive finally at

$$\frac{dE}{dt} = -\frac{G}{5c^5} \mathcal{I}^{(jk)} \mathcal{I}^{(jk)}, \quad (228)$$

which is precisely the same as Eq. (217) for the energy lost as a consequence of radiation reaction forces.

6.2 Energy Flux and Inspiring Compact Binaries

For a binary star system, \mathcal{I}^{jk} is given to lowest order by

$$\mathcal{I}^{jk} = \mu r^j r^k, \quad (229)$$

where $\mu = m_1 m_2 / (m_1 + m_2)$ is the reduced mass, and $\mathbf{r}(t) = \mathbf{x}_1 - \mathbf{x}_2$. Calculating the three time derivatives, and using the Newtonian equation of motion in place of $d\mathbf{v}/dt$, we obtain

$$\frac{dE}{dt} = -\frac{8G^3}{15c^5} \eta^2 \left(\frac{m}{r}\right)^4 (12v^2 - 11\dot{r}^2), \quad (230)$$

where $\eta := \mu/m$. Notice that the fractional change in energy over one orbital period, $\Delta E/E \sim (dE/dt)(P/E)$, can be written, using the fact that $P \sim r^{3/2}/(Gm)^{1/2}$, and that $E \sim \mu v^2 \sim G\mu m/r$, in the schematic form

$$\frac{\Delta E}{E} \sim \frac{r^{3/2}}{(Gm)^{1/2}} \left(\frac{r}{G\mu m}\right) \frac{8G^3}{15c^5} \eta^2 \left(\frac{m}{r}\right)^4 \frac{Gm}{r} \sim \eta \left(\frac{Gm}{c^2 r}\right)^{5/2}, \quad (231)$$

indicating clearly that radiation damping is a 2.5PN-order effect. Using our Keplerian orbit formulae from Sect. 2, and averaging the energy loss over one orbit of the binary system, we obtain the average rate of energy loss,

$$\frac{dE}{dt} = -\frac{32G^4}{5c^5}\eta^2\left(\frac{m}{a}\right)^5 F(e), \quad (232)$$

where a is the semimajor axis, e is the eccentricity, and $F(e)$ is given by

$$F(e) := \frac{1 + \frac{73}{24}e^2 + \frac{37}{96}e^4}{(1 - e^2)^{7/2}}. \quad (233)$$

It is useful to express these results in terms of the dimensionless variable x , given by

$$x := \left(\frac{2\pi}{P} \frac{Gm}{c^3}\right)^{2/3}. \quad (234)$$

The quantity in parentheses is the ratio of the light travel time across a distance given by Gm/c^2 to the orbit period; it is easy to see that $x = Gm/ac^2 \sim (v/c)^2$. In terms of x , we obtain

$$\frac{dE}{dt} = -\frac{32c^5}{5G}\eta^2 x^5 F(e). \quad (235)$$

Since $E \propto a^{-1}$ and $P \propto a^{3/2}$, we have that $2P^{-1}dP/dt = -3E^{-1}dE/dt$. Thus we can obtain an expression for the rate of decrease of the orbital period,

$$\frac{dP}{dt} = -\frac{96}{5}\eta x^{5/2} F(e). \quad (236)$$

This prediction of the damping of an orbit caused by gravitational radiation reaction has been confirmed spectacularly, first using the Hulse–Taylor binary pulsar PSR 1913+16, discovered in 1974, and later in a number of similar binary pulsar systems, notably the “double pulsar,” J0737-3039, discovered in 2003. For more on binary pulsar observations, see the chapter “The Role of Binary Pulsars in Testing Gravity Theories” by A. Possenti and M. Burgay.

The Hulse–Taylor binary and the double-pulsar binary are destined to have their orbits shrink to such a state that the two stars will merge to form a black hole, on a timescale that is short compared to the age of the galaxy. If systems such as this exist in the final stage of inspiral and merger today within several hundred megaparsecs, the waves they emit should be detectable by the advanced ground-based laser interferometers in the LIGO-Virgo-Geo network. In a similar vein, the waves emitted during the inspiral and merger of a pair of supermassive black holes should be detectable by a space-based interferometer such as LISA, out to very large redshifts.

However, in order to enhance the detection confidence, and to extract useful astrophysical information from the detected wave trains, it has been shown that predictions for the motion and gravitational wave field will be needed that go far

beyond the lowest, Newtonian, and post-Newtonian approximations that we have discussed in this chapter. With this motivation, many groups have calculated the higher-order corrections, mainly using techniques based on the post-Minkowskian formalism presented here. The equations of motion for spinless binaries are now known through 3.5PN order, and the effects of spin have also been calculated to rather high post-Newtonian order. Similarly, the gravitational waveform and energy flux have been calculated through 3.5PN order *beyond* the quadrupole formula shown here; again many effects of spin have also been calculated. Work has also begun on the extraordinarily difficult 4PN contributions to both the motion and the waveform. To illustrate what these results look like, we quote here the energy flux, expressed as an expansion in powers of the variable x , for spinless binaries in circular orbits, through 3.5PN order:

$$\begin{aligned} \frac{dE}{dt} = & -\frac{32c^5}{5G}\eta^2x^5\left\{1 - \left(\frac{1247}{336} + \frac{35}{12}\eta\right)x + 4\pi x^{3/2} \right. \\ & - \left(\frac{44711}{9072} - \frac{9271}{504}\eta - \frac{65}{18}\eta^2\right)x^2 - \left(\frac{8191}{672} + \frac{583}{24}\eta\right)\pi x^{5/2} \\ & + \left[\frac{6643739519}{69854400} + \frac{16}{3}\pi^2 - \frac{1712}{105}C - \frac{846}{105}\ln(16x) \right. \\ & - \left(\frac{134543}{7776} - \frac{41}{48}\pi^2\right)\eta - \frac{94403}{3024}\eta^2 - \frac{775}{324}\eta^3\left]x^3 \right. \\ & \left. - \left(\frac{16285}{504} - \frac{214745}{1728}\eta - \frac{193385}{3024}\eta^2\right)\pi x^{7/2} + O(c^{-8})\right\}, \quad (237) \end{aligned}$$

where $C = 0.577\dots$ is the Euler constant. The first term in this sequence is the leading quadrupole term of Eq. (235). Results like these will play a key role in the development of gravitational wave astronomy.

Acknowledgements This work has been supported by the National Science Foundation under grant PHY 09-65133. Much of this work was completed while the author was at Washington University in St. Louis. I am also grateful to the Institut d’Astrophysique de Paris, where part of this work was completed. This chapter is based in part on the forthcoming book *Gravitation: Newtonian, post-Newtonian, Relativistic*, by Eric Poisson and Clifford Will. I am grateful to my coauthor and our publisher Cambridge University Press for permission to use parts of that text in the present chapter.

References

1. Blanchet L. Gravitational radiation from post-Newtonian sources and inspiralling compact binaries. Living Rev Relativ. 2014;17:2 (cited on 1 July 2014): <http://www.livingreviews.org/lrr-2014-2>.
2. Merkowitz SM. Tests of gravity using lunar laser ranging. Living Rev Relativ. 2010;13:7. (cited on 1 July 2012): <http://www.livingreviews.org/lrr-2010-7>.

3. Poisson E, Will CM. Gravity: Newtonian, post-Newtonian, relativistic. Cambridge: Cambridge University Press; in 2014.
4. Sathyaprakash BS, Schutz BF. Physics, astrophysics and cosmology with gravitational waves. *Living Rev Relativ.* 2009;12:2. (cited on 1 July 2012): <http://www.livingreviews.org/lrr-2009-2>.
5. Soffel MH. Relativity in astrometry, celestial mechanics and geodesy. Berlin: Springer-Verlag; 1989.
6. Stairs IH. Testing general relativity with pulsar timing. *Living Rev Relativ.* 2003;6:5. (cited on 1 July 2012): <http://www.livingreviews.org/lrr-2003-5>.
7. Turyshev SG. Experimental tests of general relativity. *Ann Rev Nucl Particle Sci.* 2008;58:207. (arXiv:0806.1731).
8. Will CM. Theory and experiment in gravitational physics. Cambridge: Cambridge University Press; 1993.
9. Will CM. The confrontation between general relativity and experiment. *Living Rev Relativ.* 2014;17:4 (cited on 1 July 2014): <http://www.livingreviews.org/lrr-2014-4>
10. Will CM. Resource letter PTG-1: precision tests of gravity. *Am J Phys.* 2010;78:1240. (arXiv:1008.0296).
11. Will CM. On the unreasonable effectiveness of the post-Newtonian approximation in gravitational physics. *Proc Natl Acad Sci U S A.* 2011;108:5938. (arXiv:1102.5192).

Gravity: Where Do We Stand?

Peron, R.; Colpi, M.; Gorini, V.; Moschella, U. (Eds.)

2016, XIV, 484 p. 130 illus., 103 illus. in color.,

Hardcover

ISBN: 978-3-319-20223-5